Valter Moretti
Dipartimento di Matematica
Facoltà di Scienze M.F.N
Università di Trento

FISICA MATEMATICA I: Elementi di Meccanica Razionale, Meccanica Analitica e Teoria della Stabilità

Corsi di *Fisica Matematica I* per la Laurea Triennale in Matematica e
Meccanica Analitica I per la Laurea Triennale in Fisica.
Facoltà di Scienze MFN, Università di Trento,
anno accademico 2006-2007

Indice

Scopi, prerequisiti matematici. 2
Inquadramento generale della meccanica (analitica) classica. 2

1 Lo Spaziotempo della Fisica Classica e Cinematica. 9
1.1 Lo spaziotempo della fisica classica: Tempo e Spazio assoluti e linee di universo. 9
1.2 Sistemi di riferimento. 13
1.2.1 Sistemi di coordinate solidali. 15
1.3 Cinematica assoluta del punto materiale. 17
1.3.1 Derivazione di curve in spazi affini. 17
1.3.2 Grandezze cinematiche. 18
1.3.3 Cinematica per punti materiali vincolati a curve e superfici ferme. 20
1.4 Cinematica relativa del punto materiale. 26
1.4.1 Vettore $\boldsymbol{\omega}$ e formule di Poisson. 27
1.4.2 Velocità ed accelerazione al variare del riferimento. 35

2 Dinamica del punto e dei sistemi di punti materiali. 38
2.1 Primo principio della dinamica. 38
2.1.1 Sistemi di riferimento inerziali. 38
2.1.2 Trasformazioni di Galileo. 40
2.1.3 Moto relativo di riferimenti inerziali. 43
2.2 Formulazione generale della dinamica classica dei sistemi di punti materiali. . . . 45
2.2.1 Masse, Impulsi e Forze. 45
2.2.2 Sovrapposizione delle forze. 48
2.2.3 Problema fondamentale della dinamica e determinismo. 49
2.3 Situazioni dinamiche più generali. 51
2.3.1 Moto assegnato per un sottosistema: forze dipendenti dal tempo. 51
2.3.2 Vincoli geometrici: reazioni vincolari. 52
2.3.3 Dinamica in riferimenti non inerziali: forze inerziali. 57
2.4 Alcuni commenti sulla formulazione generale sulla dinamica newtoniana 61
2.4.1 Invarianza galileiana della meccanica classica. 61
2.4.2 *Un commento sul cosiddetto “Principio di Mach”. 65

3 Introduzione alla teoria delle equazioni differenziali ordinarie **66**
3.1 Sistemi di equazioni differenziali . . . 66
3.1.1 Riduzione al prim'ordine. . . . 67
3.1.2 Problema di Cauchy. . . . 68
3.1.3 Integrali primi. . . . 69
3.2 Alcune nozioni e risultati preparatori per il teoremi di esistenza e unicità . . . 69
3.2.1 Lo spazio di Banach $C^0(K;\mathbb{K}^n)$. . . . 69
3.2.2 Teorema del punto fisso in spazi metrici completi. . . . 72
3.2.3 Funzioni lipschitziane. . . . 74
3.3 Teoremi di esistenza ed unicità per il problema di Cauchy. . . . 76
3.3.1 Teorema di esistenza ed unicità locale per il problema di Cauchy. . . . 77
3.3.2 Condizione per gli integrali primi. . . . 80
3.3.3 Teorema di esistenza ed unicità globale per il problema di Cauchy. . . . 80
3.3.4 Completezza di soluzioni massimali. . . . 92
3.4 *Confronto tra equazioni differenziali, dipendenza dalle condizioni iniziali e da parametri. . . . 96
3.4.1 Lemma di Gronwall e sue conseguenze. . . . 96
3.4.2 Regolarità della dipendenza dai dati di Cauchy e questioni connesse. . . . 98
3.5 *Problema di Cauchy su varietà differenziabili. . . . 99
3.5.1 Problema di Cauchy, esistenza ed unicità globali. . . . 100
3.5.2 Completezza delle soluzioni. . . . 101
3.5.3 Gruppi di diffeomorfismi locali ad un parametro. . . . 103
3.5.4 Esistenza di integrali primi funzionalmente indipendenti. . . . 105

4 Leggi di bilancio ed integrali primi in Meccanica. **108**
4.1 Equazioni cardinali per i sistemi di punti materiali, conservazione dell'impulso e del momento angolare . . . 108
4.1.1 Massa totale, impulso totale, momento angolare totale, energia cinetica totale . . . 109
4.1.2 Equazioni cardinali. . . . 111
4.1.3 Leggi di bilancio/conservazione di impulso e momento angolare. . . . 114
4.2 Energia meccanica . . . 115
4.2.1 Teorema delle forze vive. . . . 117
4.2.2 Forze conservative. . . . 118
4.2.3 Bilancio e conservazione dell'energia meccanica. . . . 120
4.3 *La necessità della descrizione in termini di continui e di campi in meccanica classica. . . . 122

5 Introduzione alla teoria della stabilità . **125**
5.1 Punti singolari e configurazioni di equilibrio . . . 125
5.1.1 Equilibrio stabile ed instabile. . . . 127
5.1.2 Introduzione ai metodi di Liapunov per lo studio della stabilità . . . 130

5.1.3 *Ancora sulla stabilità asintotica. 135
5.1.4 Un criterio per l'instabilità basato sulla procedura di linearizzazione. . . . 137
5.2 Applicazioni a sistemi di punti materiali con forze conservative. 140
5.2.1 Il teorema di Lagrange-Dirichlet. 141
5.2.2 Un criterio per l'instabilità 145

6 Fondamenti di Meccanica Lagrangiana. 149
6.1 Sistemi vincolati ed equazioni di Eulero-Lagrange. 150
6.1.1 Spaziotempo delle configurazioni in presenza di vincoli olonomi. 150
6.1.2 Spostamenti virtuali e vincoli ideali. 155
6.1.3 Grandezze cinematiche ed energia cinetica. 159
6.2 Equazioni di Eulero-Lagrange e loro proprietà elementari. 160
6.2.1 Normalità delle equazioni di Eulero-Lagrange. 164
6.2.2 Spaziotempo degli atti di moto ed invarianza delle equazioni di Eulero-Lagrange. 166
6.2.3 Lagrangiane. 170
6.2.4 Cambiamento di riferimento inerziale e non unicità della lagrangiana. . . . 174
6.3 *Formulazione geometrico differenziale globale delle equazioni di Eulero-Lagrange. 178
6.3.1 La struttura di varietà fibrata di $\mathbb{V}^{n+1}$ e di $j^1(\mathbb{V}^{n+1})$. 178
6.3.2 Il campo vettoriale dinamico Z. 180

7 Alcuni argomenti più avanzati di Meccanica Lagrangiana. 182
7.1 Il cosiddetto "Principio di Minima Azione" per sistemi che ammettono lagrangiana. 182
7.1.1 Primi rudimenti di calcolo delle variazioni. 182
7.1.2 Il principio di minima azione. 184
7.2 I potenziali generalizzati. 187
7.2.1 Il caso della forza di Lorentz. 187
7.2.2 Generalizzazione della nozione di potenziale. 189
7.2.3 Condizioni per l'esistenza del potenziale generalizzato. 191
7.2.4 Potenziali generalizzati delle forze inerziali. 193
7.3 Configurazioni di equilibrio e stabilità 196
7.3.1 Configurazioni di equilibrio rispetto ad un riferimento. 196
7.3.2 Equilibrio stabile ed instabile, teorema di Lagrange-Dirichlet. 202

8 Simmetrie e leggi di conservazione: teoremi di Nöther e di Jacobi. 206
8.1 Il legame tra simmetria e leggi di conservazione: coordinate cicliche. 206
8.1.1 Coordinate cicliche e conservazione dei momenti coniugati. 206
8.1.2 Invarianza traslazionale e conservazione dell'impulso. 209
8.1.3 Invarianza rotazionale e conservazione del momento angolare. 210
8.2 Il legame tra simmetrie e leggi di conservazione: il teorema di Emmy Nöther. . . 212
8.2.1 Trasformazioni su $j^1(\mathbb{V}^{n+1})$. 213
8.2.2 Il teorema di Nöther in forma locale elementare. 214

8.2.3 Invarianza dell'integrale primo di Nöther per trasformazione di coordinate. 219
8.2.4 *Le trasformazioni di simmetria (debole) di un sistema lagrangiano trasformano soluzioni delle equazioni di E.-L. in soluzioni delle stesse. 220
8.3 L'integrale primo di Jacobi, invarianza sotto "traslazioni temporali" e conservazione dell'energia meccanica. 222
8.4 Commenti finali sul teorema di Nöther. 225
8.4.1 Invarianza sotto il gruppo di Galileo in meccanica lagrangiana. 225
8.4.2 Formulazione lagrangiana e teorema di Nöther oltre la meccanica classica. 225
8.5 *Formulazione generale e globale del Teorema di Nöther. 226
8.5.1 Il teorema di Nöther nella forma generale. 229
8.5.2 Il vettore di Runge-Lenz dal teorema di Nöther. 236
8.5.3 L'integrale primo di Jacobi come conseguenza del teorema di Nöther. . . . 238

9 Fondamenti di Meccanica Hamiltoniana. 239
9.1 Lo spaziotempo delle fasi e le equazioni di Hamilton. 239
9.1.1 Lo spaziotempo delle Fasi $F(\mathbb{V}^{n+1})$. 241
9.1.2 Le equazioni di Hamilton. 244
9.2 Sistemi hamiltoniani su $\mathbb{R} \times \mathbb{R}^{2n}$. 253
9.2.1 Sistemi hamiltoniani su $\mathbb{R} \times \mathbb{R}^{2n}$. 253
9.2.2 Il gruppo simplettico ed i sistemi hamiltoniani. 254
9.2.3 Il teorema di Liouville su $\mathbb{R} \times \mathbb{R}^{2n}$. 259
9.3 *La struttura di varietà fibrata di $F(\mathbb{V}^{n+1})$ e le equazioni di Hamilton come equazioni globali. 263
9.3.1 Lo spazio fibrato $F(\mathbb{V}^{n+1})$. 263
9.3.2 Trasformazione di Legendre globale come diffeomorfismo da $j^1(\mathbb{V}^{n+1})$ a $F(\mathbb{V}^{n+1})$. 265
9.3.3 Equazioni di Hamilton assegnate globalmente su $F(\mathbb{V}^{n+1})$ e campo vettoriale dinamico Z. 267

10 Alcuni argomenti più avanzati di Meccanica Hamiltoniana. 269
10.1 Trasformazioni canoniche e loro proprietà fondamentali. 269
10.1.1 Trasformazioni canoniche 270
10.1.2 Preservazione della forma delle equazioni di Hamilton. 272
10.1.3 *Struttura globale delle equazioni di Hamilton in riferimento all'atlante canonico di $F(\mathbb{V}^{n+1})$. 277
10.2 *Il teorema di Liouville in forma globale ed il teorema "del ritorno" di Poincaré. 278
10.2.1 Teorema di Liouville e l'equazione di Liouville. 278
10.2.2 Il teorema "del ritorno" (o "di ricorrenza") di Poincaré 281
10.3 Simmetrie e leggi di conservazione in meccanica di Hamilton. 285
10.3.1 Parentesi di Poisson. 285
10.3.2 Gruppi locali ad un parametro di trasformazioni canoniche attive. 289
10.3.3 Simmetrie e leggi di conservazione. 293

A Alcune nozioni matematiche rilevanti. 298
A.1 Elementi di geometria differenziale. 299
A.1.1 Varietà differenziabili. 299
A.1.2 Varietà prodotto. 302
A.1.3 Funzioni differenziabili. 302
A.1.4 Sottovarietà embedded. 303
A.1.5 Spazio tangente e cotangente. Campi vettoriali covarianti e controvarianti. 307
A.1.6 Differenziali, curve e vettori tangenti. 310
A.1.7 Pushforward, pullback, derivata di Lie. 311
A.1.8 Immersione di spazi tangenti per sottovarietà embedded. 313
A.1.9 Fibrato tangente e cotangente, varietà fibrate e sezioni. 314
A.2 Elementi di Geometria Affine. 316
A.2.1 Spazi affini. 316
A.2.2 Spazi Euclidei. 321

B Soluzioni e/o suggerimenti per risolvere gli esercizi proposti. 325
B.1 Esercizi del Capitolo 1. 325
B.2 Esercizi del Capitolo 2. 331
B.3 Esercizi del Capitolo 3. 332
B.4 Esercizi del Capitolo 4. 337
B.5 Esercizi del Capitolo 5. 343
B.6 Esercizi del Capitolo 6. 344
B.7 Esercizi del Capitolo 9. 356
B.8 Esercizi dell'Appendice A. 363

C Alcuni esercizi d'esame svolti. 371

Scopi, prerequisiti matematici ed inquadramento generale.

Il fine di queste dispense è quello di introdurre gli studenti dei corsi di laurea in Matematica e in Fisica agli argomenti ed i metodi della fisica matematica classica con particolare attenzione alla meccanica classica, della formulazione lagrangiana della meccanica classica, includendo un'introduzione alla teoria della stabilità ed alla formulazione hamiltonianan della meccanica classica. L'accento è posto in particolare sulla struttura logico-matematica delle teorie fisiche studiate che, per quanto possibile, sono presentate in modo assiomatico deduttivo partendo da un numero ridotto di ipotesi fisiche. Il linguaggio matematico usato è quello dell'analisi e della geometria differenziale elementare.
Queste dispense contengono abbondantemente il materiale didattico dei corsi di *Fisica Matematica I* per la Laurea Triennale in Matematica e di *Meccanica Analitica I* per la Laurea Triennale in Fisica.
I prerequisiti per l'utilizzo delle dispense consistono nel calcolo differenziale ed integrale di una e più variabili, delle nozioni elementari di geometria ed algebra lineare, delle nozioni elementari di topologia generale e dei fondamenti di fisica meccanica. Molte delle nozioni tecniche usate saranno brevemente richiamate prima del loro uso. *Gli esempi (che spesso sono esercizi svolti) e gli esercizi proposti sono parte integrante del corso.*

Le sezioni, i teoremi, le dimostrazioni e gli esercizi contrassegnati con un asterisco non sono strettamente fondamentali ai fini del corso in quanto si riferiscono ad argomenti (specialmente matematici) più avanzati. Sono spesso importanti per chi voglia apparofondire il formalismo.

L'ultima sezione riporta le soluzioni o suggerimenti per le soluzioni degli esercizi proposti. Come testo generale di riferimento, per approfondimenti, complementi ed esercizi, si consigliano i testi [Goldstein50], [Fasano-Marmi] e [Arnold92].

Ringraziamento. Ringrazio l'amico e collega Prof. Enrico Pagani per diversi suggerimenti sui contenuti dei corsi a cui si rivolgono queste dispense. Ringrazio (in ordine alfabetico) Marco Frego, Fabio Zanini e Chiara Zarpellon per avermi segnalato varie sviste in diverse formule presenti nel testo.

Inquadramento generale della meccanica (analitica) classica.

Dal punto di vista fisico è importante sottolineare che la descrizione della realtà fisica presentata in queste dispense ha precisi limiti di applicabilità e, generalmente parlando, deve pensarsi come un'approssimazione di qualche teoria più fondamentale. Infatti essa risulta essere inadeguata in almeno due contesti.

(a) La descrizione classica cessa di valere nel regime di velocità comparabili con quelle della luce/forti campi gravitazionali/trattazioni della fisica cosmologica. In tali contesti la descrizione

più soddisfacente, al momento nota, è data dalla *Teoria della Relatività* [1], di cui la meccanica classica è approssimazione. La rivoluzione della Relatività ha mostrato che le strutture metriche classiche (lunghezze ed intervalli di tempo) sono in realtà *relative al sistema di riferimento*, ma al contempo sono parti di una struttura metrica spaziotemporale *assoluta* che ha particolari proprietà di simmetria (almeno fino a quando si trascura la descrizione relativistica dell'interazione gravitazionale) descritte dal cosiddetto gruppo di Lorentz-Poincaré. La geometria dello spaziotempo che ne consegue si è rivelata il linguaggio matematico per poter trattare argomenti di generale interesse fisico, come la nozione di causalità. Le implicazioni di questo nuovo punto di vista sono state incredibilmente feconde ed hanno avuto influenze fondamentali nello sviluppo di tutta la fisica del 1900. La teoria della relatività ha costruito, insieme alla meccanica quantistica, il *linguaggio stesso* ed il *paradigma* della fisica teorica di un secolo intero di ricerca.

(b) La descrizione classica cessa di essere adeguata anche, rozzamente parlando, in riferimento a sistemi microscopici (scale molecolari ed inferiori). In tali contesti la descrizione più adeguata è data dalla *Meccanica Quantistica* (e dalla teoria dei campi quantistica), di cui, un'altra volta, la meccanica classica è approssimazione. Mentre il linguaggio matematico delle teorie relativistiche è ancora quello della geometria differenziale, il linguaggio matematico delle teorie quantistiche è dato dall'analisi funzionale (degli spazi di Hilbert in particolare).
Dobbiamo doverosamente rimarcare che lo schema è ancora tutt'altro che completo, visto che le teorie quantistiche e quelle relativistiche non formano un corpus coerente, in particolare vi sono diversi problemi concettuali nel conciliare la descrizione quantistica con quella data dalla relatività generale. Al momento manca una descrizione completa della struttura fisica di ciò che esiste.

La meccanica classica, d'altra parte, funziona perfettamente per le applicazioni pratiche più comuni, ma non solo. Basti pensare che le missioni *Apollo* che hanno portato l'uomo sulla luna sono state concepite completamente nell'ambito della meccanica classica, attraverso la quale sono stati costruiti tutti i modelli e sono stati fatti i calcoli.

[1] Problemi irrisolti anche nell'ambito delle Teorie Relativistiche rimangono aperti in cosmologia, in particolare in relazione al cosiddetto problema della *materia oscura*.

Capitolo 1

Lo Spaziotempo della Fisica Classica e Cinematica.

La Matematica in Fisica è come il maiale: non si butta mai via niente.

In questo capitolo introduciamo la struttura dello spaziotempo della fisica classica, la nozione di sistema di riferimento e la cinematica assoluta e relativa elementare. L'appendice A richiama le nozioni di *spazio affine* e *spazio euclideo*, alcune nozioni elementari di *topologia generale* e la nozione di varietà differenziabile dando diversi esempi ed esercizi svolti. In particolare la definizione A.1 richiama la nozione di funzione differenziabile di classe C^k.

1.1 Lo spaziotempo della fisica classica: Tempo e Spazio assoluti e linee di universo.

Un postulato fisico fondamentale, comune alle teorie fisiche classiche e relativistiche, è quello che afferma che tutto ciò che accade sia decomponibile in *eventi*. Un evento, dal punto di vista fisico corrisponde alla minima determinazione spaziotemporale possibile, individuabile dall'assegnazione di tre coordinate spaziali ed una temporale. Tali assegnazioni sono relative ai diversi osservatori, intesi qui come puri *sistemi di riferimento* (non è richiesta alcuna attività cosciente!). L'insieme degli eventi costituisce lo *spaziotempo*. Deve essere chiaro da subito che ogni evento e lo spazio tempo stesso hanno comunque una natura indipendente e pre-esistente alle loro rappresentazioni, in termini di coordinate spaziotemporali, date nei singoli sistemi di riferimento. In questo *framework*, tutto ciò che accade deve ammettere una descrizione in termini di *relazioni* o *coincidenze* tra eventi.
Si assume ulteriormente che lo spaziotempo abbia una natura *topologica* – tecnicamente la natura di *spazio topologico di Hausdorff a base numerabile* – ed una soprastante natura *differenziabile*

– tecnicamente ciò corrisponde alla presenza di una struttura di *varietà differenziabile quadridimensionale*, sostanzialmente una classe di sistemi di coordinate nell'intorno di ogni evento che identificano tale intorno con un corrispondente intorno di $\mathbb{R}^4$ –. La struttura topologica permette in particolare (ma non solo) di dare senso alle nozioni di "vicinanza o prossimità spaziale e temporale", la struttura differenziabile permette di dare senso alla nozione di curva e funzione differenziabile (rispetto alle coordinate) definite nello spaziotempo. Come vedremo tra poco, si possono in tal modo introdurre le nozioni di velocità ed accelerazione, per descrivere l'evoluzione di punti materiali. La struttura differenziabile dello spaziotempo permette, molto più in generale, di descrivere in termini di *equazioni differenziali* le equazioni che determinano l'evoluzione spaziotemporale di sistemi fisici classici (e relativistici), siano essi discreti (sistemi di punti) o continui (fluidi o campi). L'esistenza e l'unicità delle soluzioni di tali equazioni (in presenza di condizioni iniziali e/o al contorno) corrisponde al postulato fisico classico del *determinismo*. I teoremi di esistenza ed unicità sono strettamente connessi alla struttura topologica (non solo differenziabile) dello spaziotempo.
Considerando esplicitamente il caso dello spaziotempo classico, lo spaziotempo include due strutture ulteriori: il *tempo assoluto* e lo *spazio assoluto*. Il termine *assoluto* si riferisce all'indipendenza dai possibili sistemi di riferimento delle misure di angoli, distanze ed intervalli temporali.

Definizione 1.1. (**Lo spaziotempo della fisica classica**.) Lo **spaziotempo della fisica classica** è una varietà differenziabile quadridimensionale (di classe C^∞) indicata con $\mathbb{V}^4$. I punti di $\mathbb{V}^4$ sono detti **eventi**.
$\mathbb{V}^4$ è dotato di una funzione privilegiata definita a meno di una arbitraria costante additiva, $T : \mathbb{V}^4 \to \mathbb{R}$, detta **tempo assoluto** che si richiede essere differenziabile, suriettiva e ovunque nonsingolare (definizione A.6). Si suppone inoltre che:

(i) ciascuno dei sottoinsiemi a due a due disgiunti: $\Sigma_t := \big\{p \in \mathbb{V}^4 \mid T(p) = t\big\}$, detta **spazio assoluto al tempo** $t \in \mathbb{R}$, abbia una struttura di spazio euclideo tridimensionale (con spazio delle traslazioni V_t e prodotto scalare $(\cdot|\cdot)_t$) compatibile con la struttura globale di $\mathbb{V}^4$.

(ii) Se $d_t : \Sigma_t \times \Sigma_t \to \mathbb{R}$, detta (funzione) **distanza assoluta al tempo** t, è la distanza su Σ_t associata alla struttura di spazio euclideo, la funzione $\mathbb{R} \times \Sigma_t \times \Sigma_t \ni (t, P, Q) \mapsto d_t(P, Q)$ deve essere ovunque continua, ed anche differenziabile per $P \neq Q$. $\diamondsuit$

Dal punto di vosta fisico è importante notare che attualmente si specula sulla eventuale natura non continua dello spaziotempo stesso a scale molto piccole (scale di Planck $\sim 10^{-33}$cm e 10^{-43}s) in cui dovrebbe valere una qualche forma di *Quantum Gravity*. A tali scale la struttura topologico-differenziabile classica descritta nella definizione 1.1 cesserebbe di essere fisicamente appropriata.
Tornando alla descrizione classica basata sulla definizione 1.1 possiamo dire che, dal punto di vista fisico-operativo, un **punto materiale** è un sistema fisico la cui posizione è determinata, istante per istante, assegnando solo tre coordinate spaziali. Il fatto che un sistema fisico sia o no assimilabile ad un punto materiale, trascurandone l'eventuale struttura "interna", può dipendere dalla precisione dei nostri strumenti di misura e dal grado di approssimazione scelto. L'evoluzione temporale di un punto materiale nello spaziotempo costituisce un oggetto elementare detto *storia*

o *linea di universo* (*di un punto materiale*). Formalmente possiamo dare la seguente definizione.

Definizione 1.2. (**Linea di universo o storia**.) Una **storia** o **linea di universo** (**di un punto materiale**) è una curva differenziabile

$$I \ni t \mapsto \gamma(t) \in \mathbb{V}^4 \,,$$

dove I è un intervallo aperto di $\mathbb{R}$ che può essere identificato con un intervallo di tempo assoluto, nel senso che vale, *se fissata opportunamente la costante additiva della definizione di* T

$$T(\gamma(t)) = t \,,\ \forall t \in I. \tag{1.1}$$

$\diamondsuit$

Osservazioni 1.1.
(1) Si devono notare alcuni fatti su T. In primo luogo T è definita a meno di una costante additiva, ció corrisponde al fatto fisico evidente di poter fissare l'origine convenzionale del tempo a nostro piacimento. D'altra parte il fatto che T non sia definita a meno di un fattore di scala significa che è stato fatta una scelta universale dell'unità di misura del tempo. Dal punto di vista fisico l'assunzione dell'esistenza di un tempo assoluto implica che si è anche assunta l'esistenza di *orologi ideali* disponibili in ogni punto dello spaziotempo che misurino tale tempo.
(2)* Le Σ_t sono sottovarietà embedded di dimensione $4-1=3$ per il teorema dei valori regolari, teorema A.2 nell'appendice, essendo T ovunque non singolare. Riguardo alla struttura di spazio euclideo delle Σ_t, l'affermazione "compatibile con la struttura globale di $\mathbb{V}^4$" significa che la struttura differenziabile associata alla struttura di spazio affine è quella dovuta al fatto che Σ_t è sottovarietà embedded coincidono.
(3) Per costruzione valgono i fatti seguenti riguardati gli spazi assoluti.
(a) La suriettività di T assicura che $\Sigma_t \neq \emptyset$ per ogni $t \in \mathbb{R}$.
(b) Per definizione, se $t \neq t'$, gli spazi assoluti associati sono disgiunti:

$$\Sigma_t \cap \Sigma_{t'} = \emptyset \,.$$

(c) Dato che T è definito su tutto lo spaziotempo, per ogni $p \in \mathbb{V}^4$ esiste $t \in \mathbb{R}$ con $\Sigma_t \ni p$. Pertanto

$$\mathbb{V}^4 = \cup_{t\in\mathbb{R}} \Sigma_t \,.$$

In definitiva lo spaziotempo è fogliato dagli spazi assoluti e coincide con l'unione di essi.
(4) Ogni spazio assoluto Σ_t è dunque uno spazio euclideo tridimensionale reale. Dal punto di vista fisico, la distanza d_t, e quindi i prodotti scalari e gli angoli tra i segmenti in Σ_t, viene misurata con l'assegnazione di una classe di *regoli rigidi ideali*[1] che si devono supporre disponibili in ogni punto dello spaziotempo. Le traslazioni indotte dai vettori v in Σ_t devono essere

[1] Ci si può chiedere cosa intendiamo, dal punto di vista fisico, quando parliamo della classe dei regoli rigidi *ideali* e della classe degli orologi *ideali*. L'*idealità* dei regoli è relativa alla seguente proprietà della classe di essi. Scelti due regoli arbitrariamente, essi risultano coincidere se sono in quiete nello stesso posto in un arbitrario

pensate come le traslazioni fisiche dei corpi materiali. Devono quindi esistere corpi (almeno i regoli rigidi!) che siano metricamente invarianti per traslazioni fisiche come lo è d_t.
(5) Abbiamo definito le nozioni di distanza ed intervallo temporale indipendentemente dalla nozione di sistema di riferimento che dobbiamo ancora introdurre. Proprio per tale fatto la nozione di distanza e di lunghezza di un intervallo temporale sono *assoluti*.
(6) Nella definizione di linea di universo, il requisito $T(\gamma(t)) = t$ per ogni $t \in I$ assicura che il tempo assoluto possa essere usato come parametro per descrivere la linea di universo, ma anche che una linea di universo non possa intersecare più di una volta lo stesso spazio assoluto Σ_t per un fissato valore di t[2] (e non possa in particolare autointersecarsi). Come conseguenza del vincolo posto, risulta che il punto materiale, di cui la linea di universo rappresenta la storia, non "può tornare indietro nel tempo" (non può tornare a Σ_t una volta che l'ha lasciata nel passato). Questo è un requisito fisico che evita i paradossi causali della fantascienza.
(7) Possiamo pensare di dotare due punti materiali di orologi ideali. Tali orologi, a meno della scelta dell'origine del tempo, indicheranno le etichette t delle varie ipersuperfici Σ_t attraversate dalle linee di universo dei due punti. Se i due punti si incontrano nell'evento $p \in \Sigma_t$ e in tale evento si decide di sincronizzare gli orologi *a vista* in modo che segnino, in quell'evento, entrambi t, ad ogni successivo incontro dei due orologi, essi segneranno ancora lo stesso tempo $t' > t$ (indipendentemente dal loro stato di moto relativo al loro incontro). Inoltre, se sincronizziamo due orologi a distanza tramite un terzo orologio che funge da spola tra i due, la sincronizzazione rimarrà in futuro e tutte le procedure di sincronizzazione risultano essere equivalenti. Tutto questo perché gli orologi ideali, si limitano a *leggere* il tempo assoluto postulato esistere. Sappiamo dalla fisica che in realtà questo stato di cose è solo approssimato e cessa di essere verificato quando le velocità in gioco (per es. quella relativa tra due orologi istantaneamente vicini) sono prossime a quella della luce. Questo significa che l'assioma che postula il tempo assoluto ha validità fisica limitata e deve essere sostituito da qualcos'altro per avere un'immagine più fedele del mondo fisico. Similmente, se associamo dei regoli rigidi ai punti materiali, ogni qual volta tali regoli sono messi a confronto nello stesso evento p (anche con velocità relativa diversa) devono coincidere, perché si limitano a *leggere* la geometria assoluta dello spazio assoluto $\Sigma_t \ni p$. Anche in questo caso, la realtà fisica è ben diversa quando le velocità in gioco sono vicine a quella della luce e ciò significa che postulare lo spazio, e la geometria assoluti non ritrae fedelmente la realtà fisica.

---

riferimento e che questo fatto permane anche dopo che i regoli hanno subito diverse storie (incluse accelerazioni), una volta riportati in quiete relativa in un arbitrario riferimento (anche diverso dal primo). Lo stesso criterio si applica per la nozione di idealità di orologi: presi due orologi essi risultano battere il tempo nello stesso modo quando sono in quiete nello stesso posto in un riferimento e tale fatto permane anche dopo che gli orologi hanno subito diverse storie, una volta riportati in quiete relativa in un riferimento (anche diverso dal primo e anche se non risultano più essere sincronizzati se lo erano inizialmente). Queste nozioni di *idealità* per regoli ed orologi sono valide anche in fisica relativistica.

[2] Se γ interseca Σ_t per t_1 e t_2, allora $t_1 = T(\gamma(t_1)) = T(\Sigma_t) = T(\gamma(t_2)) = t_2$ per cui $\gamma(t_1) = \gamma(t_2)$.

1.2 Sistemi di riferimento.

Quello che vogliamo fare ora è introdurre gli strumenti necessari per poter assegnare, ad ogni linea di universo, una velocità ed una accelerazione. Per fare ciò dobbiamo introdurre la nozione di *sistema di riferimento*. Tradizionalmente questa parte della meccanica cade sotto il nome di **cinematica**, che è la parte della meccanica che si occupa della pura descrizione dei moti.
Un sistema di riferimento è un modo per rappresentare, in un fissato "spazio di quiete", ciò che accade nello spaziotempo, allo scorrere del tempo assoluto. Lo spazio di quiete deve essere uno spazio euclideo tridimensionale $\mathbb{E}^3$ tale che, dal punto di vista affine e metrico, sia identifica bile con ogni spazio assoluto Σ_t. Inoltre i punti P dello spazio di quiete, evolvendo nel tempo, devono descrivere linee di universo in $\mathbb{V}^4$.
Esiste però un punto di vista alterativi, in un certo senso più fisico, in cui si immagina un sistema di riferimento $\mathscr{I}$ come l'assegnazione di una classe $\{\gamma_P\}_{P\in E_{\mathscr{I}}}$ di linee di universo, parametrizzata su *tutto* il tempo assoluto e *sincronizzate*, cioé le costanti additive del parametro t che le descrive ($\gamma_P = \gamma_P(t)$) sono fissate in modo che *ogni* curva intercetti ogni Σ_τ per il valore $t = \tau$ del suo parametro. Le linee di $\mathscr{I}$ devono permeare *tutto* lo spaziotempo. Si assume che le linee γ_P *non si intersechino* tra di loro. In questo modo, ogni evento è intercettato da una sola linea di universo, l'indice P che etichetta tale linea e la distingue dalle altre, definisce "la localizzazione spaziale" dell'evento. Il valore del parametro, il tempo assoluto in cui γ_P interseca l'evento, definisce la "localizzazione temporale" dell'evento. Lo spazio di quiete di $\mathscr{I}$, in questo contesto, può pensarsi come è l'insieme degli indici P che etichettano le differenti linee di universo di $\mathscr{I}$.
Se le linee di universo evolvono in modo *rigido*, cioé conservando le distanze reciproche negli spazi Σ_t al variare di t, si può dotare lo spazio di quiete $E_{\mathscr{I}}$, come ci si aspetta dalla fisica, di una struttura metrica che lo identifica con ogni spazio Σ_t.
Noi useremo la questa seconda definizione di sistema di riferimento e mostreremo che in realtà è riconducibile alla prima.

Definizione 1.3. (**Sistemi di riferimento**.) Un **sistema di riferimento** o brevemente un **riferimento** (della fisica classica), $\mathscr{I}$ è una classe di linee di universo $\{\gamma_P\}_{P\in E_{\mathscr{I}}}$ che soddisfa i seguenti requisiti.
(i) **inestendibilità temporale**: per ogni $P \in E_{\mathscr{I}}$, il dominio di γ_P è tutto $\mathbb{R}$.
(ii) **sincronizzazione**: $\gamma_P(t) \in \Sigma_t$ per ogni $t \in \mathbb{R}$ e ogni $P \in E_{\mathscr{I}}$ (per una scelta opportuna della costante additiva del tempo assoluto).
(iii) **globalità** : $\bigcup_{P\in E_{\mathscr{I}}} \gamma_P(\mathbb{R}) = \mathbb{V}^4$.
(iv) **non intersecabilità** : $\gamma_P(\mathbb{R}) \cap \gamma_Q(\mathbb{R}) = \emptyset$ se $P \neq Q$ per ogni $P, Q \in E_{\mathscr{I}}$.
(v) **rigidità** : $d_t(\gamma_P(t), \gamma_Q(t)) = d_{t'}(\gamma_P(t'), \gamma_Q(t'))$ per ogni $P, Q \in E_{\mathscr{I}}$ e per ogni $t, t' \in \mathbb{R}$.
L'insieme degli indici $E_{\mathscr{I}}$ è detto **spazio di quiete** di $\mathscr{I}$. $\diamondsuit$

Teorema 1.1. *Sia $\mathscr{I}$ un riferimento su $\mathbb{V}^4$. Vale quanto segue.*
(a) *Si può dotare lo spazio di quiete $E_{\mathscr{I}}$ di struttura di spazio euclideo tridimensionale (con spazio delle traslazioni, prodotto scalare e distanza che indichiamo rispettivamente con $V_{\mathscr{I}}$, $(\cdot|\cdot)_{\mathscr{I}}$ e $d_{\mathscr{I}}$).*

(b) *L'applicazione* $\Pi_{\mathscr{I}} : \mathbb{V}^4 \to E_{\mathscr{I}}$ *che associa* $p \in \mathbb{V}^4$ *all'indice* P *dell'unica linea di universo di* $\mathscr{I}$ *che interseca* p *soddisfa:*

(i) $\Pi_{\mathscr{I}} \restriction_{\Sigma_t} : \Sigma_t \to E_{\mathscr{I}}$ *è un'isometria (e quindi un'isometria affine) per ogni* $t \in \mathbb{R}$

(ii) le curve $\mathbb{R} \ni t \mapsto (\Pi_{\mathscr{I}} \restriction_{\Sigma_t})^{-1}(P)$ *sono differenziabili e sono le linee di universo di* $\mathscr{I}$.

***Dimostrazione**. (a) L'idea generale è semplice ed è la seguente.
Vogliamo mettere sullo spazio delle curve $\{\gamma_P\}_{P \in E_{\mathscr{I}}}$ *una struttura di spazio euclideo in modo tale che, per ogni fissato* t*, la struttura di spazio euclideo che eredita lo spazio* $\Sigma_t = \{\gamma_P(t)\}_{P \in E_{\mathscr{I}}}$ *coincida con la preesistente struttura euclidea di* Σ_t.
La possibilità di definire tale struttura è dovuta principalmente al seguente fatto. Se $R \in \Sigma_t$ indichiamo con $\gamma_{P(t,R)}$ l'unica curva del riferimento $\mathscr{I}$ che passa per R. Per $t' \in \mathbb{R}$ fissato, l'applicazione

$$\Sigma_t \ni R \mapsto \Phi_{t't}(R) := \gamma_{P(t,R)}(t')$$

è un'isometria tra gli spazi euclidei Σ_t e $\Sigma_{t'}$ a causa della richiesta di rigidità (v) in definizione 1.3. $d\Phi_{t't} : \gamma_P(t) - \gamma_0(t) \mapsto \gamma_P(t') - \gamma_0(t')$ è quindi un isomorfismo di spazi vettoriali che conserva il prodotto scalare, essendo $\Phi_{t't}$ una isometria affine ((iv) in esercizio A.5.**4**). Useremo tali fatti tra poco.
Per prima cosa determiniamo lo spazio delle traslazioni $V_{\mathscr{I}}$ di $E_{\mathscr{I}}$. Fissati $t \in \mathbb{R}$ e $O \in E_{\mathscr{I}}$, al variare di $P \in \Sigma_t$, $(P - O)(t) := \gamma_P(t) - \gamma_O(t) \in V_t$ individua biunivocamente tutti i vettori dello spazio delle traslazioni V_t di Σ_t. Con la scelta fatta di O, se P è tenuto fisso, al variare di $t \in \mathbb{R}$ abbiamo una *curva di vettori*. Indichiamo con $P - O$ semplicemente tale curva. Notare che, per valori differenti di t, i vettori $(P - O)(t)$ appartengono a spazi vettoriali differenti. Possiamo dotare l'insieme delle curve di vettori $\{P - O\}_{P \in E_{\mathscr{I}}}$ di una struttura di spazio vettoriale, inducendola da quella di *tutti* gli spazi V_t, richiedendo che valga, per $P, Q \in E_{\mathscr{I}}$, $\alpha \in \mathbb{R}$ ed al variare di $t \in \mathbb{R}$,

$$((P - O) + (Q - O))(t) := (\gamma_P(t) - \gamma_O(t)) + (\gamma_Q(t) - \gamma_O(t)) \text{ e } (\alpha(P - O))(t) := \alpha(\gamma_P(t) - \gamma_O(t)).$$

Lasciamo al lettore il compito di provare che gli assiomi di spazio vettoriale sono soddisfatti in questo modo. Per provarlo, bisogna usare il fatto che le funzioni $d\Phi_{t't}$, che connettono V_t a $V_{t'}$ muovendosi lungo ciascuna curva di vettori, sono isomorfismi di spazi vettoriali[3]. Indichiamo con $V_{\mathscr{I}}$ lo spazio vettoriale ottenuto in questo modo, che per costruzione è isomorfo a ogni V_t tramite la funzione $P - O \mapsto \gamma_P(t) - \gamma_O(t)$. Possiamo dotare $V_{\mathscr{I}}$ di prodotto scalare definendolo come[4]

$$(P - O | Q - O)_{\mathscr{I}} := (\gamma_P(t) - \gamma_O(t) \cdot \gamma_Q(t) - \gamma_O(t))_t \,. \tag{1.2}$$

Notare che, affinché questa definizione abbia senso il secondo membro non deve dipendere da t Dato che $d\Phi_{t't}$ è un isomorfismo di spazi vettoriali che conserva il prodotto scalare, la struttura

[3] Si osservi per esempio che, se le definizioni di somma e prodotto per curve di vettori data sono ben poste, deve accadere che se $\alpha(P - O) + \beta(Q - O) = \delta(R - O)$ allora deve essere $\alpha(\gamma_P(t) - \gamma_O(t)) = \beta(\gamma_Q(t) - \gamma_O(t)) + \delta(\gamma_Q(t) - \gamma_O(t))$ per *ogni* $t \in \mathbb{R}$, e questo è in generale falso se le $d\Phi_{t't}$ non sono isomorfismi.

[4] attenzione alla notazione: $(P - O | Q - O)$ è il prodotto scalare tra il vettore $P - O$ ed il vettore $Q - O$, prima si esegue la differenza tra punti e poi si esegue il prodotto scalare, non aggiungiamo altre parentesi per non appesantire la notazione

di spazio vettoriale con prodotto scalare introdotta sopra in questo modo è ben definita. Per introdurre la struttura affine di $E_{\mathscr{I}}$ è necessario assegnare l'applicazione che associa a punti $P, Q \in E_{\mathscr{I}}$ un vettore di $V_{\mathscr{I}}$ che, ovviamente sarà indicato con $P - Q$. Il vettore $R - Q \in V_{\mathscr{I}}$ è definito come $(R - Q) := (R - O) + (O - Q)$, avendo già definito sopra $P - O$. Si noti che la definizione non dipende dalla scelta di O, in quanto, per ogni fissato t, vale:

$$(R-O)(t)+(O-Q)(t) = (\gamma_R(t)-\gamma_O(t))+(\gamma_O(t)-\gamma_Q(t)) = (\gamma_R(t)-\gamma_{O'}(t))+(\gamma_{O'}(t)-\gamma_Q(t)) .$$

e l'ultimo membro è uguale a $(R - O')(t) + (O' - Q)(t)$, da cui, per definizione

$$(R - O) + (O - Q) = (R - O') + (O' - Q) .$$

Si verifica facilmente che, con questa definizione, $E_{\mathscr{I}}$ è uno spazio affine con spazio delle traslazioni $V_{\mathscr{I}}$. Inoltre, per costruzione e per ogni fissato $t \in \mathbb{R}$, $E_{\mathscr{I}} \ni P \mapsto \gamma_P(t) \in \Sigma_t$ è un isomorfismo di spazi affini (una trasformazione affine biettiva). La distanza associata al prodotto scalare assegnato su $V_{\mathscr{I}}$ risulta soddisfare per costruzione, per ogni fissato t:

$$d_{\mathscr{I}}(P, Q) := d_t(\gamma_P(t), \gamma_Q(t)) \text{ , con } P, Q \in \mathbb{E}^3_{\mathscr{I}} \text{ e } t \in \mathbb{R}. \tag{1.3}$$

La costruzione della struttura euclidea di $E_{\mathscr{I}}$ data sopra implica immediatamente che, per un fissato t, la funzione $\Phi_t : E_{\mathscr{I}} \ni P \mapsto \gamma_P(t) \in \Sigma_t$ sia, in conclusione, un'isometria di spazi affini.
(b) (i) Per definizione di $\Pi_{\mathscr{I}}$, Φ_t coincide con l'inversa di $\Pi_{\mathscr{I}} \restriction_{\Sigma_t}$ che è pertanto isometria di spazi affini. La proprietà (ii) è ovvia per definizione di $\Pi_{\mathscr{I}}$. $\square$

1.2.1 Sistemi di coordinate solidali.

Vogliamo ora introdurre particolari carte globali su $\mathbb{V}^4$, associate ad un riferimento fissato $\mathscr{I}$, che diremo *solidali* con $\mathscr{I}$. A tal fine fissiamo, nello spazio di quiete $E_{\mathscr{I}}$, un sistema di coordinate cartesiane ortonormali x^1, x^2, x^3. Consideriamo poi la funzione $\sigma_{\mathscr{I}} : \mathbb{V}^4 \to \mathbb{R}^4$, che dopo aver proiettato ogni evento $p \in \Sigma_t$ dello spaziotempo sullo spazio di quiete di $\mathscr{I}$, $E_{\mathscr{I}}$, associa a p la quaterna ordinata di numeri dati dal valore del tempo assoluto $T(p) = t$ e dalle tre coordinate cartesiane di $\Pi_{\mathscr{I}}(p)$ nello spazio di quiete di $\mathscr{I}$:

$$\sigma_{\mathscr{I}} : p \mapsto \left(T(p), x^1\left(\Pi_{\mathscr{I}}(p)\right), x^2\left(\Pi_{\mathscr{I}}(p)\right), x^3\left(\Pi_{\mathscr{I}}(p)\right)\right) .$$

Mostriamo che $\sigma_{\mathscr{I}}$ è biettiva e quindi, in particolare, definisce una carta globale $(\mathbb{V}^4, \sigma_{\mathscr{I}})$ e di conseguenza un atlante su $\mathbb{V}^4$.

Suriettività. Scelto $(t, x^1, x^2, x^3) \in \mathbb{R}^4$ si consideri $\Pi_{\mathscr{I}} \restriction_{\Sigma_t} : \Sigma_t \to E_{\mathscr{I}}$. Questa è una funzione biettiva e quindi suriettiva in quanto affine (esercizio A.5.**4**). Quindi esiste $p \in \Sigma_t$ tale che $\Pi_{\mathscr{I}}(p) \in E_{\mathscr{I}}$ ha coordinate (x^1, x^2, x^3). Per costruzione $\sigma_{\mathscr{I}}(p) = (t, x^1, x^2, x^3)$.

Iniettività. Se $p \neq p'$ ci sono due possibilità. (1) $T(p) \neq T(p')$ per cui $\sigma_{\mathscr{I}}(p) \neq \sigma_{\mathscr{I}}(p')$. (2) $T(p) = T(p') =: t$, per cui $p, p' \in \Sigma_t$, ma $d_t(p, p') \neq 0$ perché sono punti distinti. In tal caso, essendo $\Pi_{\mathscr{I}} \restriction_{\Sigma_t}$ un'isometria sarà $\Pi_{\mathscr{I}}(p) \neq \Pi_{\mathscr{I}}(p')$ e quindi

$$\left(x^1\left(\Pi_{\mathscr{I}}(p)\right), x^2\left(\Pi_{\mathscr{I}}(p)\right), x^3\left(\Pi_{\mathscr{I}}(p)\right)\right) \neq \left(x^1\left(\Pi_{\mathscr{I}}(p')\right), x^2\left(\Pi_{\mathscr{I}}(p')\right), x^3\left(\Pi_{\mathscr{I}}(p')\right)\right) ,$$

da cui $\sigma_{\mathscr{I}}(p) \neq \sigma_{\mathscr{I}}(p')$.

Definizione 1.4. (**Coordinate cartesiane ortonormali solidali**.) Sia $\mathscr{I}$ un sistema di riferimento nello spaziotempo $\mathbb{V}^4$, siano x^1, x^2, x^3 coordinate cartesiane ortonormali nello spazio di quiete $E_{\mathscr{I}}$ e $\sigma_{\mathscr{I}} : \mathbb{V}^4 \to \mathbb{R}^4$ sia la funzione biettiva che associa ad ogni evento p dello spaziotempo i numeri $(T(p), x^1(\Pi_{\mathscr{I}}(p)), x^2(\Pi_{\mathscr{I}}(p)), x^3(\Pi_{\mathscr{I}}(p)))$. La carta globale $(\mathbb{V}^4, \sigma_{\mathscr{I}})$ è detta **sistema di coordinate cartesiane ortonormali solidali** con $\mathscr{I}$.

Osservazioni 1.2.
(1) Si noti che ci sono infiniti sistemi di coordinate cartesiane ortonormali solidali con un fissato riferimento $\mathscr{I}$. Questa infinità è parametrizzata dalla scelta dell'origine $O \in E_{\mathscr{I}}$, dalla scelta della base ortonormale in $V_{\mathscr{I}}$, ma anche della scelta dalla costante indeterminata nella definizione del tempo assoluto.
(2) Dal punto di vista fisico, le coordinate solidali con i riferimenti sono le coordinate naturali dello spaziotempo. Lo spaziotempo è assunto essere una varietà proprio per l'esistenza di diversi sistemi di coordinate, associate ai diversi riferimenti, che lo identificano con $\mathbb{R}^4$. Ci si deve quindi chiedere, a questo punto, se la struttura differenziabile indotta su $\mathbb{V}^4$ dai sistemi di coordinate solidali con i riferimenti coincida davvero con la struttura differenziabile che abbiamo a priori assunto esistere su $\mathbb{V}^4$ nella definizione 1.1. La risposta, come deve essere, è positiva in quanto vale la seguente proposizione tecnica, basata sul fatto che la funzione tempo assoluto è non singolare. La dimostrazione di essa è fornita negli esercizi.

Proposizione 1.1. *Ogni sistema di coordinate cartesiane ortonormali solidali con un arbitrariamente assegnato riferimento $\mathscr{I}$ di $\mathbb{V}^4$ appartiene alla struttura differenziabile di $\mathbb{V}^4$. Di conseguenza valgono i seguenti fatti:*
(a) *la struttura differenziabile di $\mathbb{V}^4$ coincide con quella indotta da ogni sistema di coordinate cartesiane ortonormali solidale con ogni fissato riferimento;*
(b) *per ogni fissato riferimento $\mathscr{I}$, la funzione $\Pi_{\mathscr{I}}$ è differenziabile.*

Esercizi 1.1.
1.* Dimostrare la proposizione 1.1.
2. Dimostrare che, dati due riferimenti $\mathscr{I}$ e $\mathscr{I}'$ con associate coordinate cartesiane ortonormali solidali, rispettivamente (t, x^1, x^2, x^3) e (t', x'^1, x'^2, x'^3), la funzione $\sigma_{\mathscr{I}'} \circ \sigma_{\mathscr{I}}^{-1} : \mathbb{R}^4 \to \mathbb{R}^4$, ovvero la "legge di trasformazione delle coordinate non primate nelle coordinate primate", ha la forma

$$\begin{cases} t' = & t + c\,, \\ x'^i = & \displaystyle\sum_{j=1}^{3} R^i{}_j(t)\, x^j + c^i(t)\,, \quad i = 1, 2, 3\,, \end{cases} \tag{1.4}$$

dove $c \in \mathbb{R}$ è una costante, le funzioni $\mathbb{R} \ni t \mapsto c^i(t)$ e $\mathbb{R} \ni t \mapsto R^i{}_j(t)$ sono C^∞ e, per ogni $t \in \mathbb{R}$, la matrice di coefficienti $R^i{}_j(t)$ è una matrice ortogonale reale 3×3.
3. Sia $\mathscr{I}$ un riferimento con coordinate cartesiane ortonormali solidali (t, x^1, x^2, x^3). Considerare

un secondo sistema di coordinate (globale) su $\mathbb{V}^4$, (t', x'^1, x'^2, x'^3), connesso al precedente dalle (1.4) (con funzioni $\mathbb{R} \ni t \mapsto c^i(t)$ e $\mathbb{R} \ni t \mapsto R^i{}_j(t)$ con le caratteristiche dette sopra). Mostrare che:

(i) esiste un riferimento $\mathscr{I}'$ per cui le coordinate (t', x'^1, x'^2, x'^3) sono coordinate cartesiane ortonormali solidali con esso;

(ii) Se $\mathscr{I} = \mathscr{I}'$ le funzioni $\mathbb{R} \ni t \mapsto c^i(t)$ e $\mathbb{R} \ni t \mapsto R^i{}_j(t)$ sono costanti e $c = 0$. Viceversa, se tali funzioni sono costanti e $c = 0$, i riferimenti $\mathscr{I}$ e $\mathscr{I}'$ sono lo stesso riferimento.

Osservazioni 1.3. Attraverso le coordinate cartesiane ortonormali solidali con ogni fissato riferimento $\mathscr{I}$ *non* è possibile dotare lo spaziotempo di una *struttura affine* privilegiata per la quale le coordinate cartesiane solidali con ogni riferimento (includendo le coordinate temporali!) siano sistemi di coordinate cartesiani compatibili con la struttura affine. La ragione è nella forma generale delle equazioni (1.4). Se fissiamo un riferimento $\mathscr{I}$ e consideriamo due eventi p e q, possiamo associare alla coppia (p, q) un vettore $p - q$ definito dalle componenti $(t_p - t_q, x_p - x_q, y_p - y_q, z_p - z_q)$, dove abbiamo introdotto un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$. Se cambiamo sistema di coordinate continuando ad usare coordinate solidali con lo stesso $\mathscr{I}$, scopriamo che la legge di trasformazione delle componenti di $p - q$ è correttamente una trasformazione affine, della forma (1.4), con coefficienti c^i nulli e dove i coefficienti $R^i{}_j(t)$ non dipendono dal tempo. Tuttavia se passiamo ad un altro riferimento $\mathscr{I}'$ e proviamo ad associare alla stessa coppia (p, q) un vettore $p - q$ definito dalle componenti $(t'_p - t'_q, x'_p - x'_q, y'_p - y'_q, z'_p - z'_q)$, dove abbiamo introdotto un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}'$, scopriamo che, in generale, le componenti di tale vettore non si trasformano secondo una trasformazione affine passando al precedente riferimento. Questo è dovuto alla dipendenza temporale dei coefficienti $R^i{}_j(t)$ e $c^i(t)$ in (1.4) che appare cambiando riferimento ed al fatto che siamo liberi di scgliere gli eventi p e q non contemporanei. In definitiva: non è possibile, usando i sistemi di coordinate cartesiani solidali con i riferimenti possibili nello spaziotempo, definire il vettore $p - q$ in maniera indipendente dal riferimento.

1.3 Cinematica assoluta del punto materiale.

Ci occuperemo ora della descrizione del moto di un punto materiale rispetto ad un assegnato sistema di riferimento $\mathscr{I}$, introdurremo il concetto di velocità ed accelerazione e vedremo come quest'ultima si decompone su una terna associata alla curva descritta dal punto materiale nello spazio $E_{\mathscr{I}}$.

1.3.1 Derivazione di curve in spazi affini.

Se $\mathbb{A}^n$ è uno spazio affine con spazio delle traslazioni V e $P = P(t)$ con $t \in (a, b)$ descrive una curva C^k, i vettori $(p \leq k)$ per $\tau \in (a, b)$:

$$\frac{d^n P}{dt^n}|_{t=\tau} \in V$$

sono definiti usando la struttura differenziabile di $\mathbb{A}^n$ indotta da quella affine. Fissate coordinate cartesiane di origine O e con assi $\{\mathbf{e}_i\}_{i=1,\dots,n}$, la curva $P = P(t)$ sarà individuata da una curva $x^i = P^i(t)$ in $\mathbb{R}^n$ di classe C^k. Definiamo allora:

$$\frac{d^n P}{dt^n}|_{t=\tau} := \sum_{i=1}^{n} \frac{d^n P^i}{dt^n}|_{t=\tau}\mathbf{e}_i \ . \tag{1.5}$$

Similmente, se $\mathbf{v} = \mathbf{v}(t) \in V$ con $t \in (a,b)$ descrive una curva a valori nello spazio delle traslazioni di uno spazio affine $\mathbb{A}^n$, definiamo per $\tau \in (a,b)$, purché il secondo membro esista

$$\frac{d^n \mathbf{v}}{dt^n}|_{t=\tau} := \sum_{i=1}^{n} \frac{d^n v^i}{dt^n}|_{t=\tau}\mathbf{e}_i \ . \tag{1.6}$$

dove le $v^i(t)$ sono le componenti di $\mathbf{v}(t)$ rispetto alla base $\{\mathbf{e}_i\}_{i=1,\dots,n}$ di V. Vogliamo provare che le due definizioni sono ben poste, cioé sono indipendenti dalla scelta del sistema di coordinate cartesiane nel primo caso e dalla scelta della base nel secondo.
Dall'esercizio A.4.**1**, le coordinate di $P(t)$, $P'^j(t)$ in un nuovo sistema di coordinate cartesiane su $\mathbb{A}^n$ sono connesse alle $P^i(t)$ dalla trasformazione lineare non omogenea:

$$P'^j(t) = \sum_{i=1}^{n} B^j{}_i (P^i(t) + b^i)$$

dove $\mathbf{e}'_j = \sum_k (B^{-1})^k{}_j \mathbf{e}_k$. Di conseguenza:

$$\frac{d^n P'^j}{dt^n} = \sum_{i=1}^{n} B^j{}_i \frac{d^n P^i}{dt^n} \ ,$$

e quindi

$$\sum_j \frac{d^n P'^j}{dt^n}\mathbf{e}'_j = \sum_{i,j,k} (B^{-1})^k{}_j B^j{}_i \frac{d^n P^i}{dt^n}\mathbf{e}_k = \sum_{i,k} \delta_i^k \frac{d^n P^i}{dt^n}\mathbf{e}_k = \sum_i \frac{d^n P^i}{dt^n}\mathbf{e}_i \ .$$

La dimostrazione nel secondo caso è analoga.

1.3.2 Grandezze cinematiche.

Possiamo dare le definizione fondamentali della cinematica: la velocità e l'accelerazione di un punto materiale rispetto ad un fissato riferimento $\mathscr{I}$ in cui è data la legge oraria del moto del punto.

Definizione 1.5. Consideriamo un sistema di riferimento $\mathscr{I}$ ed un punto materiale descritto da una linea di universo $I \ni t \mapsto \gamma(t) \in \mathbb{V}^4$. Possiamo rappresentare questa linea di universo nello spazio di quiete $E_{\mathscr{I}}$ come una curva, che per costruzione è differenziabile,

$$t \mapsto P(t) := \Pi_{\mathscr{I}}(\gamma(t)) \, .$$

La curva sopra scritta è detta **legge oraria** della linea di universo γ nel riferimento $\mathscr{I}$. La **velocità di γ rispetto a $\mathscr{I}$, all'istante** τ, è il vettore di $V_{\mathscr{I}}$:

$$\mathbf{v}|_{\mathscr{I}}(\tau) := \frac{dP_\gamma}{dt}|_{t=\tau} \, .$$

L'**accelerazione di γ rispetto a $\mathscr{I}$, all'istante** τ, è il vettore di $V_{\mathscr{I}}$:

$$\mathbf{a}|_{\mathscr{I}}(\tau) := \frac{d^2P_\gamma}{dt^2}|_{t=\tau} \, .$$

$\diamondsuit$

Esercizi 1.2.

1. Considerare, in un piano nello spazio di quiete del riferimento $\mathscr{I}$, $E_{\mathscr{I}}$, coordinate polari piane r, ϕ ($r > 0$, $\phi \in (-\pi, \pi)$) riferite a coordinate cartesiane ortogonali x, y. Esprimere la velocità , $\mathbf{v}$, e l'accelerazione, $\mathbf{a}$, rispetto ad $\mathscr{I}$ di un punto vincolato a muoversi nel piano usando le coordinate polari ed i versori $\mathbf{e}_r$ (tangente alla curva coordinate r ed uscente dall'origine) ed $\mathbf{e}_\varphi$ (tangente alla curva coordinata ϕ con verso dato da θ crescente). Provare in particolare che, indicando con il punto la derivata rispetto al tempo:

$$\begin{aligned} \mathbf{v} &= \dot{r}\,\mathbf{e}_r + r\dot{\phi}\,\mathbf{e}_\varphi \, , & (1.7) \\ \mathbf{a} &= (\ddot{r} - r\dot{\phi}^2)\,\mathbf{e}_r + (r\ddot{\phi} + 2\dot{r}\dot{\phi})\,\mathbf{e}_\varphi & (1.8) \end{aligned}$$

2. Considerare, nello spazio di quiete del riferimento $\mathscr{I}$, $E_{\mathscr{I}}$, coordinate polari sferiche r, ϕ, θ, ($r > 0$, $\phi \in (-\pi, \pi)$, $\theta \in (0, \pi)$) riferite a coordinate cartesiane ortogonali x, y, z. Esprimere la velocità e l'accelerazione rispetto ad $\mathscr{I}$ di un punto usando le coordinate polari ed i versori $\mathbf{e}_r$ (tangente alla curva coordinate r ed uscente dall'origine), $\mathbf{e}_\varphi$ (tangente alla curva coordinata ϕ con verso dato da ϕ crescente) ed $\mathbf{e}_\theta$ (tangente alla curva coordinata θ con verso dato da θ crescente). Provare, in particolare che:

$$\begin{aligned} \mathbf{v} &= \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta + r\dot{\phi}\sin\theta\,\mathbf{e}_\varphi \, , & (1.9) \\ \mathbf{a} &= (\ddot{r} - r\dot{\theta}^2 - r\dot{\phi}^2\sin^2\theta)\,\mathbf{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\dot{\phi}^2\sin\theta\cos\theta)\,\mathbf{e}_\theta \\ &\quad + (r\ddot{\phi}\sin\theta + 2\dot{r}\dot{\phi}\sin\theta + 2r\dot{\phi}\dot{\theta}\cos\theta)\,\mathbf{e}_\varphi \, . & (1.10) \end{aligned}$$

3. Assiderare, nello spazio di quiete del riferimento $\mathscr{I}$, $E_{\mathscr{I}}$, coordinate cilindriche r, ϕ, z ($r > 0$, $\phi \in (-\pi, \pi)$, $z \in \mathbb{R}$) riferite a coordinate cartesiane ortogonali x, y, z. Esprimere la velocità e l'accelerazione rispetto ad $\mathscr{I}$ di un punto usando le coordinate polari ed i versori $\mathbf{e}_r$

(tangente alla curva coordinate r ed uscente dall'origine), $\mathbf{e}_\varphi$ (tangente alla curva coordinata ϕ con verso dato da ϕ crescente) ed $\mathbf{e}_z$ (tangente alla curva coordinata z con verso positivo). Provare in particolare che, indicando con il punto la derivata rispetto al tempo:

$$\mathbf{v} = \dot{r}\,\mathbf{e}_r + r\dot{\phi}\,\mathbf{e}_\varphi + \dot{z}\,\mathbf{e}_z\,, \tag{1.11}$$

$$\mathbf{a} = (\ddot{r} - r\dot{\phi}^2)\,\mathbf{e}_r + (r\ddot{\phi} + 2\dot{r}\dot{\phi})\,\mathbf{e}_\varphi + \ddot{z}\,\mathbf{e}_z \tag{1.12}$$

4.* Gli esercizi 1.2.**2** e 1.2.**3**, per quanto riguarda il calcolo dell'accelerazione, possono essere risolti in maniera più rapida usando la derivata covariante rispetto alla connessione affine di Levi-Civita di $\mathbb{E}^3 = E_{\mathscr{I}}$ associata al prodotto scalare. Vale in coordinate locali curvilinee arbitrarie x^1, x^2, x^3, se $(x^1(t), x^2(t), x^3(t))$ sono le coordinate di $P(t)$ al variare del tempo $t \in \mathbb{R}$,

$$\mathbf{a} = \sum_{i=1}^{3} \left(\frac{d^2x^i}{dt^2} + \sum_{j,k=1}^{3} \Gamma^i{}_{jk} \frac{dx^j}{dt}\frac{dx^k}{dt} \right) \frac{\partial}{\partial x^i}\,. \tag{1.13}$$

dove i simboli di connessione $\Gamma^i{}_{jk}$ sono, al solito, definiti come, rispetto al tensore metrico $\sum_{ik} g_{ik} dx^i \otimes dx^k$ (ed al suo inverso di componenti g^{rs}):

$$\Gamma^i{}_{jk} = \sum_{r=1}^{3} \frac{1}{2} g^{ir} \left(\frac{\partial g_{rj}}{\partial x^k} + \frac{\partial g_{kr}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^r} \right) \tag{1.14}$$

$\frac{\partial}{\partial x^i}$ sono i vettori della base associata alle coordinate connessi ai versori tramite $\mathbf{e}_i = (g_{ii})^{-1/2}\frac{\partial}{\partial x^i}$. Per esempio, nel caso delle coordinate polari sferiche, se $x^1 = r, x^2 = \theta, x^3 = \phi$, $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ si ha: $\mathbf{e}_r = \frac{\partial}{\partial r}$, $\mathbf{e}_\theta = r^{-1}\frac{\partial}{\partial \theta}$, $\mathbf{e}_\varphi = (r\sin\theta)^{-1}\frac{\partial}{\partial \phi}$. Spiegare il perché di questo risultato.

1.3.3 Cinematica per punti materiali vincolati a curve e superfici ferme.

Spesso il moto di punti materiali è *vincolato*, cioé costretto ad avvenire in insiemi geometrici fissati, che possono essere curve o superfici. Ci occuperemo qui della situazione in cui il moto avviene su curve o superfici ferme rispetto ad un sistema di riferimento. Un punto tecnicamente importante è quello di sviluppare tecniche matematiche per poter di decomporre la velocità e l'accelerazione del punto materiale rispetto a basi vettoriali associate agli insiemi dati. Sviluppando la dinamica del punto materiale vincolato, vedremo che sarà fisicamente importante poter decomporre l'accelerazione del punto materiale studiato in *componente tangente* e *componente e normale* alla superficie o alla curva su cui avviene il moto.

Consideriamo una curva Γ, di equazione $P = P(u)$ con $u \in (a, b)$, *in quiete* rispetto al riferimento $\mathscr{I}$. In altre parole, l'equazione della curva, in coordinate cartesiane ortonormali solidali con $\mathscr{I}$, è tale che in essa non compaia esplicitamente il tempo (il parametro u non è il tempo). Richiederemo che la curva sia di classe C^1 (in modo in particolare che sia rettificabile [GiustiI]) e che sia *regolare*, ossia il vettore tangente sia ovunque non nullo su Γ:

$$\frac{dP}{du} \neq 0\,, \quad \text{per } u \in (a, b).$$

Supponiamo che un punto materiale sia vincolato a muoversi su tale curva. Notiamo che la curva Γ ammette una parmetrizzazione privilegiata che ora introduciamo. Definiamo, fissato un punto $O = P(u_0)$ sulla curva:

$$s(u) := \int_{u_0}^{u} \left|\left| \frac{dP}{du'} \right|\right| du' \,. \tag{1.15}$$

Dal teorema fondamentale del calcolo, la funzione $(a, b) \ni u \mapsto s(u)$, detta **ascissa curvilinea**, risulta essere C^1 con derivata sempre positiva data dall'integrando in (1.15). In virtù di ciò essa è strettamente crescente e quindi invertibile con inversa C^1. Di conseguenza s può essere usato per riparametrizzare Γ. Dalla teoria elementare delle curve [GiustiI] è noto che $s(u)$ altro non è che la *lunghezza* del segmento di curva tra $P(u_0)$ e $P(u)$. Tale lunghezza è definita con il segno: è positiva per punti successivi a $P(u_0)$ e negativa per punti precedenti $P(u_0)$, il senso di percorrenza essendo stato fissato dal parametro iniziale u.
Cambiamo ora parametro, per descrivere la stessa Γ, usando un nuovo parametro $v = v(u)$, nell'ipotesi che la funzione $v = v(u)$ sia C^1 con derivata strettamente positiva. In tali ipotesi la funzione $v = v(u)$ risulta essere biettiva con inversa C^1 e la curva Γ potrà essere descritta da una funzione $P(v) := P(u(v))$ con $v \in (v(a), v(b))$. Possiamo anche in questo caso definire l'ascissa curvilinea $s = s(v)$ tramite la (1.15). Nel seguito, dato che non ci sarà ambiguità , indicheremo con $\cdot$ il prodotto scalare nello spazio euclideo $E_{\mathscr{I}}$. Direttamente dalla (1.15) troviamo che

$$s(u) = \int_{u_0}^{u} \left|\left| \frac{dP}{du'} \right|\right| du' = \int_{v(u_0)}^{v(u)} \left|\left| \frac{dP}{du'} \right|\right| \frac{du'}{dv'} dv' = \int_{v(u_0)}^{v(u)} \left|\left| \frac{dP}{dv'} \right|\right| dv' = s(v(u)) \,.$$

In questo senso l'ascissa curvilinea è *invariante* sotto riparametrizzazioni di Γ (date da funzioni differenziabili con continuità e con derivata ovunque strettamente positiva).
Supponiamo ora di parametrizzare Γ con il parametro s. Scriveremo pertanto $P(s) := P(u(s))$ con $s \in (s_a, s_b)$. Per prima cosa notiamo che il vettore tangente alla curva associato al parametro $s \in (s_a, s_b)$:

$$\mathbf{t}(s) := \frac{dP}{ds} \tag{1.16}$$

ha norma unitaria. Infatti dalla (1.15), usando s come parametro, ed applicando il primo teorema fondamentale del calcolo, si ha

$$1 = \frac{ds}{ds} = \left|\left| \frac{dP}{ds} \right|\right| \,.$$

Un secondo versore associato a Γ si può definire quando $\mathbf{t}$ non è costante. Si tratta del **versore normale**:

$$\mathbf{n}(s) := \rho(s) \frac{d\mathbf{t}}{ds} \,, \quad \text{con } \rho(s) := \left|\left| \tfrac{d\mathbf{t}}{ds} \right|\right|^{-1} \,. \tag{1.17}$$

$\rho(s)$ è detto **raggio di curvatura** nel punto $P(s)$. Se $\mathbf{t}$ e $\mathbf{n}$ sono entrambi definiti, allora $\mathbf{t} \perp \mathbf{n}$:

$$\mathbf{t} \cdot \mathbf{n} = \rho \mathbf{t} \cdot \frac{d\mathbf{t}}{ds} = \frac{\rho}{2} \frac{d\mathbf{t} \cdot \mathbf{t}}{ds} = \frac{\rho}{2} \frac{d}{ds} 1 = 0 \,.$$

Se esistono $\mathbf{t}$ e $\mathbf{n}$, un terzo versore normale ad entrambi, detto **binormale**, si definisce come:

$$\mathbf{b}(s) := \mathbf{t}(s) \wedge \mathbf{n}(s) \,. \tag{1.18}$$

I versori $\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)$, quando esistono, formano una terna *ortonormale destrorsa* per costruzione. Tale terna d viene detta **terna intrinseca** (o **terna di Frenet**) di Γ nel punto $P(s)$.

Osservazioni 1.4.
(1) L'ascissa curvilinea s è definita a meno della scelta dell'origine di Γ (il punto in cui vale $s = 0$) per cui, se s è un'ascissa curvilinea, lo è anche $s'(s) := s + c$, dove $c \in \mathbb{R}$ è una costante arbitrariamente fissata. I versori della terna intrinseca, essendo costruiti con derivate in s, non risentono dell'ambiguità nella definizione di s.
(2) Per una curva Γ regolare vale $\left|\left|\frac{d\mathbf{t}}{ds}\right|\right| = 0$ per ogni $s \in (s_1, s_2)$ se e solo se Γ è riparametrizzabile (usando un parametro u con $u = u(s)$ di classe C^1 con derivata ovunque non nulla) come un segmento di retta per $s \in (s_1, s_2)$. Infatti, se Γ è riparametrizzabile come un segmento di retta tra $P(s_1)$ e $P(s_2)$, allora $P(s(u)) = P(s_1) + u(P(s_2) - P(s_1))$ tra i sue punti detti. Applicando la definizione di ascissa curvilinea, si ha subito che $s(u) = u||P(s_2) - P(s_1)||$ e quindi, come ci si aspettava intuitivamente, deve essere

$$P(s) = P(s_1) + s\frac{P(s_2) - P(s_1)}{||P(s_2) - P(s_1)||}$$

da cui $\mathbf{t} = \frac{P(s_2)-P(s_1)}{||P(s_2)-P(s_1)||}$ è costante e $\left|\left|\frac{d\mathbf{t}}{ds}\right|\right| = 0$. Se viceversa $\left|\left|\frac{d\mathbf{t}}{ds}\right|\right| = 0$, vuol dire che $\mathbf{t}$ è costante e di conseguenza, integrando l'equazione $\frac{dP(s)}{ds} = \mathbf{t}$ in (s_1, s_2), deve valere l'equazione del segmento di retta:

$$P(s) = P(s_1) + s\mathbf{t} \,.$$

Se un punto materiale (con linea di universo γ) è vincolato a muoversi sulla curva regolare Γ, ferma nel riferimento $\mathscr{I}$, la posizione del punto al variare del tempo è completamente determinata dall'ascissa curvilinea s pensata come funzione del tempo $s = s(t)$. Pertanto si devono poter esprimere la velocità e l'accelerazione del punto rispetto a $\mathscr{I}$ in termini di s e di altre caratteristiche geometriche della curva. Supporremo che Γ non sia parametrizzabile come un segmento di retta, in modo tale che esista, in ogni suo punto, la terna intrinseca. Se la curva ha equazione $P = P(s)$, la legge oraria del punto materiale sarà scrivibile come $P(t) = P(s(t))$. Applicando la definizione di velocità e di vettore tangente si ha immediatamente:

$$\mathbf{v}(t) = \frac{ds}{dt}\,\mathbf{t}(s(t)) \,, \tag{1.19}$$

dove abbiamo omesso la specificazione del riferimento $|_{\mathscr{I}}$ per semplicità. Derivando un'altra volta nel tempo si ha

$$\mathbf{a}(t) = \frac{d^2s}{dt^2}\,\mathbf{t}(s(t)) + \frac{ds}{dt}\frac{d\mathbf{t}}{ds}\frac{ds}{dt} \,,$$

da cui, usando la definizione di $\mathbf{n}$, si ha l'espressione dell'accelerazione:

$$\mathbf{a}(t) = \frac{d^2 s}{dt^2}\,\mathbf{t}(s(t)) + \frac{1}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2 \mathbf{n}(s(t))\,. \tag{1.20}$$

L'accelerazione appare decomposta in componente tangente e componente normale alla curva di vincolo, ciò sarà di centrale importanza sviluppando la dinamica del punto materiale vincolato.

Esempi 1.1.
1. Si consideri la curva (elica), $P = P(u)$, in coordinate cartesiane ortonormali data dalle equazioni parametriche $x(u) = \cos u$, $y(u) = \sin u$, $z = u$, con $u \in \mathbb{R}$. Scriviamone la terna intrinseca rispetto alla base di versori $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. L'ascissa curvilinea si ottiene integrando

$$s(u) = \int_0^s \left|\left|\frac{dP(u)}{du}\right|\right| du = \int_0^s \sqrt{(-\sin u)^2 + (\cos u)^2 + 1}du = \int_0^s \sqrt{2}du\,.$$

Quindi $s(u) = \sqrt{2}u$ se scegliamo l'origine di s nel punto di coordinate $(1, 0, 0)$. In definitiva la curva ha equazioni parametriche: $x(s) = \cos(s/\sqrt{2})$, $y(u) = \sin(s/\sqrt{2})$, $z = (s/\sqrt{2})$. Il calcolo diretto produce, per la terna intrinseca:

$$\mathbf{t}(s) = \frac{1}{\sqrt{2}}\Big(-\sin(s/\sqrt{2})\,\mathbf{e}_x + \cos(s/\sqrt{2})\,\mathbf{e}_y + \mathbf{e}_z\Big)\,, \tag{1.21}$$

$$\mathbf{n}(s) = -\Big(\cos(s/\sqrt{2})\,\mathbf{e}_x + \sin(s/\sqrt{2})\,\mathbf{e}_y\Big)\,, \tag{1.22}$$

$$\mathbf{b}(s) = \frac{1}{\sqrt{2}}\Big(\sin(s/\sqrt{2})\,\mathbf{e}_x - \cos(s/\sqrt{2})\,\mathbf{e}_y + \mathbf{e}_z\Big)\,. \tag{1.23}$$

Dato che la base $\mathbf{t}, \mathbf{n}, \mathbf{b}$ è ortonormale come $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$, la trasformazione di sopra si inverte immediatamente scambiando la matrice della trasformazione con la sua trasposta:

$$\mathbf{e}_x = -\frac{1}{\sqrt{2}}\sin(s/\sqrt{2})\mathbf{t}(s) - \cos(s/\sqrt{2})\mathbf{n}(s) + \frac{1}{\sqrt{2}}\sin(s/\sqrt{2})\mathbf{b}(s)\,, \tag{1.24}$$

$$\mathbf{e}_y = \frac{1}{\sqrt{2}}\cos(s/\sqrt{2})\mathbf{t}(s) - \sin(s/\sqrt{2})\mathbf{n}(s) - \frac{1}{\sqrt{2}}\cos(s/\sqrt{2})\mathbf{b}(s)\,, \tag{1.25}$$

$$\mathbf{e}_z = \frac{1}{\sqrt{2}}\mathbf{t}(s) + \frac{1}{\sqrt{2}}\mathbf{b}(s)\,. \tag{1.26}$$

Esercizi 1.3.
1. Si consideri la curva Γ, C^1 e regolare in $\mathbb{E}^3$, parametrizzata tramite la sua ascissa curvilinea $P = P(s)$, $s \in (a, b)$. Si dimostri che se Γ è contenuta in un piano Π e non è riparametrizzabile come un segmento di retta in ogni sottointervallo di (a, b), allora il versore binormale $\mathbf{b}$ è costante su Γ ed è perpendicolare a Π.

2. Si consideri la curva Γ, C^1 e regolare in $\mathbb{E}^3$, parametrizzata tramite la sua ascissa curvilinea $P = P(s)$, $s \in (a, b)$. Dimostrare che se il vettore binormale $\mathbf{b}$ esiste, è non nullo ed è costante su Γ, allora la curva è tutta contenuta in un piano normale a $\mathbf{b}$.

3. Dimostrare che valgono le *formule di Darboux*, per una curva C^1 regolare parametrizzata tramite la sua ascissa curvilinea $P = P(s)$, $s \in (a, b)$ (nell'ipotesi che i vettori e gli scalari scritti esistano):

$$\frac{d\mathbf{n}(s)}{ds} = -\frac{\mathbf{t}(s)}{\rho(s)} - \frac{\mathbf{b}(s)}{\tau(s)} = \Omega(s) \wedge \mathbf{n}(s) \,, \tag{1.27}$$

$$\frac{d\mathbf{b}(s)}{ds} = -\frac{\mathbf{n}(s)}{\tau(s)} = \Omega(s) \wedge \mathbf{b}(s) \,, \tag{1.28}$$

$$\frac{d\mathbf{t}(s)}{ds} = -\frac{\mathbf{n}(s)}{\rho(s)} = \Omega(s) \wedge \mathbf{t}(s) \,, \tag{1.29}$$

dove τ, definito dalla (1.28), è detto *raggio di torsione* e

$$\Omega(s) := \rho(s)^{-1}\mathbf{b}(s) - \tau(s)^{-1}\mathbf{t}(s)$$

è il *vettore di Darboux* nel punto $P(s)$.

Passiamo a considerare un punto materiale vincolato ad una superficie ferma in un riferimento $\mathscr{I}$. In altre parole, in coordinate cartesiane ortonormali solidali con $\mathscr{I}$ le equazioni che determinano S non contengono esplicitamente il tempo. Assumeremo che tale superficie S sia descritta, nello spazio euclideo $\mathbb{E}^3 = E_{\mathscr{I}}$, da equazioni parametriche $P = P(u, v)$, con $(u, v) \in D$ insieme aperto (e connesso) di $\mathbb{R}^2$. Assumeremo che la superficie sia *regolare*, cioé la funzione $P = P(u, v)$ sia di classe C^1 almeno e valga:

$$\frac{\partial P}{\partial u} \wedge \frac{\partial P}{\partial v} \neq 0 \quad \text{per ogni } (u, v) \in D \,.$$

Dall'esempio A.2.**4** segue immediatamente che S è una sottovarietà embedded di $E_{\mathscr{I}}$ di dimensione 2 e che (u, v) sono coordinate globali della sua struttura differenziabile.
Nelle ipotesi fatte i tre vettori $\frac{\partial P}{\partial u}, \frac{\partial P}{\partial v}, \frac{\partial P}{\partial u} \wedge \frac{\partial P}{\partial v}$ costituiscono, punto per punto, su S, una base vettoriale. Notare che il terzo vettore, per costruzione è sempre perpendicolare a S. Per studiare la cinematica e, più ancora, per studiare la dinamica di un punto materiale vincolato a S, è conveniente parametrizzare il moto del punto tramite le coordinate $(u, v) = (u(t), v(t))$ e quindi esprimere la velocità e l'accelerazione usando la base vettoriale sopra introdotta oppure quella dei *versori* associati ai vettori della base. Si osservi che è possibile, entro certi limiti, modificare la scelta dei parametri u, v per descrivere S in modo da semplificare i calcoli. In particolare, quando è possibile, è molto conveniente scegliere i parametri u, v in modo tale che la base di versori detta sopra sia ortonormale.

Esempi 1.2.
1. Si consideri la porzione di cono S di equazioni $0 < \kappa z = \sqrt{x^2 + y^2}$ ($\kappa > 0$ fissato) riferita a

coordinate ortonormali $\mathbf{x} = (x, y, z)$ in $\mathbb{E}^3 = E_{\mathscr{I}}$. Possiamo parametrizzare S tramite $u = x$ e $v = y$ e le equazioni parametriche risultano $x = u, y = v, z = \sqrt{u^2 + v^2}$ dove il dominio di tali coordinate è $D = \{(u, v) \in \mathbb{R}^2 \mid 0 < u^2 + v^2 < 1\}$. Si verifica immediatamente che la condizione di regolarità , in coordinate cartesiane

$$\frac{\partial \mathbf{x}}{\partial u} \wedge \frac{\partial \mathbf{x}}{\partial v} \neq 0 \quad \text{per ogni } (u, v) \in D \,,$$

è soddisfatta. Definiamo la funzione $\zeta := \sqrt{(x^2 + y^2) + z(x, y)^2} = \sqrt{(1 + \kappa^{-2})(x^2 + y^2)}$ di chiaro significato geometrico (distanza tra un punto di coordinate $\mathbf{x} \in C$ ed il vertice O del cono) ed indichiamo con ϕ l'ordinaria coordinata angolare polare piana nel piano x, y. Se riparametrizziamo la superficie con coordinate $\zeta \in (0, +\infty)$ e $\phi \in (-\pi, \pi)$ (le coordinate non sono globali perché la semiretta a $\phi = \pm\pi$ non è coperta) S, la base vettoriale associata a tali coordinate può essere fissata come ortonormale. Un punto sul cono è descritto da

$$x = (\sin\alpha)\zeta\cos\phi \,, \quad y = (\sin\alpha)\zeta\sin\phi \,, \quad z = (\cos\alpha)\zeta \,,$$

dove $\kappa = \tan\alpha$. I versori $\mathbf{e}_\zeta$ e $\mathbf{e}_\varphi$ tangenti alle curve coordinate ζ e ϕ (diretti nella direzione in cui crescono tali coordinate), ed il loro prodotto vettore $\mathbf{n} := \mathbf{e}_\zeta \wedge \mathbf{e}_\varphi$ normale a S, sono, rispetto alla base cartesiana:

$$\mathbf{e}_\zeta = \sin\alpha\,\cos\phi\,\mathbf{e}_x + \sin\alpha\,\sin\phi\,\mathbf{e}_y + \cos\alpha\,\mathbf{e}_z \,, \tag{1.30}$$

$$\mathbf{e}_\varphi = -\sin\phi\,\mathbf{e}_x + \cos\phi\,\mathbf{e}_y \,, \tag{1.31}$$

$$\mathbf{n} = -\cos\alpha\,\cos\phi\,\mathbf{e}_x - \cos\alpha\,\sin\phi\,\mathbf{e}_y + \sin\alpha\,\mathbf{e}_z \,. \tag{1.32}$$

Queste relazioni si invertono (prendendo la trasposta della matrice di trasformazione) in

$$\mathbf{e}_x = \sin\alpha\,\cos\phi\,\mathbf{e}_\zeta - \sin\phi\,\mathbf{e}_\varphi - \cos\alpha\,\cos\phi\,\mathbf{n} \,, \tag{1.33}$$

$$\mathbf{e}_y = \sin\alpha\,\sin\phi\,\mathbf{e}_\zeta + \cos\phi\,\mathbf{e}_\varphi - \cos\alpha\,\sin\phi\,\mathbf{n} \,, \tag{1.34}$$

$$\mathbf{e}_z = \cos\alpha\,\mathbf{e}_\zeta + \sin\alpha\,\mathbf{n} \,. \tag{1.35}$$

Assumiamo ora che $\zeta = \zeta(t)$, $\phi = \phi(t)$ descrivano l'evoluzione del punto $P(t)$ su S. La posizione di $P(t)$ può essere individuata dal vettore posizione: $P(t) - O = \zeta(t)\mathbf{e}_\zeta(t)$. La derivata nel tempo definisce la velocità del punto nel riferimento $\mathscr{I}$ in cui il cono S è immobile. Bisogna tenere conto del fatto che $\mathbf{e}_\zeta$ varia nel tempo. Indicando con il punto la derivata temporale si ha:

$$\mathbf{v}(t) = \dot{\zeta}(t)\mathbf{e}_\zeta(t) + \zeta(t)\dot{\mathbf{e}}_\zeta(t) \,.$$

La derivata temporale dei versori si ottiene derivando il secondo membro di (1.30)-(1.32) (assumendo $\zeta = \zeta(t)$, $\phi = \phi(t)$) e quindi usando (1.33)-(1.35) nel risultato. Il calcolo fornisce:

$$\dot{\mathbf{e}}_\zeta = \dot{\phi}\sin\alpha\,\mathbf{e}_\varphi \,, \tag{1.36}$$

$$\dot{\mathbf{e}}_\varphi = -\dot{\phi}\sin\alpha\,\mathbf{e}_\zeta + \dot{\phi}\cos\alpha\,\mathbf{n} \,, \tag{1.37}$$

$$\dot{\mathbf{n}} = -\dot{\phi}\cos\alpha\,\mathbf{e}_\varphi \,, \tag{1.38}$$

da cui

$$\mathbf{v}(t) = \dot{\zeta}(t)\mathbf{e}_\zeta(t) + \dot{\phi}(t)\zeta(t)\sin\alpha\,\mathbf{e}_\varphi(t)\,. \tag{1.39}$$

Abbiamo espresso la velocità sulla base associata alla superficie su cui il punto è vincolato. Derivando (1.39) nel tempo

$$\mathbf{a}(t) = \ddot{\zeta}(t)\mathbf{e}_\zeta(t) + \dot{\zeta}(t)\dot{\mathbf{e}}_\zeta(t) + \dot{\phi}(t)\dot{\zeta}(t)\sin\alpha\,\mathbf{e}_\varphi(t) + \ddot{\phi}(t)\zeta(t)\sin\alpha\,\mathbf{e}_\varphi + \dot{\phi}(t)\zeta(t)\sin\alpha\,\dot{\mathbf{e}}_\varphi(t)\,,$$

e procedendo come sopra (usando le espressioni ottenute per le derivate dei versori), si ottiene infine l'espressione dell'accelerazione decomposta in parte normale e parte tangente alla superficie conica:

$$\mathbf{a}(t) = \Big(\ddot{\zeta}(t) - \dot{\phi}(t)^2\zeta(t)\sin^2\alpha\Big)\,\mathbf{e}_\zeta(t) + \Big(\ddot{\phi}(t)\zeta(t) + 2\dot{\phi}(t)\dot{\zeta}(t)\Big)\sin\alpha\mathbf{e}_\varphi(t) + \frac{\dot{\phi}(t)^2\zeta(t)\sin 2\alpha}{2}\mathbf{n}(t). \tag{1.40}$$

2. Le formule degli esercizi 1.2.**2** e 1.2.**3** permettono di scrivere velocità ed accelerazione per un punto materiale vincolato a muoversi rispettivamente su una superficie sferica e su una superficie cilindrica. Descrivendo la prima in coordinate polari sferiche r, ϕ, θ ed individuando la superficie fissando $r = R$, le coordinate ϕ e θ risultano essere coordinate ammissibili sulla superficie con $\mathbf{e}_\varphi \perp \mathbf{e}_\theta$. Il versore normale alla superficie $\mathbf{n}$ risulta coincidere con $\mathbf{e}_r$. L'espressione per velocità ed accelerazione usando la base ortonormale di tali versori è quindi:

$$\mathbf{v} = R\dot{\theta}\,\mathbf{e}_\theta + R\dot{\phi}\sin\theta\,\mathbf{e}_\varphi\,, \tag{1.41}$$

$$\mathbf{a} = (R\ddot{\theta} - R\dot{\phi}^2\sin\theta\cos\theta)\,\mathbf{e}_\theta + (R\ddot{\phi}\sin\theta + 2R\dot{\phi}\dot{\theta}\cos\theta)\,\mathbf{e}_\varphi - (R\dot{\theta}^2 + R\dot{\phi}^2\sin^2\theta)\mathbf{n}\,. \tag{1.42}$$

Nel caso della superficie cilindrica descritta da $r = R$ in coordinate cilindriche r, ϕ, z, le coordinate ϕ, z sono coordinate ammissibili sulla superficie, vale $\mathbf{e}_\varphi \perp \mathbf{e}_z$ ed il versore normale è ancora $\mathbf{e}_r$. Infine:

$$\mathbf{v} = R\dot{\phi}\,\mathbf{e}_\varphi + \dot{z}\,\mathbf{e}_z\,, \tag{1.43}$$

$$\mathbf{a} = R\ddot{\phi}\,\mathbf{e}_\varphi + \ddot{z}\,\mathbf{e}_z - R\dot{\phi}^2\mathbf{n}\,. \tag{1.44}$$

1.4 Cinematica relativa del punto materiale.

In questa sezione ci occupiamo delle relazioni che intercorrono tra le varie descrizioni del moto di uno stesso punto materiale che vengono date in due (o più) differenti riferimenti. Queste relazioni sono il contenuto matematico della **cinematica relativa**.

Consideriamo un riferimento $\mathscr{I}$. Scegliamo un punto (come entità puramente geometrica) $P \in E_{\mathscr{I}}$ ed un vettore $\mathbf{v} \in V_{\mathscr{I}}$. Questi due oggetti matematici vivono nello spazio di quiete del riferimento, tuttavia possono essere identificati *biunivocamente* con analoghi oggetti matematici, P_t e $\mathbf{v}_t$ rispettivamente, in un fissato spazio assoluto Σ_t quando è fissato un tempo t. Semplicemente basta definire $P_t := (\Pi_{\mathscr{I}} \restriction_{\Sigma_t})^{-1}(P)$ e $\mathbf{v}_t := d(\Pi_{\mathscr{I}} \restriction_{\Sigma_t})^{-1}(\mathbf{v})$. Viceversa oggetti matematici P_t e $\mathbf{v}_t$ rispettivamente, in un fissato spazio assoluto possono essere pensati come nello spazio

euclideo di quiete di un riferimento $\mathscr{I}'$ invertendo le relazioni sopra scritte (ed usando $\mathscr{I}'$ diverso da $\mathscr{I}$ se necessario): $P' := \Pi_{\mathscr{I}'} \restriction_{\Sigma_t} (P_t)$ e $\mathbf{v}' := d(\Pi_{\mathscr{I}'} \restriction_{\Sigma_t})(\mathbf{v}_t)$. Tutte queste identificazioni sono ottenute tramite isomorfismi (isometrie affini e isomorfismi di spazi vettoriali rispettivamente), per cui *la strutture delle operazioni ammissibili tra gli enti considerati è preservata dalle identificazioni*, per esempio, la differenza di due punti in Σ_t diventa la differenza tra i corrispondenti punti nello spazio $E_{\mathscr{I}'}$.
Se γ è una linea di universo di un punto metariale che, nel riferimento $\mathscr{I}$, all'istante t, ha velocità $\mathbf{v}|_{\mathscr{I}}(t)$, quest'ultima può essere pensata come nello spazio assoluto Σ_t piuttosto che nello spazio euclideo di quiete di $\mathscr{I}$. Oppure può essere pensata nello spazio euclideo di quiete di un *altro* riferimento $\mathscr{I}'$. In questo modo, possiamo confrontare le velocità possedute da *uno stesso* punto materiale, ma riferite a *due differenti* sistemi di riferimento. *Per questa ragione, se non ci saranno ambiguità , d'ora in poi indicheremo con lo stesso simbolo due punti appartenenti a spazi euclidei diversi (es. $E_{\mathscr{I}}$, Σ_t), ma identificati tramite $\Pi_{\mathscr{I}}$. Useremo la stessa convenzione per i vettori dei corrispondenti spazi delle traslazioni. Similmente indicheremo i prodotti scalari con il semplice puntino $\cdot$ (omettendo anche la dipendenza temporale), eccetto nelle situazioni in cui le differenti strutture metriche verranno confrontate.*

1.4.1 Vettore $\boldsymbol{\omega}$ e formule di Poisson.

Consideriamo una funzione $\mathbb{R} \mapsto \mathbf{u}(t) \in V_t$ che associa ad ogni istante t un vettore nello spazio assoluto corrispondente Σ_t (quindi più precisamente nello spazio delle traslazioni V_t). Diremo che tale funzione è *differenziabile*, se le componenti di $t \mapsto d(\Pi_{\mathscr{I}} \restriction_{\Sigma_t})(\mathbf{u}(t))$ sono funzioni differenziabili del tempo rispetto ad una base ortonormale di $V_{\mathscr{I}}$ arbitraria.

Osservazioni 1.5. Si prova immediatamente, identificando i vettori con punti degli spazi di quiete e quindi usando le trasformazioni (1.4) tra sistemi di coordinate cartesiani solidali con differenti sistemi di riferimento, che la differenziablità di una curva $\mathbb{R} \mapsto \mathbf{u}(t) \in V_t$ non dipende dal riferimento usato.

Assegnata una funzione $\mathbb{R} \mapsto \mathbf{u}(t) \in V_t$ differenziabile, consideriamo due basi ortonormali $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in V_{\mathscr{I}}$ e $\mathbf{e'}_1, \mathbf{e'}_2, \mathbf{e'}_3 \in V_{\mathscr{I}'}$ riferite agli spazi delle traslazioni di due *differenti* sistemi di riferimento $\mathscr{I}$ e $\mathscr{I}'$ rispettivamente. Avremo che in ogni Σ_t:

$$\mathbf{u}(t) = \sum_{i=1}^{3} u^i(t)\mathbf{e}_i(t) = \sum_{j=1}^{3} u'^j(t)\mathbf{e'}_j(t) \, .$$

Similmente, nel riferimento $\mathscr{I}$, dove ogni $\mathbf{e}_i$ è un vettore costante nel tempo:

$$\mathbf{u}(t) = \sum_{i=1}^{3} u^i(t)\mathbf{e}_i = \sum_{j=1}^{3} u'^j(t)\mathbf{e'}_j(t) \, .$$

Infine, nel riferimento $\mathscr{I}'$, dove ogni $\mathbf{e}'_j$ è un vettore costante nel tempo:

$$\mathbf{u}(t) = \sum_{i=1}^{3} u^i(t)\mathbf{e}_i(t) = \sum_{j=1}^{3} u'^j(t)\mathbf{e}'_j \ .$$

Si osservi che, in base alle convenzioni di identificazione che stiamo usando, $u(t)$ della seconda formula indica propriamente $d(\Pi_{\mathscr{I}} \restriction_{\sigma_t})(\mathbf{u}(t))$.
Nel riferimento $\mathscr{I}$ possiamo derivare rispetto al tempo $\mathbf{u}(t)$ derivando semplicemente le componenti di $\mathbf{u}(t)$ rispetto alla base di elementi $\mathbf{e}_i$.

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{u}(t) := \sum_{i=1}^{3} \frac{d}{dt} u^i(t)\mathbf{e}_i \ .$$

Nel riferimento $\mathscr{I}'$ possiamo derivare rispetto al tempo $\mathbf{u}(t)$, semplicemente derivando le componenti di $\mathbf{u}(t)$ rispetto alla base di elementi $\mathbf{e}'_j$:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}'} \mathbf{u}(t) := \sum_{j=1}^{3} \frac{d}{dt} u'^j(t)\mathbf{e}'_j \ .$$

Possiamo quindi trasportare nello spaziotempo, più precisamente su ogni Σ_t, questi vettori ottenendo, rispettivamente

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{u}(t) := \sum_{i=1}^{3} \frac{du^i(t)}{dt}\mathbf{e}_i(t) \ ,$$

e

$$\left.\frac{d}{dt}\right|_{\mathscr{I}'} \mathbf{u}(t) := \sum_{j=1}^{3} \frac{du'^j(t)}{dt}\mathbf{e}'_j(t) \ .$$

Notare che ora i versori *dipendono* dal tempo, ma le derivate agiscono *solo* sulle componenti, *quando queste sono riferite alla base associata al riferimento rispetto al quale si calcola la derivata*. Enunciamo ora tutto in termini di una definizione formale.

Definizione 1.6. (**Derivata temporale rispetto ad un riferimento**.) Si consideri un curva $(a,b) \ni t \mapsto \mathbf{u}(t) \in V_t$.
(a) Diremo che tale curva è **differenziabile** se le componenti di $(a,b) \ni t \mapsto \Pi_{\mathscr{I}}(\mathbf{u}(t)) \in V_{\mathscr{I}}$ rispetto ad una base ortonormale dello spazio di quiete con un riferimento $\mathscr{I}$ sono funzioni differenziabili.
(b) In tal caso, per ogni $t \in (a,b)$, la **derivata di u rispetto al riferimento** $\mathscr{I}$ è definita come il vettore di V_t

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{u}(t) := \sum_{i=1}^{3} \frac{du^i(t)}{dt}\mathbf{e}_i(t) \ ,$$

dove $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ definiscono una base ortonormale di $V_{\mathscr{I}}$ in cui $\mathbf{u} = \sum_{i=1}^3 u^i \mathbf{e}_i$ e usiamo la convenzione per cui, se $\mathbf{v} \in V_{\mathscr{I}}$, $\mathbf{v}(t) := d(\Pi_{\mathscr{I}} \restriction_{\Sigma_t})^{-1}(\mathbf{v})$. $\diamondsuit$

Osservazioni 1.6.
(1) Per costruzione, la derivata rispetto al riferimento $\mathscr{I}$ non dipende dal sistema di coordinate ortonormali solidale con il riferimento. La prova, riducendosi allo spazio di quiete di $\mathscr{I}$ tramite $d(\Pi_{\mathscr{I}} \restriction_{\Sigma_t})$, è la stessa che abbiamo dato nella sezione 1.3.1. In tale sede abbiamo provato che la derivata di una curva e di una curva di vettori, in uno spazio euclideo, è indipendente dalla scelta del sistema di coordinate cartesiane, nel primo caso, e dalla scelta della base nel secondo.
(2) Se γ è una linea di universo in $\mathbb{V}^4$, per la definizione 1.5, la velocità di γ rispetto a $\mathscr{I}$ può essere calcolata derivando $\Pi_{\mathscr{I}}(\gamma(t))$ rappresentato in coordinate cartesiane ortogonali di origine $O_{\mathscr{I}}$ e assi $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in $E_{\mathscr{I}}$. Equivalentemente, tale velocità si ottiene derivando la curva a valori vettoriali in V_t, $t \mapsto \Pi_{\mathscr{I}}(\gamma(t)) - O_{\mathscr{I}}$. In questo approccio possiamo applicare la definizione della derivata rispetto ad un riferimento e vedere la velocità come un vettore in Σ_t dato da:

$$\mathbf{v}|_{\mathscr{I}}(t) = \left.\frac{d}{dt}\right|_{\mathscr{I}} (\gamma(t) - O_{\mathscr{I}}(t)) \,, \tag{1.45}$$

dove $O(t) = (\Pi_{\mathscr{I}} \restriction_{\Sigma_t})^{-1}(O_{\mathscr{I}})$ e $O_{\mathscr{I}} \in E_{\mathscr{I}}$.

La questione che ora ci interessa studiare è la relazione che intercorre tra l'operatore di derivazione $\frac{d}{dt}\big|_{\mathscr{I}}$ e l'operatore di derivazione $\frac{d}{dt}\big|_{\mathscr{I}'}$ se $\mathscr{I}$ è differente da $\mathscr{I}'$. La questione è di grande interesse fisico perché permette di costruire la cinematica relativa: permette di scrivere le velocità e l'accelerazione di un punto materiale in un riferimento quando sono note in un altro riferimento ed il moto relativo tra i due riferimenti è assegnato. Vale il seguente famoso teorema dovuto a Poisson.

Teorema 1.2. (**Formule di Poisson**.) *Siano $\mathscr{I}$ e $\mathscr{I}'$ due riferimenti e $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \in V_{\mathscr{I}}$, $\{\mathbf{e'}_1, \mathbf{e'}_2, \mathbf{e'}_3\} \in V_{\mathscr{I}'}$ due basi ortonormali nei rispettivi spazi delle traslazioni. Posto:*

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) := \frac{1}{2} \sum_{j=1}^{3} \mathbf{e'}_j(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e'}_j(t) \,, \tag{1.46}$$

vale:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} = \left.\frac{d}{dt}\right|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \,. \tag{1.47}$$

$\diamondsuit$

Dimostrazione. Rappresentiamo tutti vettori che consideriamo nello spazio $V_{\mathscr{I}}$. In tale spazio i vettori $\mathbf{e}_i$ non dipendono dal tempo, mentre lo dipendono i vettori $\mathbf{e'}_j$. Consideriamo allora che, per una generica curva differenziabile $t \mapsto \mathbf{u}(t)$:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{u}(t) = \left.\frac{d}{dt}\right|_{\mathscr{I}} \left(\sum_{j=1}^{3} u'^j(t) \mathbf{e'}_j(t) \right) = \sum_{j=1}^{3} \frac{du'^j}{dt} \mathbf{e'}_j(t) + \sum_{j=1}^{3} u'^j(t) \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e'}_j(t) \,.$$

Per terminare la dimostrazione è sufficiente provare che

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t) = \left(\frac{1}{2}\sum_{k=1}^{3} \mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\right) \wedge \mathbf{e}'_j(t)\,. \tag{1.48}$$

Usando la nota relazione generale $(\mathbf{a}\wedge\mathbf{b})\wedge\mathbf{c} = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{b}\cdot\mathbf{c})\mathbf{a}$ e il fatto che $\mathbf{e}'_r(t)\cdot\mathbf{e}'_s(t) = \delta_{rs}$, il secondo membro di questa presunta identità è scrivibile come

$$\frac{1}{2}\sum_{k=1}^{3}\left(\mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\right) \wedge \mathbf{e}'_j(t) = \frac{1}{2}\sum_{k=1}^{3}\left[\delta_{kj}\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t) - \left(\left(\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\right)\cdot \mathbf{e}'_j(t)\right)\mathbf{e}'_k(t)\right]\,.$$

Tuttavia

$$\left(\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\right)\cdot \mathbf{e}'_j(t) = \frac{d}{dt}(\mathbf{e}'_k(t)\cdot\mathbf{e}'_j(t)) - \mathbf{e}'_k(t)\cdot\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t) = \mathbf{0} - \mathbf{e}'_k(t)\cdot\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t)\,,$$

da cui

$$\frac{1}{2}\sum_{k=1}^{3}\left(\mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\right) \wedge \mathbf{e}'_j(t) = \frac{1}{2}\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t) + \frac{1}{2}\sum_{k=1}^{3}\left(\mathbf{e}'_k(t)\cdot\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t)\right)\mathbf{e}'_k(t)$$

cioé

$$\frac{1}{2}\sum_{k=1}^{3}\left(\mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\right) \wedge \mathbf{e}'_j(t) = \frac{1}{2}\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t) + \frac{1}{2}\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t) = \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_j(t)\,,$$

che è proprio (1.48). □

Osservazioni 1.7.
(1) La definizione di $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ non dipende in realtà dalle basi scelte negli spazi di quiete dei due riferimenti $\mathscr{I}$ e $\mathscr{I}'$, ma solo dai due riferimenti. Infatti la base scelta in $\mathscr{I}$ non gioca alcun ruolo nella (1.46) (si tenga conto anche di (1) nelle osservazioni 1.6). Per quanto riguarda la base degli $\mathbf{e}'_j$ scelta nello spazio delle traslazioni $V_{\mathscr{I}'}$ dello spazio di quiete $E_{\mathscr{I}}$ di $\mathscr{I}$, si osservi che vale quanto segue. Se $\mathbf{f}'_j = \sum_k R^k{}_j \mathbf{e}'_k$, con $j = 1,2,3$ definisce un'altra base ortonormale nello stesso spazio, allora la (1.46) riproduce comunque lo stesso vettore $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ valendo:

$$\sum_{j=1}^{3}\mathbf{f}'_j(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{f}'_j(t) = \sum_{i,j,k=1}^{3} R^k{}_j R^i{}_j \mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_i(t) = \sum_{i,k=1}^{3}\delta_{ki}\mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_i(t)$$

$$= \sum_{k=1}^{3}\mathbf{e}'_k(t) \wedge \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{e}'_k(t)\,,$$

dove abbiamo tenuto conto delle relazioni di ortonormalità $\sum_{j=1}^{3} R^k{}_j R^i{}_j = \delta_{ki}$ e dell'indipendenza dal tempo dei coefficienti $R^p{}_q$.

(2) Come al solito, i vettori $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)$ si possono vedere come appartenenti allo spazio delle traslazioni del riferimento $\mathscr{I}$, oppure $\mathscr{I}'$, oppure di un qualsiasi altro riferimento $\mathscr{I}''$, oppure – infine – dello spazio delle traslazioni di Σ_t. Questa possibilità deve essere tenuta in considerazione in tutto ciò che segue.

Vale la seguente proposizione che stabilisce, in particolare, la *legge di composizione* dei vettori $\boldsymbol{\omega}$.

Proposizione 1.2. **(Legge di composizione dei vettori $\boldsymbol{\omega}$).** *Siano $\mathscr{I}$, $\mathscr{I}'$ e $\mathscr{I}''$ tre sistemi di riferimento nello spaziotempo $\mathbb{V}^4$. Ad ogni fissato tempo $t \in \mathbb{R}$ valgono i seguenti fatti.*
(a) Legge di composizione*:*

$$\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}}(t) = \boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\,. \tag{1.49}$$

(b) Legge di inversione*:*

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = -\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'}(t)\,. \tag{1.50}$$

(c) Assolutezza della derivata di $\boldsymbol{\omega}$*:*

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \left.\frac{d}{dt}\right|_{\mathscr{I}'} \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\,. \tag{1.51}$$

$\diamondsuit$

Dimostrazione. (a) Valgono contemporaneamente le seguenti formule:

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{\mathscr{I}} &= \left.\frac{d}{dt}\right|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\wedge\,,\\
\left.\frac{d}{dt}\right|_{\mathscr{I}'} &= \left.\frac{d}{dt}\right|_{\mathscr{I}''} + \boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t)\wedge\,,\\
\left.\frac{d}{dt}\right|_{\mathscr{I}} &= \left.\frac{d}{dt}\right|_{\mathscr{I}''} + \boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}}(t)\wedge\,.
\end{aligned} \tag{1.52}$$

Inserendo la seconda nella prima si trova anche:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} = \left.\frac{d}{dt}\right|_{\mathscr{I}''} + \left(\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\right)\wedge\,,$$

che, per confronto con (1.52), fornisce, per ogni $t \in \mathbb{R}$:

$$\left(\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) - \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}''}(t)\right)\wedge \quad = 0\,.$$

Lo zero a secondo membro deve essere interpretato come *l'operatore nullo*. In altre parole, la formula di sopra si interpreta come segue. *Per ogni curva differenziabile $(a,b) \ni t \mapsto \mathbf{v}(t) \in V_t$, per ogni $t \in (a,b)$ vale:*

$$\left(\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) - \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}''}(t)\right)\wedge \mathbf{v}(t) = \mathbf{0}\,. \tag{1.53}$$

Se $t_0 \in (a,b)$, sia $\mathbf{u}_0 := (\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t_0) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t_0) - \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}''}(t_0))$. Possiamo facilmente costruire una curva differenziabile $\mathbb{R} \ni t \mapsto \mathbf{v}(t) \in V_t$ (basta lavorare in coordinate cartesiane ortonormali solidali con qualche riferimento) tale che, esattamente per $t = t_0$, valga $||\mathbf{v}(t_0)|| = c > 0$ e $\mathbf{v}(t_0) \perp \mathbf{u}_0$. La (1.53) implica per $t = t_0$ che

$$c\, ||\, \boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t_0) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t_0) - \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}''}(t_0)|| = 0\,.$$

Dato che $c > 0$ dobbiamo concludere che

$$\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}'}(t_0) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t_0) - \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}''}(t_0) = \mathbf{0}\,.$$

Dato che ciò vale per ogni $t_0 \in \mathbb{R}$, concludiamo che (a) è vera.
(b) Segue subito da (a) scegliendo $\mathscr{I}'' = \mathscr{I}$, essendo, come si prova immediatamente dalla definizione (1.46): $\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}} = \mathbf{0}$.
(c) Da (1.47) si ricava:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \left.\frac{d}{dt}\right|_{\mathscr{I}'} \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\,.$$

D'altra parte $\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'}$ e $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ sono paralleli (con segno opposto), per (b), per cui il loro prodotto vettoriale è nullo e (1.51) segue immediatamente. □

Esempi 1.3.
(1) Consideriamo due sistemi di riferimento $\mathscr{I}$ e $\mathscr{I}'$ con coordinate cartesiane ortonormali solidali rispettivamente t, x, y, z, con origine $O \in E_{\mathscr{I}}$ e assi $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in V_{\mathscr{I}}$, e t', x', y', z' con origine $O' \in E_{\mathscr{I}}$ e assi $\mathbf{e'}_1, \mathbf{e'}_2, \mathbf{e'}_3 \in V_{\mathscr{I}'}$. Supponiamo che la relazione tra questi due sistemi di coordinate sia:

$$x = x' \cos\phi(t) - y' \sin\phi(t) + X(t)\,, \tag{1.54}$$
$$y = x' \sin\phi(t) + y' \cos\phi(t) + Y(t)\,, \tag{1.55}$$
$$z = z' + Z(t)\,, \tag{1.56}$$

dove $\phi = \phi(t)$, $X = x(t)$, $Y = Y(t)$ e $Z = Z(t)$ sono funzioni differenziabili assegnate. Tenendo cono che le differenze di coordinate di un punto individuano coordinate di vettori a causa della struttura affine degli spazi di quiete, abbiamo immediatamente che la legge di trasformazione per le componenti di un vettore $u = \sum_i u^i \mathbf{e}_i = \sum_j u'^j \mathbf{e'}_j \in V_t$

$$u^1 = u'^1 \cos\phi(t) - u'^2 \sin\phi(t)\,, \tag{1.57}$$
$$u^2 = u'^1 \sin\phi(t) + u'^2 \cos\phi(t)\,, \tag{1.58}$$
$$u^3 = u'^3\,, \tag{1.59}$$

Possiamo infine ricavare la legge di trasformazione dei versori, notando che $\mathbf{e'}_k = \sum_j \delta^j_k \mathbf{e'}_j$, da cui

$$\mathbf{e'}_1 = \cos\phi(t)\, \mathbf{e}_1 + \sin\phi(t)\, \mathbf{e}_2\,, \tag{1.60}$$
$$\mathbf{e'}_2 = -\sin\phi(t)\, \mathbf{e}_1 + \cos\phi(t)\, \mathbf{e}_3\,, \tag{1.61}$$
$$\mathbf{e'}_3 = \mathbf{e}_3\,. \tag{1.62}$$

A questo punto siamo in grado di calcolare $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ direttamente dalla (1.46) ottenendo:

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \frac{d\phi(t)}{dt}\,\mathbf{e}_3\,. \tag{1.63}$$

Se ammettiamo che entrambe le terne di assi siano destrorse, notiamo che quando $\phi(t) = \nu t$ con $\nu > 0$ il moto della terna primata rispetto a quella non primata è una rotazione uniforme in senso antiorario (se guardata dall'alto) e $\boldsymbol{\omega} = \nu\mathbf{e}_3$, che è diretto verso l'alto, coincide con il vettore velocità angolare delle trattazioni elementari.
(2) Consideriamo il caso di una trasformazione di coordinate generica tra sistemi di coordinate cartesiane ortogonali solidali con $\mathscr{I}$ e $\mathscr{I}'$ rispettivamente. Sappiamo dall'esercizio 1.1.**2** che tale legge di trasformazione è :

$$\begin{cases} t &= t' + c\,, \\ x^i &= \displaystyle\sum_{j=1}^{3} R^i{}_j(t)\,x'^j + c^i(t)\,,\ \ i = 1,2,3\,, \end{cases} \tag{1.64}$$

dove $c \in \mathbb{R}$ è una costante, le funzioni $\mathbb{R} \ni t \mapsto c^i(t)$ e $\mathbb{R} \ni t \mapsto R^i{}_j(t)$ sono C^∞ e, per ogni $t \in \mathbb{R}$, la matrice di coefficienti $R^i{}_j(t)$ è una matrice ortogonale reale 3×3.
La corrispondente legge di trasformazione per i versori è , con ovvie notazioni,

$$\mathbf{e}'_j = \sum_{i=1}^{3} R^i{}_j(t)\mathbf{e}_i\,.$$

Il calcolo di $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ con la (1.46) fornisce:

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \frac{1}{2}\sum_{j=1}^{3} R^i{}_j(t)\frac{dR^k{}_j(t)}{dt}\,\mathbf{e}_i \wedge \mathbf{e}_k\,.$$

Assumendo la base ortonormale $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ di tipo destrorso vale:

$$\mathbf{e}_i \wedge \mathbf{e}_k = \sum_{h=1}^{3} \epsilon_{ikh}\,\mathbf{e}_h\,, \tag{1.65}$$

dove abbiamo introdotto il **simbolo di Ricci** ϵ_{ijk} che vale: 1 se ijk è una permutazione ciclica di 123, oppure -1 se ijk è una permutazione non ciclica di 123, 0 altrimenti. In definitiva, se valgono le leggi di trasformazione (1.64), il vettore $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)$ è scrivibile come:

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \frac{1}{2}\sum_{j,h=1}^{3} \epsilon_{ikh} R^i{}_j(t)\frac{dR^k{}_j(t)}{dt}\,\mathbf{e}_h\,. \tag{1.66}$$

(3) Un disco rigido è vincolato a stare in un piano verticale Π. In Π il disco ruota attorno al proprio asse di moto uniforme con velocità angolare $d\theta/dt = \nu$ costante, con θ angolo delle

coordinate polari piane in Π di origine O data dal centro del disco. Il piano Π ruota, rispetto al riferimento $\mathscr{I}$, attorno alla retta verticale $z \in \Pi$, che non interseca il disco, con la legge $\phi(t) = \sin(\lambda t)$, dove ϕ è il solito angolo polare nel piano normale a z. Si considera un riferimento $\mathscr{I}'$ solidale con il disco (cioé nel quale il disco è sempre in quiete) e si vuole conoscere $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$.
Il problema si risolve componendo i vettori $\boldsymbol{\omega}$. In un riferimento $\mathscr{I}''$ solidale con il piano Π, il riferimento $\mathscr{I}'$, solidale con il disco, ha un vettore $\boldsymbol{\omega}$, per l'esempio (1) dato da: $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}''} = \nu \mathbf{n}_\Pi$, dove $\mathbf{n}_\Pi$ è il versore normale al piano e *disposto in modo tale che la rotazione del disco avvenga in senso antiorario guardando il disco dalla punta di* $\mathbf{n}_\Pi$. Al solito questa orientazione dell'asse di rotazione si dice *positivo* rispetto al verso in cui l'angolo di rotazione cresce. Il riferimento $\mathscr{I}''$ è in rotazione attorno all'asse z nel giudizio di $\mathscr{I}$. Usando nuovamente l'esempio (1), abbiamo che $\boldsymbol{\omega}_{\mathscr{I}''}|_{\mathscr{I}} = \frac{d\phi}{dt}\mathbf{e}_z = \lambda\cos(\lambda t)\mathbf{e}_z$, dove $\mathbf{e}_z$ è il versore parallelo a r ed orientato positivamente rispetto all'angolo ϕ. Usando la legge di composizione dei vettori $\boldsymbol{\omega}$ si ha infine:

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \lambda\cos(\lambda t)\mathbf{e}_z + \nu\mathbf{n}_\Pi(t)\ .$$

Volendo si può esplicitare la dipendenza temporale di $\mathbf{n}_\Pi$, tenendo conto che esso ruota solidalmente al piano Π. Per fare ciò si considerino altri due versori mutuamente ortogonali $\mathbf{e}_x$, $\mathbf{e}_y$, tali che formino con $\mathbf{e}_z$ (nell'ordine detto), una terna ortonormale destrorsa solidale con $\mathscr{I}$. In questo caso, con le ipotesi fatte: $\mathbf{n}_\Pi(t) = \cos(\lambda t + \alpha)\,\mathbf{e}_x + \sin(\lambda t + \alpha)\,\mathbf{e}_y$ per qualche costante α. In definitiva:

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) = \nu\cos(\lambda t + \alpha)\,\mathbf{e}_x + \nu\sin(\lambda t + \alpha)\,\mathbf{e}_y + \lambda\cos(\lambda t)\mathbf{e}_z\ .$$

Esercizi 1.4.

1. Si considerino due sistemi di coordinate ortonormali destrorse solidali rispettivamente con i riferimenti $\mathscr{I}$ e $\hat{\mathscr{I}}$ e con assi, rispettivamente, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ e $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, $\hat{\mathbf{e}}_3$ ed origini, rispettivamente, O e $\hat{O}$. Si supponga che $O = \hat{O}$ (in ogni Σ_t) ad ogni istante. Introduciamo i cosiddetti **angoli di Eulero**: $\psi \in [0, 2\pi)$, $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$ che individuano la terna di $\mathscr{I}$ partendo da quella di $\hat{\mathscr{I}}$ con la seguente procedura.
(i) Si ruota la terna associata ad $\hat{\mathscr{I}}$ attorno ad $\hat{\mathbf{e}}_3$, di un angolo ψ, in senso positivo rispetto a $\hat{\mathbf{e}}_3$. Alla fine della rotazione l'asse $\hat{\mathbf{e}}_1$ si è spostato sulla cosiddetta **retta dei nodi**, individuata dal versore $\mathbf{N}$ nel piano normale a $\hat{\mathbf{e}}_3$.
(ii) La terna ottenuta in (i) viene ora ruotata attorno alla retta dei nodi in senso positivo di un angolo θ. In questo modo l'asse $\hat{\mathbf{e}}_3$ si sposta su un nuovo asse, che indichiamo con $\mathbf{e}_3$.
(iii) La terna ottenuta in (ii) viene ruotata di un angolo ϕ in senso positivo attorno a $\mathbf{e}_3$. In questo modo $\mathbf{N}$ si sposta sul nuovo asse, che indichiamo con $\mathbf{e}_1$.
La terna finale ottenuta con questa procedura è quella associata a $\mathscr{I}$: $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. $\mathbf{e}_3$ e $\mathbf{e}_1$ sono gli assi ottenuti sopra, in (ii) e (iii) rispettivamente, e $\mathbf{e}_2$ è l'asse raggiunto dall'asse $\hat{\mathbf{e}}_2$ della terna iniziale alla fine delle tre rotazioni. Per costruzione $\mathbf{e}_2 = \mathbf{e}_3 \wedge \mathbf{e}_1$.
Si esprima terna associata ad $\mathscr{I}$ in funzione di quella associata a $\hat{\mathscr{I}}$ usando gli angoli di Eulero.

Provare cioé che vale:

$$
\begin{aligned}
\mathbf{e}_1 &= (\cos\phi\cos\psi - \sin\phi\sin\psi\cos\theta)\,\hat{\mathbf{e}}_1 + (\cos\phi\sin\psi + \sin\phi\cos\theta\cos\psi)\,\hat{\mathbf{e}}_2 + \sin\phi\sin\theta\,\hat{\mathbf{e}}_3\,,\\
\mathbf{e}_2 &= -(\sin\phi\cos\psi + \cos\phi\sin\psi\cos\theta)\,\hat{\mathbf{e}}_1 + (-\sin\phi\sin\psi + \cos\phi\cos\theta\cos\psi)\,\hat{\mathbf{e}}_2 + \cos\phi\sin\theta\,\hat{\mathbf{e}}_3\,,\\
\mathbf{e}_3 &= \sin\theta\,\sin\psi\,\hat{\mathbf{e}}_1 - \sin\theta\,\cos\psi\,\hat{\mathbf{e}}_2 + \cos\theta\,\hat{\mathbf{e}}_3\,.
\end{aligned}
$$

2. In riferimento all'esercizio precedente, provare che:

$$
\begin{aligned}
\hat{\mathbf{e}}_1 &= (\cos\phi\cos\psi - \sin\phi\sin\psi\cos\theta)\,\mathbf{e}_1 - (\sin\phi\cos\psi + \cos\phi\sin\psi\cos\theta)\,\hat{\mathbf{e}}_2 + \sin\theta\,\sin\psi\,\mathbf{e}_3\,,\\
\hat{\mathbf{e}}_2 &= (\cos\phi\sin\psi + \sin\phi\cos\theta\cos\psi)\,\mathbf{e}_1 + (-\sin\phi\sin\psi + \cos\phi\cos\theta\cos\psi)\,\mathbf{e}_2 - \sin\theta\,\cos\psi\,\mathbf{e}_3\,,\\
\hat{\mathbf{e}}_3 &= \sin\phi\sin\theta\,\mathbf{e}_1 + \cos\phi\sin\theta\,\mathbf{e}_2 + \cos\theta\,\mathbf{e}_3\,.
\end{aligned}
$$

3. In riferimento ai due esercizi precedenti, ammettiamo che la terna $\mathscr{I}$ abbia un moto determinato, rispetto alla terna solidale con $\hat{\mathscr{I}}$, da funzioni note $\phi = \phi(t), \psi = \psi(t), \theta = \theta(t)$. Si esprima $\boldsymbol{\omega}_{\hat{\mathscr{I}}}|_{\mathscr{I}}$ in funzione delle derivate temporali degli angoli di Eulero e degli angoli di Eulero stessi nella base solidale con $\mathscr{I}$. Verificare che, indicando con il punto le derivate temporali degli angoli di Eulero:

$$
\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = (\dot{\psi}\,\sin\theta\sin\phi + \dot{\theta}\,\cos\phi)\,\mathbf{e}_1 + (\dot{\psi}\,\sin\theta\cos\phi - \dot{\theta}\,\sin\phi)\,\mathbf{e}_2 + (\dot{\phi}\,+\dot{\psi}\,\cos\theta)\,\mathbf{e}_3\,.
$$

***4**. Consideriamo $R \in SO(3)$ come operatore che agisce sui vettori di $\mathbb{E}^3$. Per ogni R detto esiste sempre un versore $\mathbf{u}$ tale che $R(\mathbf{u}) = \mathbf{u}$, tale versore è chiamato **asse di rotazione di** R. (In riferimento a coordinate cartesiane ortonormali R è individuato da una matrice a coefficienti reali. Si può sempre pensare tale matrice come caso particolare di una matrice complessa unitaria a coefficienti reali. Dal teorema spettrale è noto che ogni matrice unitaria è diagonalizzabile e i suoi autovalori sono della forma e^{ic} per qualche $c \in \mathbb{R}$. Dato che il polinomio caratteristico della matrice considerata è reale, gli autovalori devono essere: $\lambda_1 = e^{ib} \in \mathbb{R}$, cioé $\lambda_1 = \pm 1$, unitamente a due autovalori complessi coniugati: $\lambda_2 = e^{ia}$ e $\lambda_3 = e^{-ia}$. Dato che il determinante della matrice è pari al prodotto $\lambda_1\lambda_2\lambda_3$ e sappiamo che, per $R \in SO(3)$, tale determinante vale 1, facendo tutti i casi possibili concludiamo facilmente che $\lambda_1 = 1$.)
Si dimostri che se la relazione tra i versori di due riferimenti è , all'istante t, $\mathbf{e}'_k(t) = R_t(\,\mathbf{e}_k)$ per $k = 1, 2, 3$, allora vale la relazione notevole:

$$
R_t(\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)) = \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\,.
$$

In altre parole, il vettore $\boldsymbol{\omega}$ è un **asse istantaneo di rotazione**.

1.4.2 Velocità ed accelerazione al variare del riferimento.

Useremo ora le formule di Poisson, ed il vettore $\boldsymbol{\omega}$, per scrivere le formule che connettono le velocità e le accelerazioni di un punto materiale calcolate rispetto a due differenti riferimenti.
Consideriamo due riferimenti $\mathscr{I}$ e $\mathscr{I}'$ con sistemi di coordinate cartesiane ortonormali solidali rispettivamente con origine O e assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, e con origine O' e assi $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$. Sia poi una

linea di universo γ descrivente il moto di un punto materiale nello spaziotempo. Nel seguito, come consuetudine, $P(t) := \gamma(t)$. Vale allora banalmente:

$$P - O = P - O' \; + \; O' - O \, . \tag{1.67}$$

Usando l'identità di Poisson

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} = \left.\frac{d}{dt}\right|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \, ,$$

tenendo conto dell'espressione (1.45) per la velocità di P rispetto a $\mathscr{I}$ e $\mathscr{I}'$, derivando (1.67) nel tempo, giungiamo alla relazione (omettiamo di scrivere esplicitamente il tempo t per non appesantire la notazione):

$$\mathbf{v}_P|_{\mathscr{I}} = \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + \left.\frac{d}{dt}\right|_{\mathscr{I}} (O' - O) \, ,$$

ossia

$$\mathbf{v}_P|_{\mathscr{I}} = \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + \mathbf{v}_{O'}|_{\mathscr{I}} \, . \tag{1.68}$$

Conviene riscrivere l'equazione di sopra come:

$$\left\{ \begin{array}{ll} \mathbf{v}_P|_{\mathscr{I}} = & \mathbf{v}_P|_{\mathscr{I}'} + \mathbf{v}_P^{(tr)} \, , \\ \mathbf{v}_P^{(tr)} := & \mathbf{v}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') \, , \end{array} \right. \tag{1.69}$$

dove $\mathbf{v}_P^{(tr)}$ è la **velocità di trascinamento** di $\mathscr{I}'$ rispetto a $\mathscr{I}$ in P. Si tratta della velocità che $\mathscr{I}$ assegna al punto P quando è pensato fermo in $\mathscr{I}'$. Derivando ancora (1.68) rispetto a $\mathscr{I}$:

$$\mathbf{a}_P|_{\mathscr{I}} = \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{v}_P|_{\mathscr{I}'} + \left.\frac{d}{dt}\right|_{\mathscr{I}} \big(\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')\big) + \left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{v}_{O'}|_{\mathscr{I}} \, .$$

Applicando (1.47) al primo ed al secondo addendo a secondo membro, $\mathbf{a}_P|_{\mathscr{I}}$ uguaglia:

$$\mathbf{a}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} + \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \big(\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')\big) + \mathbf{a}_{O'}|_{\mathscr{I}} \, ,$$

dove $\dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} := d\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}/dt = d\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'}/dt$. Raccogliendo i vari termini:

$$\left\{ \begin{array}{ll} \mathbf{a}_P|_{\mathscr{I}} = & \mathbf{a}_P|_{\mathscr{I}'} + \mathbf{a}_P^{(tr)} + \mathbf{a}_P^{(C)} \, , \\ \mathbf{a}_P^{(tr)} := & \mathbf{a}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')) + \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') \, , \\ \mathbf{a}_P^{(C)} := & 2\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} \, , \end{array} \right. \tag{1.70}$$

dove $\mathbf{a}_P^{(tr)}$ è l'**accelerazione di trascinamento** di $\mathscr{I}'$ rispetto a $\mathscr{I}$ in P. Si tratta della accelerazione che $\mathscr{I}$ assegna al punto P quando quest'ultimo è pensato come in quiete (con accelerazione nulla) con $\mathscr{I}'$. $\mathbf{a}_P^{(C)}$ è infine la ben nota **accelerazione di Coriolis**.

Osservazioni 1.8.
(1) Precisiamo che si può provare che $\mathbf{v}_P^{(tr)}$, $\mathbf{a}_P^{(tr)}$ e $\mathbf{a}_P^{(C)}$ *non* dipendono dalla scelta di O'. Riguardo ad $\mathbf{a}_P^{(C)}$ non c'è , in realtà , nulla da provare! Dimostriamo l'indipendenza dalla scelta di O' per la velocità di trascinamento. Per l'accelerazione di trascinamento si può procedere analogamente. Se $O'_1 \neq O'$ è in quiete con $\mathscr{I}'$, da (1.68), ponendo $P = O'_1$ segue che vale l'identità : $\mathbf{v}_{O'_1}|_{\mathscr{I}} = \mathbf{0} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (O'_1 - O') + \mathbf{v}_{O'}|_{\mathscr{I}}$, cioé :

$$\mathbf{v}_{O'_1}|_{\mathscr{I}} - \mathbf{v}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (O' - O'_1) = \mathbf{0}$$

Usando la seconda delle (1.69), valutando la velocità di trascinamento $\mathbf{v}_P^{(tr1)}$ dello stesso P, ma riferendosi a O'_1 invece che ad O', si ricava subito che:

$$\mathbf{v}_P^{(tr1)} - \mathbf{v}_P^{(tr)} = \mathbf{v}_{O'_1}|_{\mathscr{I}} - \mathbf{v}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (O' - O'_1) = \mathbf{0}\,,$$

per cui $\mathbf{v}_P^{(tr1)} = \mathbf{v}_P^{(tr)}$.
(2) Se P è un punto materiale in quiete in $\mathscr{I}'$ (con origine O') la cui curva di universo è differenziabile, vale, usando lo sviluppo di Taylor nel tempo in (1.69), la relazione che esprime la posizione di P nello spazio di quiete di $\mathscr{I}$ al tempo $t + \Delta t$ in funzione di quella al tempo t :

$$P(t + \Delta t) = P(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (P(t) - O'(t))\Delta t + \mathbf{v}_{O'}|_{\mathscr{I}}(t)\Delta t + \Delta t\mathbf{O}(\Delta t) \tag{1.71}$$

dove $\mathbf{O}(\Delta t)$ è una funzione infinitesima per $\Delta t \to 0$. La (1.71) mostra che, *al primo ordine in* Δt, il moto di P in $\mathscr{I}$ è un moto dato da una traslazione con velocità $\mathbf{v}_{O'}|_{\mathscr{I}}(t)$, unitamente ad una rotazione destrorsa attorno ad $O'(t)$ rispetto all'asse $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)$, di un angolo $|\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)|\Delta t$. In questo modo si può pensare $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ come un *asse istantaneo di rotazione* (esiste una seconda interpretazione più matematica dovuta all'esercizio 1.4.4).

Capitolo 2

Dinamica del punto e dei sistemi di punti materiali.

La meccanica classica si spinge oltre alla cinematica occupandosi di **dinamica**, ossia del problema della determinazione del moto di un corpo quando sono assegnate le "condizioni iniziali" (lo stato di moto del sistema ad un tempo fissato) e le "interazioni" con altri corpi, intese come le cause del moto stesso. Le interazioni vengono descritte, in fisica classica, dal concetto di *forza* che sarà introdotto tra breve. Dal punto di vista matematico, le forze sono delle funzioni del moto del sistema. La determinazione del moto è ricondotta in ultimi termini alla soluzione di un *sistema di equazioni differenziali* costruito con le funzioni descriventi le forze e con particolari costanti fisiche associate ai punti materiali del sistema fisico, dette *masse* dei punti stessi.

2.1 Primo principio della dinamica.

Per poter introdurre la dinamica è necessario selezionare una classe privilegiata di sistemi di riferimento detti *inerziali*. Vediamo prima di tutto come tale classe viene determinata dal punto di vista fisico. L'idea di fondo è di studiare il moto dei corpi quando sono posti a distanza molto grande dagli altri corpi dell'universo. Ricordiamo che la nostra idealizzazione di corpo è quella del *punto materiale*, cioè un corpo fisico la cui struttura interna non sia rilevante e la cui evoluzione spaziotemporale sia descrivibile da una linea di universo.

2.1.1 Sistemi di riferimento inerziali.

Avendo a disposizione un unico punto materiale nell'universo e non avendo selezionato alcun riferimento privilegiato, non hanno alcun senso fisico proposizioni riguardanti lo stato di moto del punto: lo stato di moto del punto dipende dalla scelta del riferimento, ed il moto può essere fissato arbitrariamente scegliendo opportunamente il riferimento. Se consideriamo invece un

insieme di più di un punto materiale, in generale non è possibile trovare un riferimento in cui poter assegnare, rispetto ad esso, uno stato di moto scelto a piacimento a *ciascuno* dei punti materiali *contemporaneamente*.
Il **Principio d'inerzia** o **Primo principio della dinamica** dichiara quale sia lo stato di moto di un numero arbitrario di corpi quando si trovano a distanze grandissime, reciprocamente e da tutti gli altri corpi dell'universo. Punti materiali che soddisfano tale requisito di "lontananza" (in una porzione della loro linea di universo) sono detti **isolati** (in tale porzione di linea di universo). Proprio perché non è in generale possibile fissare lo stato di moto di tanti punti materiali contemporaneamente mediate una scelta opportuna del riferimento, il principio d'inerzia ha un contenuto fisico altamente non banale.

C1. Principio d'inerzia.
Esiste una classe di sistemi di riferimento in ciascuno dei quali tutti *i punti materiali isolati si muovono con moto* **rettilineo uniforme**, *cioè a velocità costante in modulo, direzione e verso, dipendente dal punto materiale considerato. Tali sistemi di riferimento si dicono* **sistemi di riferimento inerziali**

Osservazioni.
(1) Ci siamo riferiti ad una classe di sistemi di riferimento inerziali e non ad uno solo perché, come sarà mostrato tra poco, se $\mathscr{I}$ è un sistema di riferimento inerziale e $\mathscr{I}$ e $\mathscr{I}'$ hanno moto relativo rettilineo uniforme allora $\mathscr{I}'$ stesso è inerziale.
(2) È chiaro che dal punto di vista pratico è molto difficile stabilire con certezza se un riferimento sia inerziale o no. I riferimenti inerziali posso comunque essere determinati a posteriori sviluppando la teoria fino a livelli più accessibili dal punto di vista sperimentale[1].
(3) L'esistenza di sistemi di riferimento inerziali è in realtà negata da sviluppi successivi della fisica: dalla Teoria della Relatività Generale. Tuttavia anche in tale ambito si vede che sistemi di riferimento con ottima approssimazione inerziali possono esser comunque definiti in regioni dell'universo relativamente piccole dal punto di vista cosmologico, ma più grandi dello stesso sistema solare, e contenenti una quantità di massa di ordini di grandezza inferiori alla massa totale dell'universo.
A tal fine supporremo di selezionare nell'universo una regione spaziale Ω come detto in cui svilupperemo la teoria della meccanica classica e quando parleremo di corpi **sufficientemente lontani** o **isolati** intenderemo corpi in Ω che distino l'uno dall'altro una distanza pari al raggio della regione stessa o più.
(4) Si noti che, nell'enunciare il primo principio della dinamica, abbiamo accuratamente evitato di parlare di *forze* per evitare un circolo vizioso, in un certo senso "tradizionale", che purtroppo

[1] Per esempio, una volta sviluppata la dinamica ed introdotto il concetto di centro di massa, si prova che preso un insieme di corpi sufficientemente lontani dagli altri ma non isolati tra di essi, il centro di massa del sistema deve essere in quiete in un riferimento inerziale. Se uno dei corpi del sistema C ha massa molto maggiore degli altri, il suo moto deve essere vicino a quello del centro di massa, per cui c'è un riferimento inerziale in cui il corpo C è con buona approssimazione in quiete. Tale situazione si ha per il sistema solare, dove C è il sole.

si trova in diversi testi.

2.1.2 Trasformazioni di Galileo.

Studieremo ora il moto relativo tra riferimenti inerziali. Per prima cosa stabiliamo quando un riferimento è in moto *rettilineo uniforme rispetto ad un altro riferimento.*

Definizione 2.1. Consideriamo due sistemi di riferimento $\mathscr{I}$ e $\mathscr{I}'$, diremo che $\mathscr{I}$ è in moto **rettilineo uniforme rispetto al riferimento** $\mathscr{I}'$ se la velocità di trascinamento $\mathbf{v}_P^{(tr)}$ di $P \in E_{\mathscr{I}}$ rispetto a $\mathscr{I}'$ soddisfa:
(i) è indipendente da $P \in E_{\mathscr{I}}$,
(ii) è costante nel tempo.
Tale velocità di trascinamento, indicata con $\mathbf{v}_{\mathscr{I}}|_{\mathscr{I}'}$ è detta **velocità di trascinamento di $\mathscr{I}$ rispetto a $\mathscr{I}'$.** $\diamondsuit$

Abbiamo la seguente utile proposizione che caratterizza il moto rettilineo uniforme tra riferimenti.

Proposizione 2.1. *Siano $\mathscr{I}$ e $\mathscr{I}'$ due sistemi di riferimento in $\mathbb{V}^4$. Valgono i seguenti fatti.*
(a) *Se $\mathscr{I}$ è in moto rettilineo uniforme rispetto a $\mathscr{I}'$, siano t, x^1, x^2, x^3, t', x'^1, x'^2, x'^3 coordinate cartesiane ortonormali solidali rispettivamente con $\mathscr{I}$ e $\mathscr{I}'$ con origini, rispettivamente $O \in E_{\mathscr{I}}$, $O' \in E_{\mathscr{I}'}$ e assi, rispettivamente $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in V_{\mathscr{I}}$, $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 \in V_{\mathscr{I}'}$, allora la legge di trasformazione tra questi sistemi di coordinate è una* **trasformazione di Galileo**, *cioé una trasformazione da $\mathbb{R}^4$ in $\mathbb{R}^4$ di forma:*

$$\begin{cases} t' = & t + c\,, \\ x'^i = & c^i + tv^i + \sum_{j=1}^{3} R^i{}_j x^j\,, \quad i=1,2,3. \end{cases} \tag{2.1}$$

dove $c \in \mathbb{R}$, $c^i \in \mathbb{R}$ e $v^i \in \mathbb{R}$ sono costanti, e i coefficienti costanti $R^i{}_j$ definiscono una matrice ortogonale reale 3×3. Vale ulteriormente

$$\mathbf{v}_{\mathscr{I}}|_{\mathscr{I}'} := \sum_{i=1}^{3} v^i \mathbf{e}'_i\,. \tag{2.2}$$

(b) *Se viceversa, rispetto a una coppia di sistemi di coordinate cartesiane ortonormali solidali, rispettivamente con $\mathscr{I}$ e $\mathscr{I}'$, valgono le relazioni (2.1), allora $\mathscr{I}$ è in moto rettilineo uniforme rispetto a $\mathscr{I}'$ con velocità di trascinamento data da (2.2).*
(c) *Se il riferimento $\mathscr{I}$ è in moto rettilineo uniforme rispetto al riferimento $\mathscr{I}'$, allora $\mathscr{I}'$ è in moto rettilineo uniforme rispetto a $\mathscr{I}$ e $\mathbf{v}_{\mathscr{I}}|_{\mathscr{I}'} = -\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}}$.*
(d) *Se $\mathscr{I}$ è in moto rettilineo uniforme rispetto a $\mathscr{I}'$ e $\mathscr{I}'$ è in moto rettilineo uniforme rispetto*

a $\mathscr{I}''$, allora $\mathscr{I}$ è in moto rettilineo uniforme rispetto a $\mathscr{I}''$.
(e) *Se $\mathscr{I}$ è in moto rettilineo uniforme rispetto a $\mathscr{I}'$ allora $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$, $\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'}$, l'accelerazione di trascinamento e di Coriolis di $\mathscr{I}'$ rispetto a $\mathscr{I}$ e le analoghe accelerazioni di $\mathscr{I}$ rispetto a $\mathscr{I}'$ sono tutte nulle ad ogni istante.* $\diamondsuit$

Dimostrazione. (a) Nel caso generale, la trasformazione di coordinate tra i due riferimenti $\mathscr{I}$ e $\mathscr{I}'$ deve avere la forma (vedi esercizio 1.1.**2**):

$$\begin{cases} t' = & t + c\,, \\ x'^i = & \displaystyle\sum_{j=1}^{3} R^i{}_j(t)\, x^j + c^i(t)\,, \;\; i = 1,2,3\,, \end{cases} \tag{2.3}$$

dove le funzioni sono tutte C^∞ e i coefficienti $R^i{}_j(t)$, per ogni $t \in \mathbb{R}$ definiscono una matrice ortogonale. La linea di universo del punto $O \in E_{\mathscr{I}}$, corrispondente a $(0,0,0)$ per $\mathscr{I}$, ha quindi equazione $x'^i(t') = c^i(t'-c)$ per $\mathscr{I}'$. Derivando nel tempo, per definizione si devono ottenere le componenti della velocità di trascinamento di O rispetto a $\mathscr{I}'$ espresse nella base $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 \in V_{\mathscr{I}'}$. Per ipotesi queste componenti non devono dipendere dal tempo. Concludiamo che, indicando con v^i le costanti $dc^i(t'-c)/dt'$, (2.3) si riscrive:

$$\begin{cases} t' = & t + c\,, \\ x'^i = & \displaystyle\sum_{j=1}^{3} R^i{}_j(t)\, x^j + v^i t + c^i\,, \;\; i = 1,2,3\,, \end{cases} \tag{2.4}$$

dove le c^i sono ulteriori costanti. La velocità di trascinamento del generico punto $P \in E_{\mathscr{I}}$ di coordinate (x^1, x^2, x^3) per $\mathscr{I}$, avrà , dalla (2.4), componenti $v^i + \sum_j x^j dR^i{}_j(t)/dt$ rispetto alla base $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$. Per ipotesi queste componenti non possono dipendere dal punto, cioé dalle coordinate x^i pertanto, derivando rispetto alla generica coordinata x^k, otteniamo che devono valere le condizioni $dR^i{}_k(t)/dt = 0$ per ogni scelta di i e k e per ogni tempo t. Concludiamo che vale (2.2).
(b) Si consideri una linea di universo $t \mapsto P(t) \in \mathbb{V}^4$ (con $P(t) \in \Sigma_t$) che descriva l'evoluzione di un punto P dello spazio di quiete di $\mathscr{I}$ individuato nelle coordinate di $\mathscr{I}$ da $(y^1, y^2, y^3) \in \mathbb{R}^3$. In tali coordinate (solidali con $\mathscr{I}$) la linea di universo sarà quindi descritta da $x^i(t) = y^i$ costante e $t = t$. Nel riferimento $\mathscr{I}'$ la stessa linea di universo sarà descritta da

$$x'^i(t') = c^i + (t'-c)v^i + \sum_j R^i{}_j y^j\,.$$

Le coordinate c^i determinano un punto Q in $E_{\mathscr{I}'}$ a $t' = c$. Derivando nel tempo, rispetto a $\mathscr{I}'$, il vettore

$$P(t) - Q := \sum_i (x'^i(t') - c^i)\mathbf{e}'_i$$

si ottiene la velocità di del punto P rispetto a $\mathscr{I}'$. Il risultato è banalmente il vettore costante nel tempo ed indipendente da $P \in E_{\mathscr{I}}$: $\sum_{i=1}^3 v^i \mathbf{e}'_i$.

(c) Nelle ipotesi fatte vale (2.1) in coordinate solidali con $\mathscr{I}'$ e $\mathscr{I}$ rispettivamente. Possiamo invertire (2.1) ottenendo

$$\begin{cases} t = & t' - c\,, \\ x^j = & \displaystyle\sum_{i=1}^{3} (R^{-1})^j{}_i \left(cv^i - c^i\right) - t' \sum_{i=1}^{3} (R^{-1})^j{}_i v^i + \sum_{i=1}^{3} (R^{-1})^j{}_i x'^i\,, \quad \text{i=1,2,3.} \end{cases} \tag{2.5}$$

Per il punto (b), $\mathscr{I}'$ è in moto rettilineo uniforme rispetto a $\mathscr{I}$ e

$$\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}} = -\sum_{i,j=1}^{3} (R^{-1})^j{}_i v^i\, \mathbf{e}_j\,.$$

Essendo per costruzione $\sum_{j=1}^{3} (R^{-1})^j{}_i\, \mathbf{e}_j = \mathbf{e}'_i$ si ha che vale (2.2).
(d) Si prova immediatamente scegliendo tre sistemi di coordinate cartesiane ortonormali solidali rispettivamente con i tre riferimenti. La trasformazione tra le coordinate associate a $\mathscr{I}$ e $\mathscr{I}'$ e la trasformazione tra le coordinate associate a $\mathscr{I}'$ e $\mathscr{I}''$ devono essere della forma (2.1) per (a). La trasformazione tra le coordinate associate a $\mathscr{I}$ e $\mathscr{I}''$ si ottiene allora componendo le precedenti trasformazioni. Per verifica diretta si trova subito che la trasformazione ottenuta è ancora di tipo (2.1) per cui, da (b), $\mathscr{I}$ è in moto rettilineo uniforme per $\mathscr{I}''$.
(e) Segue per computo diretto da (2.1) e da (c). □

La trasformazione (2.1) è individuata dalla quaterna $(c, \vec{c}, \vec{v}, R) \in \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times O(3)$, dove $\vec{c} := (c^1, c^2, c^3)$, $\vec{v} := (v^1, v^2, v^3)$ e R denota la matrice del gruppo $O(3)$ delle matrici ortogonali reali 3×3, individuata dai coefficienti $R^i{}_j$. La quaterna corrisponde a 10 parametri reali (si tenga conto che ogni matrice di $O(3)$ è determinata da 3 numeri reali). Dato che la composizione di funzioni è associativa, che la composizione di trasformazioni di Galileo è ancora una trasformazione di Galileo, l'inversa di una trasformazione di Galileo è una trasformazione di Galileo ed, infine, la trasformazione di Galileo con parametri $(0, \vec{0}, \vec{0}, I)$ è la trasformazione identità, concludiamo che l'insieme delle trasformazioni (2.1), viste come trasformazioni da $\mathbb{R}^4$ a $\mathbb{R}^4$, al variare di $(c, \vec{c}, \vec{v}, R) \in \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times O(3)$, ha la struttura algebrica di gruppo rispetto alla composizione di funzioni. Questo gruppo si chiama **gruppo di Galileo**.

Osservazioni 2.1.
(1) Attraverso le coordinate cartesiane ortonormali solidali con ogni fissato riferimento *inerziale* $\mathscr{I}$ è possibile dotare lo spaziotempo di una *struttura affine* privilegiata. La ragione è nella forma generale delle equazioni che definiscono il gruppo di Galileo (2.4). Se fissiamo un riferimento inerziale $\mathscr{I}$ e consideriamo due eventi p e q, possiamo associare alla coppia (p, q) un vettore $p - q$ definito dalle componenti $(t_p - t_q, x_p - x_q, y_p - y_q, z_p - z_q)$, dove abbiamo introdotto un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$. Se passiamo ad un altro riferimento *inerziale* $\mathscr{I}'$ e proviamo ad associare alla stessa coppia (p, q) un vettore $p - q$ definito dalle componenti $(t'_p - t'_q, x'_p - x'_q, y'_p - y'_q, z'_p - z'_q)$, dove abbiamo introdotto un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}'$, scopriamo le componenti di tale vettore si trasformano

secondo una trasformazione affine passando al precedente riferimento. Questo è dovuto alla non dipendenza temporale dei coefficienti $R^i{}_j$ e c^i in (2.4) che appare cambiando riferimento ed al fatto che siamo liberi di scgliere gli eventi p e q non contemporanei. In definitiva: i sistemi di coordinate cartesiani solidali con i riferimenti inerziali nello spaziotempo, definiscono il vettore $p-q$ in maniera indipendente dal riferimento inerziale, dando luogo ad una struttura affine sullo spaziotempo individuata dall'applicazione che associa ad ogni coppia di eventi (p,q) il vettore $p-q$.
(2)* Con un ampliamento del formalismo ed introducendo la *connessione affine* su $\mathbb{V}^4$ individuata dalla struttura affine descritta sopra, si può verificare che le *geodetiche* di tale connessione che soddisfano l'ulteriore requisito $\langle V, dT\rangle \neq 0$ ovunque, dove V è il vettore tangente in ogni evento alla geodetica, sono nient'altro che i *moti inerziali dei punti materiali.* Il tempo assoluto è un parametro affine di tali geodetiche. In questo senso la meccanica classica e le teorie relativistiche sono più vicine di quanto si possa immaginare.

Esercizi 2.1.

1. Siano assegnati tre sistemi di riferimento $\mathscr{I}$, $\mathscr{I}'$ e $\mathscr{I}''$ tali che, rispetto a coordinate ortonormali solidali con ciascuno di essi, la trasformazione di coordinate da $\mathscr{I}'$ a $\mathscr{I}$ è di Galileo ed è individuata dalla quadrupla $(c_1, \vec{c_1}, \vec{v_1}, R_1)$, la trasformazione di coordinate da $\mathscr{I}''$ a $\mathscr{I}'$ è di Galileo individuata dalla quadrupla $(c_2, \vec{c_2}, \vec{v_2}, R_2)$. Provare che la trasformazione di coordinate da $\mathscr{I}''$ a $\mathscr{I}$ ottenuta componendo le precedenti due è la trasformazione di Galileo individuata dalla quadrupla:

$$(c_1 + c_2,\ R_2\vec{c_1} + c_1\vec{v_2} + \vec{c_2},\ R_2\vec{v_1} + \vec{v_2},\ R_2R_1)\ .$$

2. Provare che $\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times O(3)$ dotato della legge di composizione

$$(c_2, \vec{c_2}, \vec{v_2}, R_2) \circ (c_1, \vec{c_1}, \vec{v_1}, R_1) := (c_1 + c_2,\ R_2\vec{c_1} + c_1\vec{v_2} + \vec{c_2},\ R_2\vec{v_1} + \vec{v_2},\ R_2R_1)\ ,$$

ha la struttura di gruppo dove, in particolare, l'elemento neutro è dato da $(0, \vec{0}, \vec{0}, I)$ e l'elemento inverso è

$$(c, \vec{c}, \vec{v}, R)^{-1} = \big(-c,\ -R^{-1}(\vec{c} + c\vec{v}),\ -R^{-1}\vec{v},\ R^{-1}\big)\ .$$

3.* Provare che il gruppo di Galileo è isomorfo ad un sottogruppo del gruppo $GL(5, \mathbb{R})$, delle matrici reali non singolari 5×5.

2.1.3 Moto relativo di riferimenti inerziali.

Quello che vogliamo provare ora è il seguente fondamentale risultato che riguarda il moto relativo tra sistemi di riferimento inerziali.

Teorema 2.1. *Assumiamo che, per ogni riferimento $\mathscr{I}$ ed ogni evento $p \in \mathbb{V}^4$, sia disponibile un punto materiale la cui linea di universo attraversi p con velocità (rispetto a $\mathscr{I}$) di valore e direzione arbitaria e che il punto sia isolato per qualche intervallo temporale finito attorno a p. Valgono allora i fatti seguenti.*
(a) *Se $\mathscr{I}$ è un riferimento inerziale, un altro riferimento $\mathscr{I}'$ è anch'esso inerziale se e solo se*

è in moto rettilineo uniforme rispetto a $\mathscr{I}$.
(b) *Le trasformazioni tra coordinate cartesiane ortonormali solidali con riferimenti inerziali sono le trasformazioni del gruppo di Galileo.*
(c) *Se $\mathscr{I}, \mathscr{I}'$ sono entrambi inerziali, $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$, l'accelerazione di trascinamento e di Coriolis di $\mathscr{I}'$ rispetto a $\mathscr{I}$ sono sempre nulle.* $\diamondsuit$

Dimostrazione. (a) supponiamo che $\mathscr{I}$ sia inerziale e che $\mathscr{I}'$ sia in moto rettilineo uniforme rispetto a $\mathscr{I}$. In tal caso, riferendosi a sistemi di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ e $\mathscr{I}'$ rispettivamente, avremo la legge di trasformazione

$$\begin{cases} t' = & t + c\,, \\ x'^i = & c^i + tv^i + \displaystyle\sum_{j=1}^{3} R^i{}_j x^j\,, \quad \text{i=1,2,3.} \end{cases} \tag{2.6}$$

Se γ è la linea di universo di un punto isolato, per $\mathscr{I}$ deve dare luogo ad una curva $P = P(t)$ con velocità costante $\mathbf{v}_P|_{\mathscr{I}} = \sum_{j=1}^{3} u^j\,\mathbf{e}_i$. In coordinate x^j, solidali con $\mathscr{I}$, la legge oraria varrà dunque forma $x^j(t) = tu^j + x_0^j$, per $j = 1, 2, 3$ e $t \in \mathbb{R}$, dove le x_0^j sono costanti. Per (2.6), la legge oraria di γ nelle coordinate x'^i solidali con $\mathscr{I}'$ sarà:

$$x'^i(t') = c^i + (t' - c)v^i + \sum_{j=1}^{3} R^i{}_j((t' - c)u^j + x_0^j)\,.$$

Il calcolo della velocità fornisce immediatamente il vettore costante $\mathbf{v}_P|_{\mathscr{I}'} = \sum_{i=1}^{3}(v^i + R^i{}_j)u^j\,\mathbf{e}'_i$. Quindi *ogni punto materiale isolato* si muove, per $\mathscr{I}'$, di moto rettilineo uniforme: $\mathscr{I}'$ è inerziale. Supponiamo viceversa che $\mathscr{I}$ e $\mathscr{I}'$ siano entrambi inerziali. Vogliamo studiare il moto, in entrambi i riferimenti, di un punto materiale isolato P, tenendo conto delle relazioni (1.70):

$$\begin{cases} \mathbf{a}_P|_{\mathscr{I}} = & \mathbf{a}_P|_{\mathscr{I}'} + \mathbf{a}_P^{(tr)} + \mathbf{a}_P^{(C)}\,, \\ \mathbf{a}_P^{(tr)} := & \mathbf{a}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')) + \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')\,, \\ \mathbf{a}_P^{(C)} := & 2\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'}\,. \end{cases}$$

O' è un punto arbitrario in quiete con $\mathscr{I}'$. Nel caso in esame, dato che P è isolato ed entrambi i riferimenti sono inerziali, deve essere $\mathbf{a}_P|_{\mathscr{I}} = \mathbf{a}_P|_{\mathscr{I}'} = \mathbf{0}$ ad ogni istante (nell'intervallo di tempo in cui il punto è isolato). Sostituendo nella prima formula riportata sopra si ha che, ad ogni istante e per ogni punto materiale isolato P, deve valere:

$$\mathbf{a}_P^{(tr)} + \mathbf{a}_P^{(C)} = \mathbf{0}\,.$$

ossia

$$\mathbf{a}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')) + \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + 2\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} = \mathbf{0}$$

Scegliendo il punto materiale isolato con velocità nulla in $\mathscr{I}'$ e coincidente con O', concludiamo che deve essere, all'istante considerato: $\mathbf{a}_{O'}|_{\mathscr{I}} = \mathbf{0}$. Dato che possiamo fare lo stesso ragionamento per ogni istante, il risultato trovato deve essere valido ad ogni istante. In definitiva deve essere:

$$\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')) + \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + 2\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} = \mathbf{0}\,, \tag{2.7}$$

ad ogni istante e per ogni punto materiale isolato P. Scegliendo nuovamente P coincidente con O' ed animato di velocità $\mathbf{v}_P|_{\mathscr{I}'} \neq \mathbf{0}$ perpendicolare al valore di $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ nell'istante considerato, si ottiene da (2.7):

$$||\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}||\,||\mathbf{v}_P|_{\mathscr{I}'}|| = 0$$

da cui $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} = \mathbf{0}$ nell'istante considerato e quindi in ogni istante. Questo risultato inserito nelle (1.69): $\mathbf{v}_P|_{\mathscr{I}} = \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + \mathbf{v}_{O'}|_{\mathscr{I}}$, prova che ogni punto P in quiete con $\mathscr{I}'$ ha velocità costante rispetto a $\mathscr{I}$, pari a $\mathbf{v}_{O'}|_{\mathscr{I}}$ indipendentemente da P. D'altra parte, avendo anche trovato sopra che $\frac{d\mathbf{v}_{O'}|_{\mathscr{I}}}{dt}|_{\mathscr{I}}(= \mathbf{a}_{O'}|_{\mathscr{I}}) = \mathbf{0}$ per ogni istante, concludiamo che la velocità dei punti in quiete con $\mathscr{I}'$ è *costante nel tempo e non dipende dal punto*. Per definizione, $\mathscr{I}'$ è dunque in moto rettilineo uniforme rispetto a $\mathscr{I}$.
(b) la prova è conseguenza immediata di (a) di questo teorema e dei primi due punti della proposizione 2.1.
(c) Segue immediatamente da (b). □

Esercizi 2.2.

1. Si consideri la *grande ruota* di un *luna park*. Si assuma approssimativamente che il riferimento solidale con la terra, $\mathscr{I}$, sia un riferimento inerziale e si consideri un secondo riferimento $\hat{\mathscr{I}}$ solidale con un sedile (libero di rimanere nella posizione orizzontale quando la ruota gira) della grande ruota (in movimento). Quanto vale il vettore $\boldsymbol{\omega}_{\hat{\mathscr{I}}}|_{\mathscr{I}}$? Il riferimento $\hat{\mathscr{I}}$ è inerziale?

2.2 Formulazione generale della dinamica classica dei sistemi di punti materiali.

Veniamo a spiegare come procede la costruzione della dinamica una volta assunta l'esistenza della classe di riferimenti inerziali $\mathfrak{I}$ con le caratteristiche dette. Considerando punti materiali *isolati*, cioé sufficientemente lontani gli uni dagli altri, sappiamo che il loro moto, descritto dal primo principio della dinamica, è rettilineo uniforme in ogni riferimento inerziale. Il problema è ora di spiegare cosa accada al moto dei punti quando essi *non* siano più isolati. L'esperienza mostra che il loro moto diventa accelerato.

2.2.1 Masse, Impulsi e Forze.

Se abbiamo un sistema di punti materiali, diremo che esso è **isolato** quando i punti di esso sono lontani dagli altri corpi dell'universo non appartenenti all'insieme (i punti dell'insieme possono

essere invece arbitrariamente vicini gli uni agli altri).
Per procedere con la formulazione della dinamica, si assume che i punti materiali abbiano associata a ciascuno di essi una grandezza fisica chiamata **massa** che sia additiva e si conservi. Attraverso la massa si può enunciare la legge di conservazione dell'*impulso* o *quantità di moto*. Abbiamo il seguente principio che formalizza quanto detto.

C2. Impulso e massa.
Si assume che per ogni punto materiale Q esista una costante detta **massa**, $m > 0$, *tale che valgano i due seguenti fatti.*
(a) Principio di conservazione impulso. *In ogni riferimento inerziale $\mathscr{I}$ e per ogni coppia di punti materiali Q, Q', costituenti un sistema isolato*

$$m\mathbf{v}_Q|_{\mathscr{I}} + m'\mathbf{v}_{Q'}|_{\mathscr{I}} \ ,$$

è un vettore costante nel tempo, anche se non lo sono separatamente $\mathbf{v}_Q|_{\mathscr{I}}$ e $\mathbf{v}_{Q'}|_{\mathscr{I}}$. Il vettore $\mathbf{p} := m\mathbf{v}_Q|_{\mathscr{I}}$ è detto **impulso** *o anche* **quantità di moto** *del punto materiale Q rispetto a $\mathscr{I}$.*
(b) Principio di conservazione/additività della massa. *Nei processi in cui un punto materiale si decompone in più punti materiali (oppure più punti materiali si fondono in un unico punto materiale), la massa dell'unico punto materiale iniziale (finale) è pari alla somma delle masse dei costituenti finali (iniziali).*

Osservazioni 2.2.
(1) Le masse sono indipendenti dal riferimento, mentre gli impulsi non lo sono e si trasformano, cambiando riferimento, tenendo conto della legge (1.69):

$$\mathbf{p}_{\mathscr{I}} = \mathbf{p}_{\mathscr{I}'} + m\mathbf{v}_Q^{(tr)} \ . \tag{2.8}$$

Tuttavia, se entrambi $\mathscr{I}$ e $\mathscr{I}'$ sono inerziali, a causa del teorema 2.1, che sancisce in particolare che la velocità di trascinamento è costante nel tempo ed indipendente dal punto, e coincide con $\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}}$, la formula di sopra si semplifica in

$$\mathbf{p}_{\mathscr{I}} = \mathbf{p}_{\mathscr{I}'} + m\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}} \ . \tag{2.9}$$

(2) Possiamo assumere che la massa un punto materiale di riferimento P_1 valga 1 ed, attraverso il principio di conservazione dell'impulso, misurare le masse di ogni altro punto materiale P_2 facendolo interagire con P_1, assumendo tutti gli altri corpi lontani. m_{P_2} è l'unico numero che soddisfa, per $t \neq t'$ istanti arbitrari: $m_{P_2}(\mathbf{v}_{P_2}|_{\mathscr{I}}(t') - \mathbf{v}_{P_2}|_{\mathscr{I}}(t)) = 1(\mathbf{v}_{P_1}|_{\mathscr{I}}(t') - \mathbf{v}_{P_2}|_{\mathscr{I}}(t))$, quando i punti non hanno velocità costante nel tempo (a causa della loro interazione).

Passiamo quindi ad introdurre la nozione di *forza* per gestire le interazioni tra punti materiali vicini.

C3. Secondo principio della dinamica (prima parte).
Si assume che presi due punti materiali Q e Q' di masse m e m' rispettivamente, costituenti

un sistema isolato e per ogni riferimento inerziale $\mathscr{I}$, esistano due funzioni differenziabili dette **funzioni di forza** *o semplicemente* **forze** *agenti sui punti materiali,* $\mathbf{F}_{\mathscr{I}} : \Sigma_t \times \Sigma_t \times V_t \times V_t \to V_t$ *e* $\mathbf{F}'_{\mathscr{I}} : \Sigma_t \times \Sigma_t \times V_t \times V_t \to V_t$ *tali che, per ogni* $t \in \mathbb{R}$*, valga il* **secondo principio della dinamica***,*

$$\begin{cases} m\mathbf{a}_Q|_{\mathscr{I}}(t) = & \mathbf{F}_{\mathscr{I}}\left(Q(t), Q'(t), \mathbf{v}_Q|_{\mathscr{I}}(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t)\right) \ , \\ m'\mathbf{a}_{Q'}|_{\mathscr{I}}(t) = & \mathbf{F}'_{\mathscr{I}}\left(Q'(t), Q(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t), \mathbf{v}_Q|_{\mathscr{I}}(t)\right) \ . \end{cases} \tag{2.10}$$

Più precisamente $\mathbf{F}_{\mathscr{I}}$ *è detta la forza che* **agisce su** Q **a causa di** Q' *e* $F'_{\mathscr{I}}$ *è detta forza che* **agisce su** Q' **a causa di** Q*.* $\mathbf{F}_{\mathscr{I}}$ *è detta* **reazione** *associata a* $\mathbf{F}'_{\mathscr{I}}$ *e* $\mathbf{F}'_{\mathscr{I}}$ *è detta* **reazione** *associata a* $\mathbf{F}_{\mathscr{I}}$*.*
Al variare del riferimento inerziale, si assume che le funzioni di forza siano **invarianti***, cioé siano tali da soddisfare la relazione:*

$$\mathbf{F}_{\mathscr{I}}\left(P_1, P_2, \mathbf{u}_1, \mathbf{u}_2\right) = \mathbf{F}_{\mathscr{I}'}\left(P_1, P_2, \mathbf{u}'_1, \mathbf{u}'_2\right) \ , \tag{2.11}$$

per $P_1, P_2 \in \Sigma_t$ *e* $\mathbf{u}_1, \mathbf{u}_2 \in V_t$ *e dove abbiamo definito* $\mathbf{u}' := \mathbf{u} + \mathbf{u}_{\mathscr{I}}|_{\mathscr{I}'}$ *per ogni* $\mathbf{u} \in V_t$*.*

Come diremo più in generale tra poco, riferendosi a coordinate x^1, x^2, x^3 cartesiane ortonormali solidali con $\mathscr{I}$, le equazioni di sopra diventano un *sistema di equazioni differenziali del secondo ordine* che *determina completamente* le leggi orarie dei punti, una volta note posizioni e velocità di essi ad un istante t_0 arbitrario.
Dal punto di vista fisico, una volta formulata il secondo principio, bisogna determinare le varie leggi di forza che esistono in natura e che dipendono da ulteriori caratteristiche dei punti materiali (per esempio la carica elettrica che è associata alla legge di forza di Coulomb).

Osservazioni 2.3.
(1) Per non violare la prima legge della dinamica, ogni forza fisicamente sensata deve decrescere ed annullarsi quando la distanza tra i punti cresce, in modo da fornire accelerazioni nulle per corpi sufficientemente lontani.
(2) Il requisito di invarianza (2.11) segue dalla richiesta di validità delle (2.10) in ogni riferimento inerziale. In fatti, sappiamo che se $\hat{\mathscr{I}}$ è un secondo riferimento inerziale, esso sarà legato al precedente da una trasformazione di Galileo e pertanto $\boldsymbol{\omega}_{\hat{\mathscr{I}}}|_{\mathscr{I}} = \boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = 0$ e l'accelerazione di trascinamento di ogni punto in quiete in uno dei due riferimenti è nulla se riferita all'altro. Usando (1.70) troviamo che *l'accelerazione di punto P ad un istante di tempo fissato non varia se si cambia riferimento, passando da un riferimento inerziale $\mathscr{I}$ ad un altro riferimento inerziale* $\hat{\mathscr{I}}$: $\mathbf{a}_Q|_{\hat{\mathscr{I}}} = \mathbf{a}_Q|_{\mathscr{I}}$ e $\mathbf{a}_{Q'}|_{\hat{\mathscr{I}}} = \mathbf{a}_{Q'}|_{\mathscr{I}}$. Di conseguenza le (2.10) si possono riscrivere, dove i primi membri sono riferiti a $\hat{\mathscr{I}}$, ma i secondi a $\mathscr{I}$:

$$\begin{cases} m\mathbf{a}_Q|_{\hat{\mathscr{I}}}(t) = & \mathbf{F}_{\mathscr{I}}\left(Q(t), Q'(t), \mathbf{v}_Q|_{\mathscr{I}}(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t)\right) \ , \\ m'\mathbf{a}_{Q'}|_{\hat{\mathscr{I}}}(t) = & \mathbf{F}'_{\mathscr{I}}\left(Q'(t), Q(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t), \mathbf{v}_Q|_{\mathscr{I}}(t)\right) \ . \end{cases}$$

Ma, dato che il secondo principio vale anche in $\hat{\mathscr{I}}$:

$$\begin{cases} m\mathbf{a}_Q|_{\hat{\mathscr{I}}}(t) = & \mathbf{F}_{\hat{\mathscr{I}}}\left(Q(t), Q'(t), \mathbf{v}_Q|_{\hat{\mathscr{I}}}(t), \mathbf{v}_{Q'}|_{\hat{\mathscr{I}}}(t)\right) \ , \\ m'\mathbf{a}_{Q'}|_{\hat{\mathscr{I}}}(t) = & \mathbf{F}'_{\hat{\mathscr{I}}}\left(Q'(t), Q(t), \mathbf{v}_{Q'}|_{\hat{\mathscr{I}}}(t), \mathbf{v}_Q|_{\hat{\mathscr{I}}}(t)\right) \ . \end{cases}$$

Visto che m, m' non dipendono dal riferimento, concludiamo che deve valere (2.11), se si assume la possibilità di dare ai punti materiali e in un istante dato, qualsiasi posizione e velocità .

In virtù della validità della (2.11), d'ora in avanti ometteremo l'indice $_{\mathscr{I}}$ che specifica il riferimento inerziale nel quale sono assegnate le funzioni di forza. Nello stesso modo ometteremo la specificazione $|_{\mathscr{I}}$ nelle accelerazioni quando è sottointeso che sono valutate rispetto ad un riferimento inerziale (non importa quale per quanto spiegato sopra!).

La conservazione dell'impulso unitamente al secondo principio della dinamica, implica immediatamente quello che, nella presentazione tradizionale, viene chiamato **terzo principio della dinamica**[2] o **principio di azione e reazione** che mette in relazione la forza agente su Q con quella agente su Q':

$$\mathbf{F}\left(Q(t), Q'(t), \mathbf{v}_Q|_{\mathscr{I}}(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t)\right) = -\mathbf{F}'\left(Q'(t), Q(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t), \mathbf{v}_Q|_{\mathscr{I}}(t)\right) \ .$$

Il principio di azione e reazione **in forma forte** richiede che le funzioni di forza siano applicate lungo la congiungente $Q(t) - Q'(t)$; tale richiesta servirà ad ottenere, come teorema, la conservazione del momento angolare per sistemi di punti:

C3. Secondo principio della dinamica (seconda parte). *Per ogni $t \in \mathbb{R}$, se $Q(t) - Q'(t)$ non è il vettore nullo,*

$$\mathbf{F}\left(Q(t), Q'(t), \mathbf{v}_Q|_{\mathscr{I}}(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t)\right)$$

e

$$\mathbf{F}'\left(Q'(t), Q(t), \mathbf{v}_{Q'}|_{\mathscr{I}}(t), \mathbf{v}_Q|_{\mathscr{I}}(t)\right)$$

sono diretti lungo la direzione di tale vettore (con versi opposti).

2.2.2 Sovrapposizione delle forze.

Se un punto materiale, Q, interagisce contemporaneamente con più punti materiali si assume che la funzione forza totale applicata a Q sia data dalla somma delle funzioni forza che si hanno considerando tutte le coppie di punti possibili con Q fisso.

C4. Principio di sovrapposizione delle forze.
Si assume che se è dato un sistema di N punti materiali $Q_1, \ldots, Q_N$, valga per ogni punto Q_k, in ogni riferimento inerziale $\mathscr{I}$, il **secondo principio della dinamica in forma completa**

$$m_k \mathbf{a}_{Q_k}(t) = \sum_{k \neq i = 1}^{i=N} \mathbf{F}_{ki}\left(Q_k(t), Q_i(t), \mathbf{v}_{Q_k}|_{\mathscr{I}}(t), \mathbf{v}_{Q_i}|_{\mathscr{I}}(t)\right) \ ,$$

[2] Per noi è un semplice corollario dei nostri principi.

dove la funzione $\mathbf{F}_{ki}$ è la forza agente su Q_k e dovuta a Q_i, ottenuta allontanando sufficientemente tutti gli altri punti del sistema.

Se $\mathbf{x}_k$ denota la terna di coordinate determinanti Q_k in riferimento ad un sistema di coordinate cartesiane ortonormali solidali con il riferimento $\mathscr{I}$, l'insieme delle equazioni

$$\frac{d^2\mathbf{x}_k(t)}{dt^2} = m_k^{-1}\mathbf{F}_i\left(\mathbf{x}_1(t),\ldots,\mathbf{x}_k(t),\frac{d\mathbf{x}_1(t)}{dt},\ldots,\frac{d\mathbf{x}_N(t)}{dt}\right)\ , \quad k=1,2,\ldots,N\ , \tag{2.12}$$

dove, per $i=1,\ldots,N$,

$$\mathbf{F}_i\left(\mathbf{x}_1(t),\ldots,\mathbf{x}_k(t),\frac{d\mathbf{x}_1(t)}{dt},\ldots,\frac{d\mathbf{x}_N(t)}{dt}\right) := \sum_{k\neq i=1}^{i=N}\mathbf{F}_{ki}\left(\mathbf{x}_k(t),\mathbf{x}_i(t),\frac{d\mathbf{x}_k(t)}{dt},\frac{d\mathbf{x}_i(t)}{dt}\right)\ , \tag{2.13}$$

costituisce ancora un sistema di equazioni differenziali in *forma normale*, per cui si applicano le stesse considerazioni sull'esistenza e l'unicità delle soluzioni già esposte nel caso di due corpi.

2.2.3 Problema fondamentale della dinamica e determinismo.

Il problema di determinare il **moto del sistema**, cioé la *legge oraria*, dei punti materiali di un sistema, quando sono assegnate le leggi di forza agenti su ciascun punto e sono date *condizioni iniziali*, cioé posizione e velocità di ogni punto del sistema ad un certo istante rispetto ad un fissato riferimento inerziale, si dice **problema fondamentale della dinamica**. In definitiva, con alcune generalizzazioni molto importanti che diremo tra poco, il problema si riduce a risolvere, in coordinate cartesiane ortogonali solidali con un riferimento inerziale $\mathscr{I}$, il sistema di equazioni differenziali (2.12) quando sono assegnati i dati iniziali $\mathbf{x}_k(t_0)$ e $\frac{d\mathbf{x}_k(t)}{dt}|_{t=t_0}$ per $k=1,2,\ldots,N$. Tutta la teoria è basata sul seguente teorema che discuteremo, indebolendone le ipotesi e migliorandone la tesi, e proveremo nel prossimo capitolo:

Teorema 2.2. *Se $\mathbf{F}: I\times D\times D'\to\mathbb{R}^n$ con $D,D'\subset\mathbb{R}^n$ aperti non vuoti e $I\subset\mathbb{R}$ intervallo aperto, è una funzione con $\mathbf{F}\in C^1(I\times D\times D';\mathbb{R}^n)$, si consideri l'equazione differenziale*

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F}\left(t,\mathbf{x},\frac{d\mathbf{x}}{dt}\right)\ . \tag{2.14}$$

Per ogni $(t_0,\mathbf{x}_0,\dot{\mathbf{x}}_0)\in\mathbb{R}\times D\times D'$ esiste un'unica funzione di classe C^2, $\mathbf{x}=\mathbf{x}(t)$, definita in un intorno di t_0 che soddisfi le condizioni iniziali $\mathbf{x}(t_0)=\mathbf{x}_0$ e $\frac{d\mathbf{x}}{dt}|_{t=t_0}=\dot{\mathbf{x}}_0$ e che risolva (2.14). $\diamondsuit$

Osservazioni 2.4.
(1) Assumendo opportune ipotesi di regolarità sulle funzioni di forza, almeno localmente, il secondo principio della dinamica determina il moto, *nel futuro, ma anche nel passato*, se sono assegnate condizioni iniziali. Questo risultato matematico prende il nome di **determinismo** ed ha enorme importanza in fisica classica.

(2) Dal punto di vista euristico la validità del teorema 2.2 è plausibile se si rafforzano ulteriormente le ipotesi, assumendo $\mathbf{F}$ di classe C^∞. Infatti, inserendo le condizioni iniziali $\mathbf{x}(t_0) = \mathbf{x}_0$ e $\frac{d\mathbf{x}}{dt}|_{t=t_0} = \dot{\mathbf{x}}_0$ nel secondo membro di (2.14), il primo membro fornisce la derivata seconda della soluzione incognita $\mathbf{x}(t)$ in t_0. Tenendo conto di tale informazione e reiterando la procedura, si ottiene la derivata terza di $\mathbf{x}(t)$ in t_0. Iternado all'infinito la procedura si ottengono tutte le derivate di $\mathbf{x}(t)$ in t_0, $\frac{d^n\mathbf{x}}{dt^n}|_{t_0}$. A questo punto, ci si aspetta che, se la serie:

$$\sum_{n=0}^{+\infty} \left.\frac{d^n\mathbf{x}}{dt^n}\right|_{t_0} \frac{(t-t_0)^n}{n!}$$

converge in un intorno di t_0, la funzione di t ottenuta in tal modo sia una soluzione di (2.14) con le assegnate condizioni iniziali. È chiaro che, nella classe delle soluzioni analitiche, se non vuota, questa è l'unica soluzione locale con le assegnate condizioni iniziali.
Si osservi ancora che se, per l'equazione differenziale considerata, non fosse stato possibile separare a primo membro la derivata seconda lasciando a secondo membro una funzione delle derivate di ordine più basso, la procedura euristica presentata sopra non avrebbe avuto garanzia di funzionare. Dimostreremo nel prossimo capitolo il teorema 2.2 usando un'altra procedura che richiede meno ipotesi.

Esempi 2.1.
1. Tre punti materiali P_1, P_2, P_3 sono connessi, a due a due, da molle di lunghezza a riposo nulla e costante elastica $k > 0$. Vogliamo scrivere le leggi di forza per ciascuno di essi ed il sistema di equazioni differenziali che ne determina il moto.
La legge di forza di una molla di estremi P e Q e di lunghezza nulla a riposo è semplicemente

$$\mathbf{F}_P(P,Q) = -\kappa(P-Q)\,, \tag{2.15}$$

dove f_P è la forza esercitata dalla molla in P. Avremo dunque che, per $i \neq k$:

$$\mathbf{F}_{ki}(P_k, P_i) = -\kappa(P_k - P_i)\,.$$

Notare che il principio di azione e reazione è automaticamente soddisfatto. Il sistema di equazioni che determina il moto per i tre punti è dunque:

$$m_k \mathbf{a}_{P_k}(t) = \sum_{k\neq i=1}^{i=3} -\kappa(P_k - P_i) \qquad k = 1,2,3\,. \tag{2.16}$$

2. Consideriamo ancora tre punti P_1, P_2, P_3 sottoposti alla forza gravitazionale. Vogliamo scrivere il sistema di equazioni differenziali che ne determina il moto.
La forza gravitazionale che si esercita tra due punti materiali P, Q di masse m_P, m_Q rispettivamente, ha la struttura

$$\mathbf{F}_P(P,Q) = -G\frac{m_P m_Q}{||P-Q||^3}(P-Q)\,, \tag{2.17}$$

dove $G > 0$ è la costante gravitazionale.
Il sistema di equazioni che determina il moto per i tre punti è dunque:

$$m_k \mathbf{a}_{P_k}(t) = \sum_{k \neq i=1}^{i=3} -G \frac{m_{P_k} m_{P_i}}{||P_k - P_i||^3}(P_k - P_i) \quad k = 1, 2, 3 \,. \tag{2.18}$$

Si ha una formula analoga nel caso di forza coulombiana, ma in tal caso le masse nella legge di forza devono essere sostituite dalle cariche elettriche dei punti materiali ed il segno della costante G deve essere negativo.

2.3 Situazioni dinamiche più generali.

Non sempre lo schema newtoniano esposto nelle precedenti sezioni può essere implementato completamente nella forma vista. In particolare il problema fondamentale della dinamica deve, in alcuni casi essere impostato in modo differente. Vediamo tre casi di grande rilevanza.

2.3.1 Moto assegnato per un sottosistema: forze dipendenti dal tempo.

Il primo caso interessante e quando, in un sistema di N punti materiali interagenti tra di essi, alcuni di loro, $P_{k+1}, \ldots, P_N$, hanno moto in qualche modo assegnato[3]. I rimanenti punti materiali, $P_1, P_2, \ldots, P_k$ avranno funzioni di forza della forma, per $i = 1, \ldots, k$,

$$\mathbf{F}_i = \mathbf{F}_i(t, P_1, \ldots, P_k, \mathbf{v}_{P_1}, \ldots, \mathbf{v}_{P_k}) \,.$$

La dipendenza temporale esplicita tiene conto del moto, assegnato, dei rimanenti punti $P_{k+1}, \ldots, P_N$.
Il secondo principio della dinamica, per il sistema dei punti P_i, con $i = 1, \ldots, k$, può essere esteso alla forma più generale

$$m_i \mathbf{a}_{P_i} = \mathbf{F}_i(t, P_1, \ldots, P_k, \mathbf{v}_{P_1}, \ldots, \mathbf{v}_{P_k}) \,, \quad \text{con } i = 1, \ldots, k. \tag{2.19}$$

Dal punto di vista matematico, passando in coordinate cartesiane ortonormali solidali con un riferimento inerziale, il sistema di equazioni differenziali che nasce dal sistema (2.19) rientra nella stessa teoria del caso già visto del teorema 2.2, e, se le funzioni di forza sono sufficientemente regolari, il sistema

$$\frac{d^2 \mathbf{x}_i(t)}{dt^2} = m_i^{-1} \mathbf{F}_i \left(t, \mathbf{x}_1(t), \ldots, \mathbf{x}_k(t), \frac{d\mathbf{x}_1(t)}{dt}, \ldots, \frac{d\mathbf{x}_k(t)}{dt} \right) \,, \quad i = 1, \ldots, k \,, \tag{2.20}$$

ammette una ad una sola soluzione quando sono assegnate le posizioni e le velocità dei punti ad un istante. Si osservi che abbiamo un sistema di $3k$ equazioni scalari e le funzioni incognite sono

[3] Questo è un caso è tipico quando si hanno alcuni punti materiali costituenti un sistema isolato e uno di essi ha una massa enormemente più grande di quella dei rimanenti. Con buona approssimazione, tanto migliore quanto la sua massa è superiore a quella degli altri, tale corpo si muove di moto rettilineo uniforme in un (qualunque) riferimento inerziale.

effettivamente $3k$.
In questo approccio si possono fare ricadere situazioni in cui un punto materiale interagisce con un sistema più complesso dato da un fluido. In tal caso l'azione dinamica dell'ambiente esterno sul punto è ancora rappresentata da una funzione forza dipendente esplicitamente dal tempo.

Esempi 2.2.
(1) Il primo esempio, del tutto banale, è quello di un punto materiale P di massa m sottoposto alla forza di una molla di lunghezza nulla a riposo e di costante elastica κ, di cui uno dei due estremi, O, è vincolato a rimanere immobile in un riferimento inerziale $\mathscr{I}$. (Dal punto di vista pratico ciò si ottiene attaccando tale estremo ad un corpo di massa $M \gg m$.) In questo caso l'equazione che determina il moto di P è semplicemente

$$m\mathbf{a}_P = -\kappa(P - O)\,. \tag{2.21}$$

(2) Considerando un punto materiale P di massa m immerso in un liquido in quiete nel riferimento $\mathscr{I}'$, la forza agente sul punto materiale a causa dell'**attrito viscoso** del liquido è espressa, in certi regimi di velocità , dalla forma funzionale valida nel riferimento inerziale $\mathscr{I}$

$$\mathbf{F}(t, P, \mathbf{v}_P|_{\mathscr{I}}) = -\gamma \mathbf{v}_P|_{\mathscr{I}'}\,. \tag{2.22}$$

dove $\gamma > 0$ è un coefficiente che descrive l'attrito tra punto materiale e liquido. Se $\mathscr{I}' \neq \mathscr{I}$ e $\mathscr{I}'$ è inerziale, possiamo lavorare direttamente nel riferimento $\mathscr{I}'$ e la forma della forza di sopra si semplifica ottenendo alla fine le equazioni del moto (assumendo che sul punto non agiscano altre forze)

$$m\mathbf{a}_P = -\gamma \mathbf{v}_P|_{\mathscr{I}'}\,. \tag{2.23}$$

Questa equazione può essere risolta come equazione differenziale con procedure standard, una volta trascritta in coordinate cartesiane solidali con $\mathscr{I}'$.
3. In riferimento all'esempio precedente, assumiamo ora che $\mathscr{I}'$ *non* sia inerziale, ma il suo moto sia assegnato rispetto al riferimento inerziale $\mathscr{I}$. In questo caso non possiamo ridurci a lavorare in $\mathscr{I}'$, perché in esso gli assiomi della dinamica non sono verificati (discuteremo in seguito cosa accade volendo riformulare la meccanica in riferimenti non inerziali). Rimaniamo in $\mathscr{I}$ ed esprimiamo $\mathbf{v}_P|_{\mathscr{I}'}$ in funzione di $\mathbf{v}_P|_{\mathscr{I}}$. Dalle equazioni (1.69) abbiamo che $\mathbf{v}_P|_{\mathscr{I}'} = \mathbf{v}_P|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'} \wedge (P - O) + \mathbf{v}_O|_{\mathscr{I}'}$ dove $O \in E_{\mathscr{I}}$ è un punto in quiete con $\mathscr{I}$ (tipicamente l'origine delle coordinate cartesiane ortonormali solidali con $\mathscr{I}$). Le *funzioni del tempo* $\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'} = \boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'}(t)$ e $\mathbf{v}_O|_{\mathscr{I}'} = \mathbf{v}_O|_{\mathscr{I}'}(t)$ sono note perché si dal fatto che il moto relativo di $\mathscr{I}$ e $\mathscr{I}'$ è assegnato. In definitiva, le equazioni del moto per P (assumendo che sul punto non agiscano altre forze)

$$m\mathbf{a}_P = -\gamma \mathbf{v}_P|_{\mathscr{I}} - \gamma \boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'} \wedge (P - O) - \gamma \mathbf{v}_O|_{\mathscr{I}'}\,. \tag{2.24}$$

2.3.2 Vincoli geometrici: reazioni vincolari.

Un altro caso, di grande interesse matematico, della situazione in cui non è possibile usare completamente lo schema della dinamica esposto sopra, è quello in cui punti materiali $P_1, P_2, \ldots, P_k$

devono anche soddisfare *equazioni di vincolo* di tipo geometrico/cinematico. Per esempio, si può richiedere che uno o più punti sottoposto a varie funzioni di forza debba avere comunque moto confinato su una curva o su una superficie, oppure che siano soddisfatte richieste sulla velocità del punto (o dei punti) oppure, infine, che siano soddisfatte delle richieste di carattere metrico, per esempio il moto di due o più punti, sottoposti a forze assegnate, deve comunque essere tale che la distanza tra i punti sia fissata e costante nel tempo. In tutte queste situazioni si assume ancora che, a causa del vincolo, vi sia una forza (in aggiunta a quella già presenti) agente sul punto materiale. Tale forza, detta **reazione vincolare**, è indicata con la lettera ϕ. Si suppone che tale forza soddisfi, per ipotesi, il principio di azione e reazione (terzo principio) rispetto alla struttura del vincolo e che contribuisca all'equazione generalizzata per il secondo principio[4]:

$$m_i \mathbf{a}_{P_i} = \mathbf{F}_i(t, P_1, \ldots, P_k, \mathbf{v}_{P_1}, \ldots, \mathbf{v}_{P_k}) + \boldsymbol{\phi}_i \ , \quad \text{con } i = 1, \ldots, k. \tag{2.25}$$

Non è però assegnata la forza $\boldsymbol{\phi}$ come funzione della posizione e velocità del punto sottoposto al vincolo, *anche se in linea di principio si assume che ciò sarebbe possibile se si esaminasse in dettaglio la struttura fisica del vincolo.* In pratica, però , la forza $\boldsymbol{\phi}$ risulta essere un *incognita* del problema del moto come lo è la legge oraria del moto stesso. Affinché il problema complessivo sia risolvibile, bisogna accompagnare l'equazione differenziale (più in generale il sistema di equazioni differenziali), derivante dall'applicazione del secondo principio, con alcune informazioni supplettive. In primo luogo l'equazione del vincolo stesso. Non è detto che ciò sia sufficiente e deve essere spesso aggiunta una **caratterizzazione costitutiva del vincolo** che fissa qualche relazione tra le componenti della reazione vincolare. L'uso delle relazioni aggiuntive serve, quando possibile, a determinare un sistema di equazioni differenziali, *che non contenga le reazioni vincolari incognite* $\boldsymbol{\phi}$, per il quale siano noti teoremi di esistenza ed unicità se fissati condizioni iniziali. Questo sistema, quando esiste, viene detto **sistema di equazioni pure di movimento**. Le reazioni vincolari $\boldsymbol{\phi}$ vengono successivamente determinate, quando possibile, una volta nota l'equazione oraria e facendo uso del secondo e del terzo principio della dinamica. Vediamo alcuni esempi, nel caso di un *singolo* punto materiale, per chiarire il significato di quanto detto. Negli esercizi tratteremo il caso di più punti materiali. Il caso di uno o più punti materiali può essere trattato in modo del tutto generale introducendo la *formulazione Lagrangiana* della meccanica ed il postulato dei vincoli ideali che non tratteremo a questo livello.

Esempi 2.3.
1. Consideriamo un punto vincolato ad una curva regolare Γ, diversa da un segmento di retta, solidale con un riferimento inerziale $\mathscr{I}$ e di equazione $P = P(u)$. Se $\mathbf{F} = \mathbf{F}(t, P, \mathbf{v}_P|_{\mathscr{I}})$ denota la legge di forza complessiva agente su P di espressione assegnata, escludendo la reazione vincolare $\boldsymbol{\phi}$, il secondo principio della dinamica fornisce:

$$m\mathbf{a}_P = \mathbf{F}(t, P, \mathbf{v}_P) + \boldsymbol{\phi}(t) \ ,$$

che non è in grado di determinare il moto anche assegnando l'equazione della curva Γ come

[4]In taluni testi si distingue tra *forze attive*, che sono quelle di cui è nota la forma funzionale da noi indicate con $\mathbf{F}$, e le *forze reattive*, $\boldsymbol{\phi}$, di cui non è data la forma funzionale e che sono dovute ai vincoli.

$P = P(u)$. È conveniente, dato che il moto avviene su Γ, descrivere P tramite l'ascissa curvilinea s parametrizzata, a sua volta, nel tempo: $s = s(t)$. Usando (1.20), in riferimento alla *terna intrinseca* della curva Γ introdotta nella sezione 1.3.3, il secondo principio può essere riscritto

$$m\frac{d^2s}{dt^2}\,\mathbf{t}(s(t)) + \frac{m}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2\mathbf{n}(s(t)) = F^t(t, P(s(t)), \mathbf{v}_P)\mathbf{t}(s(t)) + F^n(t, P(s(t)), \mathbf{v}_P)\mathbf{n}(s(t))$$

$$+F^b(t, P(s(t)), \mathbf{v}_P)\mathbf{b}(s(t)) + \phi^t(t)\mathbf{t}(s(t)) + \phi^n(t)\mathbf{n}(s(t)) + \phi^b(t)\mathbf{b}(s(t))\,.$$

Anche la velocità che compare come argomento conviene che sia riscritta in termini dell'ascissa curvilinea e del versore tangente come in (1.19): $\mathbf{v}(t) = \frac{ds}{dt}\,\mathbf{t}(s(t))$. In altri termini, il secondo principio fornisce il sistema di tre equazioni

$$m\frac{d^2s}{dt^2} = F^t\left(t, s(t), \frac{ds}{dt}\right) + \phi^t(t)\,, \tag{2.26}$$

$$\frac{m}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2 = F^n\left(t, s(t), \frac{ds}{dt}\right) + \phi^n(t)\,, \tag{2.27}$$

$$0 = F^b\left(t, s(t), \frac{ds}{dt}\right) + \phi^b(t)\,. \tag{2.28}$$

La più semplice caratterizzazione costitutive del vincolo dato dalla curva Γ è quella di **curva liscia**. Questa caratterizzazione costitutiva si esplica nella richiesta che la componente di $\boldsymbol{\phi}$ tangente a Γ sia sempre nulla. In tal caso le equazioni di sopra si semplificano in

$$m\frac{d^2s}{dt^2} = F^t\left(t, s(t), \frac{ds}{dt}\right)\,, \tag{2.29}$$

$$\frac{m}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2 = F^n\left(t, s(t), \frac{ds}{dt}\right) + \phi^n(t)\,, \tag{2.30}$$

$$0 = F^b\left(t, s(t), \frac{ds}{dt}\right) + \phi^b(t)\,. \tag{2.31}$$

L'equazione (2.29) non contiene le funzioni incognite componenti della reazione vincolare inoltre essa ricade nella classe di equazioni differenziali trattate nel teorema 2.2. Pertanto, assegnando posizione e velocità iniziale (tangente a Γ!), queste vengono tradotte in condizioni iniziali per la funzione $s = s(t)$ che si determina univocamente dall'equazione (2.29). L'equazione (2.29) è pertanto un'**equazione pura di movimento** per il sistema considerato.
Una volta nota la legge oraria $s = s(t)$, sostituendo tale funzione a primo membro di (2.30) e (2.31), si determinano le funzioni incognite ϕ^n e ϕ^b.
2. Consideriamo nuovamente un punto vincolato ad una curva regolare Γ, diversa da un segmento di retta, solidale con un riferimento inerziale $\mathscr{I}$ e di equazione $P = P(u)$, sottoposto alla forza $\mathbf{F} = \mathbf{F}(t, P, \mathbf{v}_P|_{\mathscr{I}})$ assegnata, in modo di arrivare nuovamente alle equazioni (2.28). Una differente e più fisica relazione costitutiva del vincolo dato dalla curva Γ è quella di **curva scabra**. Questa caratterizzazione costitutiva si esplica nella richiesta che la componente tangente ϕ^t della

reazione vincolare $\boldsymbol{\phi}$ agente su P a causa di Γ sia connessa alla componente normale $\phi^n\mathbf{n}+\phi^b\mathbf{b}$ dalla relazione

$$|\phi^t(t)| \leq \mu_s\sqrt{(\phi^n(t))^2+(\phi^b(t))^2}\,, \tag{2.32}$$

fino a quando il punto P è fermo rispetto a Γ (cioè è in quiete in $\mathscr{I}$) e dove $\mu_s > 0$ è un coefficiente detto **coefficiente d'attrito statico**. Quando il punto è in movimento vale invece

$$\phi^t(s(t)) = -\frac{\frac{ds}{dt}}{\left|\frac{ds}{dt}\right|}\,\mu_d\,\sqrt{(\phi^n(t))^2+(\phi^b(t))^2}\,, \tag{2.33}$$

dove $\mu_d > 0$ è un coefficiente detto **coefficiente d'attrito dinamico** con, solitamente, $\mu_d < \mu_s$. Il coefficiente $\frac{ds}{dt}/\left|\frac{ds}{dt}\right|$ moltiplicato per $\mathbf{t}$ individua il versore della velocità del punto per cui, nella fase di movimento, $\phi^t\mathbf{t}$ è sempre diretto lungo la velocità ma con verso opposto. Nel caso in esame, assumendo inizialmente il punto in movimento, l'equazione pura di movimento si ottiene come segue. Da (2.27) e (2.28) ricaviamo

$$\frac{m}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2 - F^n\left(t,s(t),\frac{ds}{dt}\right) = \phi^n(t)\,, \tag{2.34}$$

$$-F^b\left(t,s(t),\frac{ds}{dt}\right) = \phi^b(t)\,. \tag{2.35}$$

Da queste abbiamo che, per (2.33), vale:

$$\phi^t(s(t)) = -\frac{\frac{ds}{dt}}{\left|\frac{ds}{dt}\right|}\sqrt{\left[\frac{m}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2 - F^n\left(t,s,\frac{ds}{dt}\right)\right]^2 + F^b\left(t,s(t),\frac{ds}{dt}\right)^2}\,.$$

Inserendo l'espressione ottenuta per $\phi(s(t))$ nella (2.26) si ottiene l'equazione che non contiene le reazioni vincolari incognite, ma solo la funzione $s = s(t)$ incognita, ed è quindi l'equazione pura di movimento,

$$\frac{d^2s}{dt^2} = m^{-1}F^t\left(t,s(t),\frac{ds}{dt}\right) - \frac{\frac{ds}{dt}}{m\left|\frac{ds}{dt}\right|}\sqrt{\left[\frac{m}{\rho(s(t))}\left(\frac{ds}{dt}\right)^2 - F^n\left(t,s,\frac{ds}{dt}\right)\right]^2 + F^b\left(t,s(t),\frac{ds}{dt}\right)^2}\,. \tag{2.36}$$

Quest'equazione ricade ancora nel dal teorema 2.2 purché , oltre che la curva Γ, la funzione della forza attiva sia regolare sufficientemente e purché la soluzione $s = s(t)$ non annulli la sua derivata prima (si osservi che in virtù di ciò il coefficiente $\frac{ds}{dt}/|\frac{ds}{dt}|$ è costante e vale ± 1 durante il moto). Una volta determinata la funzione $s = s(t)$ per fissate condizioni iniziali, si ricavano la reazione vincolare in funzione del tempo, $\boldsymbol{\phi} = \boldsymbol{\phi}(t)$, ottenendola direttamente dalle equazioni (2.26)-(2.27), avendo sostituito in $s(t)$ la soluzione dell'equazione pura di movimento. Se per $t = t_1$ vale, per la soluzione considerata, $ds/dt|_{t=t_1} = 0$, il moto del punto si ferma in $s_1 = s(t_1)$ e,

a partire da quell'istante, il suo stato di moto deve essere studiato usando la relazione costitutiva (2.32), unitamente alle equazioni (2.26)-(2.27) semplificate in

$$
\begin{aligned}
-F^t\,(t, s_1, 0) &= \phi^t(t)\,, && (2.37)\\
-F^n\,(t, s_1, 0) &= \phi^n(t)\,, && (2.38)\\
-F^b\,(t, s_1, 0) &= \phi^b(t)\,. && (2.39)
\end{aligned}
$$

Fino a quando, al variare di t il vincolo:

$$
|F^t\,(t, s_1, 0)\,| < \mu_s \sqrt{F^n\,(t, s_1, 0)^2 + F^b\,(t, s_1, 0)^2}\,, \tag{2.40}
$$

è soddisfatto il punto rimane in quiete. Se per $t = t_2$ i due membri della disuguaglianza coincidono, il punto ricomincia a muoversi e il moto deve essere studiato con l'equazione (2.36) con condizioni iniziali $s(t_2) = s_1$ e $ds/dt|_{t=t_2} = 0$. Il coefficiente singolare $\frac{ds}{dt}/|\frac{ds}{dt}|$ deve essere rimpiazzato dal segno di $F^t\,(t_2, s_1, 0)$, perché si assume che il moto cominci nella direzione della forza che ha vinto l'attrito statico.

3. Passiamo a considerare il caso in cui un punto materiale P di massa m si muove vincolato ad una superficie S, che assumeremo sferica di raggio R e centro O (ma quando diciamo è generale) in quiete nel riferimento inerziale $\mathscr{I}$. Se $\mathbf{F} = \mathbf{F}(t, P, \mathbf{v}_P|_{\mathscr{I}})$ denota la legge di forza complessiva agente su P di espressione assegnata, escludendo la reazione vincolare $\boldsymbol{\phi}$, il secondo principio della dinamica fornisce:

$$
m\mathbf{a}_P = \mathbf{F}(t, P, \mathbf{v}_P) + \boldsymbol{\phi}(t)\,.
$$

Possiamo parametrizzare il punto sulla sfera con coordinate polari standard θ e ϕ ed esprimere queste in funzione del tempo $\theta = \theta(t)$, $\varphi = \varphi(t)$. Cercheremo l'equazione pura di movimento in queste variabili una volta imposta una caratterizzazione costitutiva del vincolo. Ancora una volta è conveniente decomporre questa equazione nella terna di vettori associata al vincolo, data in questo caso dalla (1.42). In definitiva il secondo principio si traduce in, dove è sottointeso che $\theta = \theta(t)$, $\varphi = \varphi(t)$ e, per brevità notazionale, le derivate rispetto al tempo sono indicate con il punto:

$$
\begin{aligned}
&m(R\ddot{\theta} - R\dot{\varphi}^2 \sin\theta\cos\theta)\,\mathbf{e}_\theta + m(R\ddot{\varphi}\sin\theta + 2R\dot{\varphi}\dot{\theta}\cos\theta)\,\mathbf{e}_\varphi - m(R\dot{\theta}^2 + R\dot{\varphi}^2\sin^2\theta)\mathbf{n}\\
&= F^n(t, P((\theta,\varphi)), \mathbf{v}_P(\theta,\varphi))\,\mathbf{e}_r + F^\theta(t, P((\theta,\varphi)), \mathbf{v}_P(\theta,\varphi))\,\mathbf{e}_\theta + F^\varphi(t, P(\theta,\varphi), \mathbf{v}_P(\theta,\varphi))\mathbf{e}_\varphi\\
&\qquad + \phi^n(t)\,\mathbf{e}_r + \phi^\theta(t)\mathbf{n}(s(t))\,\mathbf{e}_\theta + \phi^\varphi(t)\mathbf{e}_\varphi
\end{aligned}
$$

Separando le equazioni sulle tre componenti e facendo uso di (1.41)

$$
\mathbf{v}_P = R\dot{\theta}\,\mathbf{e}_\theta + R\dot{\phi}\sin\theta\,\mathbf{e}_\varphi\,,
$$

si trova infine il seguente sistema di equazioni differenziali

$$m\left(R\frac{d^2\theta}{dt^2} - R\left(\frac{d\varphi}{dt}\right)^2 \sin\theta(t)\ \cos\theta(t)\right) = F^\theta\left(t, \theta(t), \varphi(t)), \frac{d\theta}{dt}, \frac{d\varphi}{dt}\right) + \phi^\theta(t)\,, \quad (2.41)$$

$$m\left(R\frac{d^2\varphi}{dt^2}\sin\theta(t) + 2R\frac{d\theta}{dt}\frac{d\varphi}{dt}\cos\theta(t)\right) = F^\varphi\left(t, \theta(t), \varphi(t)), \frac{d\theta}{dt}, \frac{d\varphi}{dt}\right) + \phi^\varphi(t)\,, \quad (2.42)$$

$$-m\left(R\left(\frac{d\theta}{dt}\right)^2 + R\left(\frac{d\varphi}{dt}\right)^2 \sin^2\theta(t)\right) = F^n\left(t, \theta(t), \varphi(t)), \frac{d\theta}{dt}, \frac{d\varphi}{dt}\right) + \phi^n(t)\,. \quad (2.43)$$

La caratterizzazione costitutiva di **superficie liscia** richiede che le componenti di $\boldsymbol{\phi}$ tangenti alla superficie siano nulle. Inserendo questa richiesta nel sistema di sopra si trova il **sistema di equazioni pure di movimento**

$$m\left(R\frac{d^2\theta}{dt^2} - R\left(\frac{d\varphi}{dt}\right)^2 \sin\theta(t)\ \cos\theta(t)\right) = F^\theta\left(t, \theta(t), \varphi(t)), \frac{d\theta}{dt}, \frac{d\varphi}{dt}\right)\,, \quad (2.44)$$

$$m\left(R\frac{d^2\varphi}{dt^2}\sin\theta(t) + 2R\frac{d\theta}{dt}\frac{d\varphi}{dt}\cos\theta(t)\right) = F^\varphi\left(t, \theta(t), \varphi(t)), \frac{d\theta}{dt}, \frac{d\varphi}{dt}\right)\,, \quad (2.45)$$

in cui non appare la reazione vincolare incognita ma solo le funzioni $\theta = \theta(t)$ e $\varphi = \varphi(t)$ che individuano il moto del punto. In effetti, il sistema di sopra, con debite ipotesi di regolarità ammette una sola soluzione per fissate condizioni iniziali $\theta(t_0) = \theta_0$ e $\varphi(t_0) = \varphi_0$ in quanto ricade nei casi considerati dal teorema[5] 2.2. La reazione vincolare, ridotta alla sola componente normale alla superficie si ottiene in funzione del tempo, quando è stato determinato il moto, direttamente dalla (2.43). Come commento finale dobbiamo sottolineare che le coordinate θ, φ *non ricoprono completamente la sfera*, pertanto lo studio completo del moto del punto sulla sfera richiede l'utilizzo di più di una carta.

2.3.3 Dinamica in riferimenti non inerziali: forze inerziali.

Cosa succede se si tenta di riformulare le equazioni della dinamica in un riferimento non inerziale $\mathscr{I}'$ di cui è assegnato il moto rispetto al riferimento inerziale $\mathscr{I}$? È possibile ripristinare lo schema basato sul secondo principio della dinamica in modo da avere ancora un sistema di equazioni differenziali che assicuri, sotto certe ipotesi, il determinismo?
Per semplicità trattiamo il caso di un punto materiale P. Se $\mathscr{I}$ è un riferimento inerziale avremo, al solito, che il secondo principio della dinamica (tenendo conto delle varie estensioni descritte sopra) si esprime come:

$$m_P \mathbf{a}_P|_{\mathscr{I}} = \mathbf{F}\left(t, P, \mathbf{v}_P|_{\mathscr{I}}\right) + \boldsymbol{\phi}\,, \quad (2.46)$$

[5] Si osservi che se $sin\theta = 0$ il sistema di equazioni non può essere scritto in forma normale, tuttavia $\sin\theta = 0$ definisce punti, $\theta = 0, \pi$ in cui il sistema di coordinate polari sferiche non è ben definito: dove le coordinate sono ben definite, il sistema è scrivibile in forma normale.

Abbiamo scritto esplicitamente $|_{\mathscr{I}}$ per enfatizzare il fatto che l'accelerazione è valutata rispetto al riferimento (inerziale) $\mathscr{I}$. Possiamo ora esprimere velocità ed accelerazione di P rispetto a $\mathscr{I}$ in termini delle analoghe quantità valutate rispetto a $\mathscr{I}'$ per mezzo delle equazioni (1.69)-(1.70). Se $O' \in E_{\mathscr{I}'}$ è in quiete con $\mathscr{I}'$ si ha che:

$$m_P \mathbf{a}_p|_{\mathscr{I}'} + m_P \left(\mathbf{a}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O')) + \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + 2\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} \right)$$
$$= \mathbf{F}\left(t, P, \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') + \mathbf{v}_{O'}|_{\mathscr{I}}\right) + \boldsymbol{\phi}$$

Conviene riscrivere l'equazione di sopra come:

$$m_P \mathbf{a}_p|_{\mathscr{I}'} = \mathbf{F}\left(t, P, \mathbf{v}_P|_{\mathscr{I}'}\right) + \boldsymbol{\phi}(t) + \mathbf{F}_{\mathscr{I}'}\left(t, P, \mathbf{v}_P|_{\mathscr{I}'}\right) \ , \tag{2.47}$$

avendo definito

$$\mathbf{F}_{\mathscr{I}'}\left(t, P, \mathbf{v}_P|_{\mathscr{I}'}\right) := -m_P \mathbf{a}_{O'}|_{\mathscr{I}}(t) - m_P \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (P - O')) - 2m_P \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \mathbf{v}_P|_{\mathscr{I}'}$$
$$- m_P \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O') \ . \tag{2.48}$$

Con una certa improprietà di scrittura, sopra $\mathbf{F}\left(t, P, \mathbf{v}_P|_{\mathscr{I}'}\right)$ indica in realtà la funzione a valori vettoriali, $\mathbf{F}\left(t, P, \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (P - O'(t)) + \mathbf{v}_{O'}|_{\mathscr{I}}(t)\right)$, in cui *le funzioni del tempo:* $\mathbf{v}_{O'}|_{\mathscr{I}} = \mathbf{v}_{O'}|_{\mathscr{I}}(t)$ *e* $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} = \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)$ *sono note, dato che il moto di* $\mathscr{I}'$ *rispetto a* $\mathscr{I}$ *è assegnato.*

Da tutto ciò si evince che: se vogliamo riavere la formulazione della dinamica in termini di "$F = ma$" in un sistema di riferimento non inerziale $\mathscr{I}'$, compatibilmente con la meccanica sviluppata in sistemi di riferimento inerziali, siamo costretti ad aggiungere delle nuove funzioni di forza. Queste nuove funzioni di forza, rappresentate dal secondo membro di (2.48), sono correttamente funzione della posizione e della velocità del punto nel riferimento *non inerziale.* Ciò assicura la il determinismo in base al teorema 2.2 purché tali funzioni siano sufficientemente regolari. Tali funzioni dipendono dal moto, che deve essere noto, del riferimento non inerziale $\mathscr{I}'$ rispetto ad un riferimento inerziale ($\mathscr{I}$ nel nostro caso).

Le funzioni di forza $\mathbf{F}_{\mathscr{I}'}$ vengono dette **forze apparenti** o **forze inerziali** agenti su P; per contrasto, quelle introdotte nelle sezioni precedenti lavorando in riferimenti inerziali (includendo le reazioni vincolari) si dicono **forze vere**. Si osservi che le forze inerziali su P *non soddisfano il principio di azione e reazione*, non ha senso dire che siano causate da qualche altro punto materiale P' (qualche altro sistema materiale) ed, infine, hanno valori che *dipendono dal riferimento* al contrario delle forze vere che hanno lo stesso valore in ogni riferimento, *inerziale e non.*

Nell'ambito della meccanica classica le forze inerziali non rappresentano interazioni, ma sono solo un espediente matematico. Nell'ambito della Teoria della Relatività Generale questo punto di vista verrà completamente ribaltato, mostrando che le forze inerziali, o meglio i corrispondenti matematici della nuova formulazione della dinamica, descrivono un'interazione che ha la stessa natura dell'interazione gravitazionale.

Esiste una classificazione più specifica delle forze inerziali:

$$\mathbf{F}_c(t, P) := -m_P \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (P - O'))$$

si dice **forza centrifuga** mentre

$$\mathbf{F}^{(Coriolis)}(t, \mathbf{v}_P|_{\mathscr{I}'}) := -2m_P \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \mathbf{v}_P|_{\mathscr{I}'}$$

si dice **forza di Coriolis**. Si osservi che la seconda non agisce se il punto materiale P ha velocità nulla in $\mathscr{I}'$. La prima invece, pensando $\mathbf{F}_c$ applicata a P e $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ uscente da O', è nel piano contenente $P - O'$ e $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$, è perpendicolare a $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ e punta nella direzione uscente dall'asse $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$. In fine la sua norma vale $m||\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}||^2 R$ dove $R > 0$ è il *braccio* di $P - O'$ rispetto a $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$ (pari a $||P - O'||\,|\sin\alpha|$ dove α è l'angolo tra $P - O'$ e $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$).

Esempi 2.4.
1. L'esempio più semplice che possiamo dare è sicuramente il seguente. Consideriamo un punto materiale P di massa m immobile nel riferimento inerziale $\mathscr{I}$. Consideriamo un secondo riferimento $\mathscr{I}'$, non inerziale, il cui moto rispetto a $\mathscr{I}$ è di pura rotazione uniforme attorno al'asse z, con velocità angolare $\omega > 0$. In altre parole, rispetto a coordinate cartesiane ortogonali x, y, z e x', y', z' solidali rispettivamente con $\mathscr{I}$ e $\mathscr{I}'$, valgono le trasformazioni $z' = z$, $x = x'\cos(\omega t) + y'\sin(\omega t)$ e $y = -x'\sin(\omega t) + y'\cos(\omega t)$. In questo modo $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} = \omega\,\mathbf{e}_z$. Nel riferimento $\mathscr{I}'$ il moto di P è descritto come un moto rotatorio uniforme attorno al punto $O \equiv O'$ di coordinate $(0, 0, 0)$ per entrambi i riferimenti. Se $(x_p, y_p, 0)$ sono le coordinate di P per $\mathscr{I}'$, il punto P in $\mathscr{I}$ ha legge oraria $x'_P(t) = x\cos(\omega t) - y\sin(\omega t)$, $y'_P(t) = x\sin(\omega t) + y\cos(\omega t)$ (e $z'_P(t) = 0$). Quindi, in particolare $\mathbf{v}_{\mathscr{I}'}(t) = \boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'} \wedge (P(t) - O)$.
Sappiamo dalla fisica elementare che, dal punto di vista dinamico, il moto rotatorio uniforme per il punto materiale P avviene se la forza complessiva che agisce su P è *centripeta*, cioé parallela ed equiversa a $-(P(t) - O)$[6], ed ha modulo $\omega^2 R$, dove R è la distanza da P dall'asse di rotazione. La forza apparente $\mathbf{F}_c(t, P)$ è invece *centrifuga*, per cui da sola non è in grado di fornire la forza centripeta necessaria. In realtà il calcolo diretto di tutta la forza inerziale che appare in $\mathscr{I}'$ tramite la (2.48) produce direttamente

$$\mathbf{F}_{\mathscr{I}'}(t, P, \mathbf{v}_P|_{\mathscr{I}'}) := -m\mathbf{a}_O|_{\mathscr{I}}(t) - m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O)) - 2m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'}$$
$$-m\dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O)$$

(ricordare che $O \equiv O'$ nella situazione considerata) dove il primo e l'ultimo addendo a secondo membro sono nulli per ipotesi, mentre la forza centrifuga e quella di Coriolis sono uguali, rispettivamente a (notare che $\omega \perp (P - O)$ nella situazione in esame):

$$\mathbf{F}_c(t, P) = -m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P(t) - O)) = -m\omega\,\mathbf{e}_z \wedge (\omega\,\mathbf{e}_z \wedge (P(t) - O)) = m\omega^2(P(t) - O)\,,$$

$$\mathbf{F}^{(Coriolis)}(t, \mathbf{v}_P|_{\mathscr{I}'}) = -2m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} = -2m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}'} \wedge (P(t) - O)) = -2m\omega^2(P(t) - O).$$

La somma delle due forze produce proprio la forza centripeta necessaria pari a $-m\omega^2(P(t) - O)$.
2. Consideriamo un punto materiale P di massa m vincolato alla circonferenza liscia Γ di raggio R centrata nell'origine O', e giacente nel piano $y' = 0$, di un sistema di coordinate cartesiane

[6] Attenzione che questo è vero nel caso considerato, perché $P - O$ è perpendicolare a $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$.

ortonormali x', y', x' solidali con il sistema di riferimento $\mathscr{I}'$. Tale riferimento non è inerziale e vale $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} = \omega\, \mathbf{e}_{z'}$, con $\omega > 0$ costante, dove $\mathscr{I}$ è un riferimento inerziale con coordinate cartesiane ortonormali solidali x, y, z di origine $O \equiv O'$ e con asse z sempre coincidente con l'asse z' di $\mathscr{I}'$. Si suppone che P sia sottoposto alla forza peso $-mg\, \mathbf{e}_{z'}$, oltre che alla reazione vincolare, $\boldsymbol{\phi}$, dovuta a Γ. Vogliamo ricavare l'equazione pura di movimento del punto nel riferimento $\mathscr{I}'$.
Conviene usare la terna intrinseca di Γ. Parametrizziamo la circonferenza con l'angolo φ tra $P - O$ e $\mathbf{e}_z = \mathbf{e}_{z'}$, orientato positivamente rispetto a $\mathbf{e}_{y'}$. Con tale scelta per l'angolo φ, introduciamo coordinate polari piane φ, r nel piano z', x'. Il versore tangente $\mathbf{t}$ coincide con $\mathbf{e}_\varphi$, il vettore normale $\mathbf{n}$ coincide con $-\mathbf{e}_r$ ed il vettore binormale coincide con $\mathbf{e}_{y'}$. Il raggio di curvatura vale semplicemente $\rho = R$ costantemente. Un calcolo immediato fornisce la funzione ascissa curvilinea $s = s(\varphi)$ nella forma $s = R\varphi$, avendo è scelto la sua origine nel punto più in alto in cui Γ interseca l'asse z'.
Le equazioni pure di movimento si ottengono proiettando l'espressione del secondo principio della dinamica lungo il versore tangente a Γ, dopo avere espresso velocità ed accelerazione in termini dell'ascissa curvilinea. L'espressione del secondo principio della dinamica è dunque, nel riferimento non inerziale $\mathscr{I}'$,

$$m\mathbf{a}_P|_{\mathscr{I}'} = -mg\, \mathbf{e}_{z'} + \boldsymbol{\phi} + \mathbf{F}_{\mathscr{I}'}$$

e, passando alla coordinata ascissa curvilinea come in (1) in esempi 2.3,

$$m\frac{d^2 s}{dt^2}\ \mathbf{e}_\varphi(s(t)) - \frac{m}{R}\left(\frac{ds}{dt}\right)^2\ \mathbf{e}_r(s(t)) = -mg\, \mathbf{e}_{z'} + \boldsymbol{\phi}(t) + \mathbf{F}_{\mathscr{I}'}(t, P(t), \mathbf{v}_{\mathscr{I}'}(t))\,.$$

L'equazione pura di movimento si ottiene prendendo il prodotto scalare di ambo membri con $\mathbf{e}_\varphi = -\sin\varphi\, \mathbf{e}_{z'} + \cos\varphi\, \mathbf{e}_{x'}$. In questo modo, tenendo conto che la reazione vincolare $\boldsymbol{\phi}$ è normale a Γ e che $\varphi = s/R$, si ha

$$m\frac{d^2 s}{dt^2} = mg\sin\left(\frac{s(t)}{R}\right) + [-m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P - O)) - 2m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'}] \cdot\ \mathbf{e}_\varphi\,, \quad (2.49)$$

dove abbiamo trascurato il primo e l'ultimo addendo a secondo membro di (2.48), essendo essi nulli nelle nostre ipotesi. Vale ancora:

$$P(t) - O = R\left(\sin\left(\frac{s(t)}{R}\right)\ \mathbf{e}_{x'} + \cos\left(\frac{s(t)}{R}\right)\ \mathbf{e}_{z'}\right)$$

da cui

$$\mathbf{v}_P|_{\mathscr{I}}(t) = \frac{ds}{dt}\left(\cos\left(\frac{s(t)}{R}\right)\ \mathbf{e}_{x'} - \sin\left(\frac{s(t)}{R}\right)\ \mathbf{e}_{z'}\right)\,.$$

Essendo $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} = \omega\, \mathbf{e}_{z'}$, il prodotto scalare a secondo membro in (2.49) vale, con qualche passaggio (in particolare l'ultimo addendo nel prodotto scalare in (2.49) fornisce contributo nullo),

$$m\omega^2 R\frac{ds}{dt}\sin\left(\frac{s(t)}{R}\right)\ \cos\left(\frac{s(t)}{R}\right)\,.$$

L'equazione pura di movimento risulta quindi essere:

$$m\frac{d^2s}{dt^2} - m\omega^2 R\,\sin\left(\frac{s(t)}{R}\right)\,\cos\left(\frac{s(t)}{R}\right) + mg\sin\left(\frac{s(t)}{R}\right) = 0\,. \tag{2.50}$$

Esercizi 2.3.

1. Si considerino due punti materiali P e Q, entrambi di massa m, vincolati sulla curva liscia Γ di equazione $x = \cos u$, $y = \sin u$, $z = u$ con $u \in \mathbb{R}$, dove x, y, z sono coordinate cartesiane ortonormali solidali con un riferimento inerziale $\mathscr{I}$. Si supponga che i punti siano sottoposti, oltre alla reazione vincolare dovuta alla curva Γ, alla forza peso $-mg\,\mathbf{e}_z$ e che siano legati l'un l'altro attraverso una molla ideale di lunghezza nulla a riposo e costante elastica κ. Si scriva il sistema delle equazioni pure di movimento dei due punti.

2. In riferimento all'esempio 2.4.**2**, si dimostri che la quantità :

$$E = \frac{1}{2}m\left(\frac{ds}{dt}\right)^2 + \frac{m\omega^2R^2}{4}\,\cos\left(\frac{2s(t)}{R}\right) - mgR\cos\left(\frac{s(t)}{R}\right)\,,$$

è costante durante ogni moto del sistema (il valore costante dipende dalla soluzione considerata). Qual'è il significato fisico di E?

2.4 Alcuni commenti sulla formulazione generale sulla dinamica newtoniana

2.4.1 Invarianza galileiana della meccanica classica.

Per come sono state formulate, le leggi della dinamica C2-C4 hanno la proprietà notevole di avere la stessa forma quando sono scritte in coordinate cartesiane ortonormali solidali con un sistema di riferimento inerziale, al variare del sistema inerziale. La prova è immediata considerando le varie leggi separatamente. Tale proprietà di invarianza viene estesa anche alla forma funzionale delle forze ed ha un significato fisico profondo. Vediamo come si impone e come si formalizza tale principio. Torniamo al caso di due punti materiali, quanto segue si può banalmente estendere al caso di più punti materiali usando il principio di sovrapposizione delle forze. Ora interpreteremo l'azione del gruppo di Galileo in senso *attivo*. In altre parole interpretando la generica trasformazione di Galileo:

$$\begin{cases} t' = & t + c\,, \\ x'^i = & c^i + tv^i + \displaystyle\sum_{j=1}^{3} R^i{}_j x^j\,, \quad \text{i=1,2,3.} \end{cases} \tag{2.51}$$

stiamo usando un *unico* sistema di coordinate cartesiane ortogonali solidali con un *fissato* riferimento inerziale $\mathscr{I}$: le coordinate non primate sono quelle di un evento p, quelle primate (nello stesso sistema di coordinate) sono quelle dell'evento p' (diverso dal primo) ottenuto dopo aver

fatto agire la trasformazione su p. Il gruppo di Galileo è composto da alcuni sottogruppi che elenchiamo[7]. Il gruppo delle **traslazioni temporali**, che include le trasformazioni, con $c \in \mathbb{R}$:

$$\begin{cases} t' = & t + c\,, \\ x'^i = & x^i\,, \quad \text{i=1,2,3}\,; \end{cases} \tag{2.52}$$

il gruppo delle **traslazioni spaziali**, corrispondenti alle trasformazioni, per ogni terna di numeri $(c^1, c^2, c^3) \in \mathbb{R}^3$,

$$\begin{cases} t' = & t\,, \\ x'^i = & c^i + x^i\,, \quad \text{i=1,2,3}\,. \end{cases} \tag{2.53}$$

Si osservi che tale gruppo altro non è che il gruppo delle traslazioni standard $V_{\mathscr{I}}$ dello spazio euclideo di quiete $E_{\mathscr{I}}$, essendo (c^1, c^2, c^3) la rappresentazione in componenti di una traslazione. Il gruppo delle **rotazioni**, corrispondenti alle trasformazioni, con $R \in O(3)$,

$$\begin{cases} t' = & t\,, \\ x'^i = & \displaystyle\sum_{j=1}^{3} R^i{}_j x^j\,, \quad \text{i=1,2,3}\,; \end{cases} \tag{2.54}$$

ed, infine, il gruppo delle **trasformazioni pure di Galileo**, corrispondenti alle trasformazioni, con $(v^1, v^2, v^3) \in \mathbb{R}^3$:

$$\begin{cases} t' = & t\,, \\ x'^i = & tv^i + x^i\,, \quad \text{i=1,2,3.} \end{cases} \tag{2.55}$$

Nell'assegnare la forma generale delle funzioni di forza che si esercitano *due punti materiali isolati dal resto da tutti gli altri*, si è assunta l'indipendenza esplicita dal tempo assoluto t. Tale è equivalente all'invarianza delle funzioni di forza sotto l'azione delle traslazioni temporali: per ogni $\Delta t \in \mathbb{R}$,

$$\mathbf{F}\left(t + \Delta t, Q, Q', \mathbf{v}_Q|_{\mathscr{I}}, \mathbf{v}_{Q'}|_{\mathscr{I}}\right) = \mathbf{F}\left(t, Q, Q', \mathbf{v}_Q|_{\mathscr{I}}, \mathbf{v}_{Q'}|_{\mathscr{I}}\right)\ .$$

Questa richiesta di invarianza cade sotto il nome di postulato di **omogeneità temporale**. Analogamente si assumono altri tre postulati. Uno è quello di **omogeneità spaziale** che richiede che le funzioni forza siano invarianti sotto l'azione del gruppo delle traslazioni spaziali. In altre parole, per ogni vettore $\mathbf{u}$ deve valere:

$$\mathbf{F}\left(t, Q + \mathbf{u}, Q' + \mathbf{u}, \mathbf{v}_Q|_{\mathscr{I}}, \mathbf{v}_{Q'}|_{\mathscr{I}}\right) = \mathbf{F}\left(t, Q, Q', \mathbf{v}_Q|_{\mathscr{I}}, \mathbf{v}_{Q'}|_{\mathscr{I}}\right)\ .$$

È ovvio che ciò implichi che le funzioni di forza possano dipendere *solo dai vettori differenza* $Q(t) - Q'(t)$, cioè dalle posizioni relative dei punti materiali ad ogni fissato tempo. Un altro

[7] Ricordiamo che un sottoinsieme S di un gruppo G è detto **sottogruppo** di G se è chiuso rispetto alla legge di composizione gruppale, rispetto al calcolo dell'elemento inverso e contiene l'elemento neutro di G. In tal caso ovviamente, S è a sua volta un gruppo rispetto alla legge di composizione di G. La verifica che gli insiemi di trasformazioni seguenti formano ciascuno un sottogruppo del gruppo di Galileo è immediata.

postulato è quello dell'**isotropia spaziale** che richiede che se $g_{O,R}$ è una qualsiasi isometria, associata alla matrice ortogonale R ed al punto O che ruota i punti di Σ_t attorno ad $O \in \Sigma_t$, della rotazione R (in altre parole $g_{O,R}(P) := O + R(P - O)$), valga[8]:

$$\mathbf{F}\left(t, g_{O,R}(Q), g_{O,R}(Q'), R\mathbf{v}_Q|_{\mathscr{I}}, R\mathbf{v}_{Q'}|_{\mathscr{I}}\right) = (R\mathbf{F})\left(t, Q, Q', \mathbf{v}_Q|_{\mathscr{I}}, \mathbf{v}_{Q'}|_{\mathscr{I}}\right) .$$

Ovviamente questo requisito corrisponde all'invarianza delle funzioni forza sotto l'azione del gruppo delle rotazioni. Tale principio non è valido in questa versione elementare se si assume una nozione più complessa di punto materiale in cui al punto è associato, nel suo sistema di quiete, qualche vettore di significato fisico (momento elettrico, momento magnetico, *spin*, ecc.)
Infine il requisito di **invarianza galileiana pura** assume che le leggi di forza siano invarianti sotto trasformazioni di Galileo pure. Per ogni vettore $\mathbf{V}$ deve essere:

$$\mathbf{F}\left(t, Q + t\mathbf{V}, Q' + t\mathbf{V}, \mathbf{v}_Q|_{\mathscr{I}} + \mathbf{V}, \mathbf{v}_{Q'}|_{\mathscr{I}} + \mathbf{V}\right) = \mathbf{F}\left(t, Q, Q', \mathbf{v}_Q|_{\mathscr{I}}, \mathbf{v}_{Q'}|_{\mathscr{I}}\right) .$$

Tale principio, unitamente al principio di omogeneità spaziale, impone che nelle funzioni di forza appaia solo la *differenza delle velocità* dei punti materiali rispetto a $\mathscr{I}$.
Dato che ogni elemento del gruppo di Galileo si costruisce moltiplicando elementi dei 4 sottogruppi menzionati sopra, risulta che le leggi di forza di coppie di punti materiali (isolati dal resto dell'universo) siano invarianti sotto l'azione di ogni elemento del gruppo di Galileo.
La proprietà di invarianza sotto il gruppo di Galileo viene assunta, in meccanica classica, a rango di *principio generale* che deve essere soddisfatto da ogni legge formulata in meccanica (non solo dalle leggi di forza):

C5. Principio di relatività galileiana. *Tutte le leggi delle meccanica classica, in particolare le leggi di forza, sono invarianti sotto l'azione attiva del gruppo di Galileo.*

Osservazioni 2.5.
(1) Il principio di relatività galileiana ha un'importantissima conseguenza sulle soluzioni delle equazioni del moto per un sistema di punti materiali. Consideriamo un esperimento in cui si studiano n punti materiali in coordinate cartesiane ortonormali solidali con un sistema di riferimento inerziale $\mathscr{I}$. Fissiamo delle condizioni iniziali per i punti assegnando posizioni e velocità di essi al tempo generico t_0 e consideriamo la soluzione ottenuta dalle equazioni di Newton tenendo conto del principio di sovrapposizione. Dal principio di invarianza galileiana e dal fatto che le equazioni di Newton ammettono una sola soluzione quando sono assegnate condizioni iniziali, segue facilmente il seguente risultato (la cui prova è lasciata per esercizio). Se facciamo agire una trasformazione g del gruppo di Galileo su una soluzione del problema del moto (quindi la stessa trasformazione su ogni punto e per ogni istante), l'evoluzione temporale degli n punti individuata in questo modo non è altro che la soluzione dello stesso problema del moto, ma con condizioni iniziali trasformate secondo l'azione di g.
(2) Esiste una formulazione "duale" del principio di invarianza galileiana, basata sul fatto che

[8] Si tenga conto che $\mathbf{F}$ è un vettore, per cui l'azione delle rotazioni agisce anche sulla sua direzione. L'equazione dice semplicemente che "ruotare gli argomenti della funzione $\mathbf{F}$ è equivalente a ruotare il vettore $\mathbf{F}$".

ogni trasformazione del gruppo di Galileo si può interpretare passivamente: come legge di trasformazione tra le coordinate di due differenti sistemi di riferimento inerziali. È facile convincersi che, da questo punto di vista, il principio suddetto può essere riformulato come segue.

C5.' Principio di relatività galileiana. *Tutte le leggi della meccanica hanno la stessa forma in coordinate cartesiane ortonormali solidali con ogni sistema di riferimento inerziale. In particolare, se in coordinate cartesiane ortonormali solidali con $\mathscr{I} \in \mathfrak{I}$ la legge di forza sul k-esimo punto materiale è definita dalle funzioni componenti*

$$f^i_{Q_k} = \mathfrak{F}^i(x^1_{Q_1}, x^2_{Q_1}, x^3_{Q_1}, \ldots x^1_{Q_N}, x^2_{Q_N}, x^3_{Q_N}, v^1_{Q_1}, v^2_{Q_1}, v^3_{Q_1} \ldots, v^1_{Q_N}, v^2_{Q_N}, v^3_{Q_N}) \,,$$

nelle coordinate cartesiane ortonormali solidali con un altro $\mathscr{I}' \in \mathfrak{I}$, la legge di forza sul k-esimo punto ha ancora componenti della forma

$$f'^i_{Q_k} = \mathfrak{F}^i(x'^1_{Q_1}, x'^2_{Q_1}, x'^3_{Q_1}, \ldots x'^1_{Q_N}, x'^2_{Q_N}, x'^3_{Q_N}, v'^1_{Q_1}, v'^2_{Q_1}, v'^3_{Q_1} \ldots, v'^1_{Q_N}, v'^2_{Q_N}, v'^3_{Q_N}) \,,$$

essendo le funzioni $\mathfrak{F}^i$ le stesse di sopra.

Ora l'osservazione (1) può essere riformulata in *termini passivi* come segue. Consideriamo un esperimento in cui si studiano n punti materiali (tra di essi interagenti ma isolati con l'esterno) in coordinate cartesiane ortonormali solidali con un sistema di riferimento inerziale $\mathscr{I}$. Fissiamo delle condizioni iniziali per i punti assegnando posizioni e velocità di essi al tempo generico t_0 e consideriamo la soluzione ottenuta dalle equazioni di Newton tenendo conto del principio di sovrapposizione. Passiamo ad un nuovo riferimento inerziale $\mathscr{I}' \neq \mathscr{I}$, assegnando, per lo stesso sistema fisico di prima (isolato all'esterno ma internamente interagente), condizioni iniziali, che in componenti siano identiche alle componenti delle condizioni iniziali date in $\mathscr{I}$. Risulta che l'evoluzione del sistema fisico, letta nelle coordinate di $\mathscr{I}'$ con le nuove condizioni iniziali, sarà *esattamente* la stessa che nel primo caso letta in coordinate di $\mathscr{I}$ e con le vecchie condizioni iniziali. In questo senso *non è in alcun modo possibile privilegiare un riferimento inerziali rispetto agli altri tramite i risultati di esperimenti di pura meccanica.*

In definitiva vediamo che il principio di invarianza galileiano sancisce che tutti i riferimenti inerziali sono completamente equivalenti ai fini della formulazione e dell'applicazione della meccanica classica e che quindi non ha alcun senso fisico privilegiarne uno nei confronti degli altri. Per questo motivo lo stesso principio viene spesso enunciato dai fisici in modo matematicamente più ambiguo come segue:

C5". Principio di invarianza galileiana.(Formulazione fisica). *Non è in alcun modo possibile privilegiare un sistema di riferimento nella classe dei sistemi di riferimento inerziali tramite i risultati di esperimenti di pura meccanica.*

(3) In una formulazione più avanzata della meccanica e della fisica in generale, si può mostrare il legame profondo che connette le proprietà di invarianza sotto gruppi di trasformazione, con l'esistenza di quantità fisiche che si conservano durante l'evoluzione del sistema (in particolare: impulso, momento angolare ed energia). Vedremo ciò diffusamente nella parte di queste dispense relative al teorema di Nother in meccanica lagrangiana.

2.4.2 *Un commento sul cosiddetto "Principio di Mach".

Nella visone di Mach ed Einstein il moto inerziale a velocità costante di un punto materiale isolato, cioé lontano da tutti gli altri corpi dell'universo deve essere in qualche modo *imposto* delle altre masse dell'universo, sia pur lontanissime, tramite qualche forma di interazione. Questo in essenza è il contenuto fisico del "principio di Mach".
In ogni caso, tale interazione *non può essere descritta in termini di forze* perché queste ultime descrivono, per definizione, le interazioni tra corpi "vicini". Per tale motivo il principio d'inerzia non può avere spiegazione dinamica all'interno della meccanica classica e deve essere assunto come principio.
Nell'ottica del Principio di Mach i sistemi di riferimento inerziali altro non sarebbero che "segnaposti" per le masse lontane, e solo rispetto a tali corpi lontani, o rispetto al loro "moto medio", avviene il moto rettilineo uniforme dei punti materiali isolati[9].
Per Mach ed Einstein i sistemi di riferimento inerziali non avrebbero quindi senso di esistere se l'universo fosse vuoto o contenesse un unico corpo (al contrario di Newton che dichiarò nei *Principia* che essi esisterebbero ugualmente nel famoso discorso riguardo al cosiddetto "secchio di Newton"). Einstein speculò a lungo sul significato fisico un'eventuale interazione che rendesse conto del moto inerziale, pensando che avesse la stessa natura della gravitazione. Tale idea però non può essere sviluppata nella meccanica classica proprio perché l'interazione gravitazionale è classicamente descritta da una forza. Anche tramite questo tipo di speculazioni Einstein giunse a formulare la teoria della Relatività Generale dove, almeno in parte, il Principio di Mach può essere sviluppato anche se non completamente.

[9] In relazione con ció si osservi che l'esperienza prova che un ottimo sistema di riferimento inerziale è quello in quiete con le lontanissime cosiddette "stelle fisse".

Capitolo 3

Introduzione alla teoria delle equazioni differenziali ordinarie

In questo capitolo introdurremo la teoria elementare dei *sistemi dinamici* in termini di *sistemi di equazioni differenziali ordinarie*. Dal punto di vista fisico, gli strumenti ed i risultati matematici che introdurremo sono fondamentali per dare senso all'idea di *determinismo* della meccanica classica. Infatti, uno dei fini sarà quello di dimostrare il teorema 2.2 che estenderemo, indebolendo le ipotesi e potenziando la tesi. La dimostrazione di tale teorema sancisce il successo della dinamica newtoniana ai fini della descrizione e della previsione dei fenomeni fisici nell'ambito della meccanica classica.
In tutto il capitolo $\mathbb{K}$ indica indifferentemente il campo $\mathbb{R}$ oppure $\mathbb{C}$. In entrambi i casi, $\mathbb{K}^n$ indica lo spazio vettoriale con campo $\mathbb{K}$ e di dimensione n costruito in modo standard.

3.1 Sistemi di equazioni differenziali

La nozione fondamentale è quella di *sistema di equazioni differenziali del prim'ordine in forma normale*.
Sia $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ (con $N = 1, 2, \ldots$, fissato) un insieme aperto non vuoto e sia $\mathbf{f} : \Omega \to \mathbb{K}^n$ una funzione arbitraria. Un **sistema di equazioni differenziali ordinarie del prim'ordine in forma normale** è un'equazione:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}\left(t, \mathbf{x}(t)\right) \ , \tag{3.1}$$

per la quale si cercano **soluzioni**, cioé funzioni differenziabili della forma $\mathbf{x} : I \to \mathbb{K}^n$, con $I \subset \mathbb{R}$ intervallo aperto, che risolvono (3.1) su I.
Il sistema è detto propriamente **equazione differenziale ordinaria del prim'ordine in forma normale** se $n = 1$. Il sistema è detto **autonomo** se $\mathbf{f}$ non dipende esplicitamente dalla variabile t. Infine, se $\mathbf{x} : I \to \mathbb{K}^n$ è una soluzione di (3.1):
(a) l'insieme $\{(t, \mathbf{x}(t)) \in \mathbb{R} \times \mathbb{K}^n \mid t \in I\} \subset \Omega$ è detto **curva integrale** associata alla soluzione;

(b) l'insieme $\{\mathbf{x}(t) \in \mathbb{K}^n \mid t \in I\}$ è detto **orbita** associata alla soluzione.

3.1.1 Riduzione al prim'ordine.

Si può definire analogamente un sistema di equazioni differenziali ordinarie di grado $k > 1$ in forma normale come:

$$\frac{d^k\mathbf{x}}{dt^k} = \mathbf{f}\left(t, \mathbf{x}(t), \frac{d\mathbf{x}}{dt}, \ldots, \frac{d^{k-1}\mathbf{x}}{dt^{k-1}}\right) , \tag{3.2}$$

in cui la funzione incognita soluzione $\mathbf{x} = \mathbf{x}(t)$ è richiesta essere differenziabile fino all'ordine k su qualche intervallo $I \ni t$. la condizione di *normalità* è la richiesta che la derivata di ordine più alto compaia, *da sola*, a primo membro. Trattando la dinamica newtoniana, come abbiamo visto nel capitolo precedente, si arriva a scrivere un sistema di equazioni differenziali del secondo ordine in forma normale, quando si applica il secondo principio della dinamica ad un sistema di punti materiali (tenendo conto dei vari casi visti nella sezione 2.3 incluso il caso di equazioni pure di movimento). La legge oraria dei punti materiali costituenti il sistema è una soluzione del sistema di equazioni differenziali. I sistemi di equazioni differenziali, per descrivere l'evoluzione di alcuni sistemi, nascono in vari contesti della fisica, ma anche di altre discipline come la biologia, la chimica, la biomatematica (dinamica delle popolazioni). In generale si tratta di sistemi di equazioni differenziali di ordine differente dal secondo. Il secondo ordine è tipico della fisica meccanica e, più in generale, della fisica.
Studiare il caso $k = 1$ non è restrittivo in quanto ci si può sempre ridurre a tale situazione partendo da $k \in \mathbb{N}$ con $k > 1$. Si consideri, a tal fine, il sistema (3.2) con $k > 1$. Definiamo le nuove variabili

$$\mathbf{y}_r := \frac{d^r\mathbf{x}}{dt^r} , \quad \text{per } r = 1, 2, \ldots, k-1.$$

Con queste variabili, il sistema (3.2) di ordine k è equivalente al sistema del prim'ordine, nella nuova variabile $\mathbf{X} = (\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1}) \in \mathbb{R}^{kn}$,

$$\frac{d\mathbf{X}}{dt} = \mathbf{g}\left(t, \mathbf{X}(t)\right) , \tag{3.3}$$

dove

$$\mathbf{g} : (t, (\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1})) \mapsto (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{f}\left(t, \mathbf{x}(t), \mathbf{y}_1(t), \ldots, \mathbf{y}_{k-1}(t)\right)) .$$

In questo modo, il sistema (3.3) si scrive esplicitamente:

$$\begin{aligned}
\frac{d\mathbf{y}_{k-1}}{dt} &= \mathbf{f}\left(t, \mathbf{x}(t), \mathbf{y}_1(t), \ldots, \mathbf{y}_{k-1}(t)\right) , \\
\frac{d\mathbf{y}_{k-2}}{dt} &= \mathbf{y}_{k-1} , \\
\cdots &= \cdots \\
\frac{d\mathbf{y}_1}{dt} &= \mathbf{y}_2 , \\
\frac{d\mathbf{x}}{dt} &= \mathbf{y}_1 .
\end{aligned}$$

Risulta allora evidente che ogni soluzione $\mathbf{X} = \mathbf{X}(t)$ di (3.3), con $t \in I$, definisce una soluzione di (3.2) $\mathbf{x} = \mathbf{x}(t)$, con $t \in I$, estraendo la prima componente della funzione $\mathbf{X}$; viceversa, ogni soluzione di (3.2) $\mathbf{x} = \mathbf{x}(t)$, con $t \in I$, definisce una soluzione di (3.3) data da $\mathbf{X} = (\mathbf{x}(t), d\mathbf{x}/dt, \ldots, d^{k-1}\mathbf{x}/dt^{k-1})$ con $t \in I$.
Riferendoci alle soluzioni $\mathbf{x} = \mathbf{x}(t)$ e $\mathbf{X} = \mathbf{X}(t)$, con $t \in I$ discusse nell'osservazione precedente, notiamo ancora che $\mathbf{x} = \mathbf{x}(t)$ soddisfa le $k-1$ condizioni, per $t_0 \in I$,

$$\mathbf{x}(t_0) = \mathbf{x}_0 \ , \frac{d\mathbf{x}}{dt}(t_0) = \mathbf{x}_0^{(1)} \ , \ldots , \ \frac{d\mathbf{x}^{k-1}}{dt^{k-1}}(t_0) = \mathbf{x}_0^{(k-1)} \ .$$

se e solo se $\mathbf{X} = \mathbf{X}(t)$ soddisfa la condizione

$$\mathbf{X}(t_0) = (\mathbf{x}_0, \mathbf{x}_0^{(1)} \ldots, \mathbf{x}_0^{(k-1)}) \ .$$

3.1.2 Problema di Cauchy.

Passiamo ad introdurre la seconda definizione fondamentale, quella di *problema di Cauchy*, che introduce la nozione di *condizioni iniziali*.
Se $\mathbf{f} : \Omega \to \mathbb{K}^n$, $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ aperto non vuoto, il **problema di Cauchy** per un sistema di equazioni differenziali del prim'ordine in forma normale:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}\left(t, \mathbf{x}(t)\right) \ , \tag{3.4}$$

consiste nel determinare le soluzioni del sistema (3.4) che soddisfano ulteriormente un'assegnata **condizione iniziale**

$$\mathbf{x}(t_0) = \mathbf{x}_0 \ ,$$

dove $(t_0, \mathbf{x}_0) \in \Omega$.
In altre parole si cercano le funzioni differenziabili della forma $\mathbf{x} : I \to \mathbb{K}^n$, con $I \subset \mathbb{R}$ intervallo aperto che contiene t_0, che risolvono (3.4) su I ed ulteriormente verificano $\mathbf{x}(t_0) = \mathbf{x}_0$.

Osservazioni 3.1. Possiamo formulare analogamente il problema di Cauchy per un sistema di equazioni differenziali di ordine $k > 1$.

$$\frac{d^k\mathbf{x}}{dt^k} = \mathbf{f}\left(t, \mathbf{x}(t), \frac{d\mathbf{x}}{dt}, \ldots, \frac{d^{k-1}\mathbf{x}}{dt^{k-1}}\right) \ ,$$

cercando le soluzioni del sistema che soddisfino $k-1$ **condizioni iniziali**:

$$\mathbf{x}(t_0) = \mathbf{x}_0 \ , \frac{d\mathbf{x}}{dt}(t_0) = \mathbf{x}_0^{(1)} \ , \ldots , \ \frac{d\mathbf{x}^{k-1}}{dt^{k-1}}(t_0) = \mathbf{x}_0^{(k-1)} \ .$$

In base a quanto spiegato nella sezione 3.1.1, questo problema di Cauchy è completamente equivalente ad un problema di Cauchy per un'equazione differenziale del prim'ordine: c'è una corrispondenza biunivoca tra le soluzioni del problema di Cauchy di ordine k e quello di ordine 1 associato. Per questo motivo, nel seguito, studieremo i problemi di esistenza ed unicità per sistemi di Cauchy del prim'ordine unicamente.

3.1.3 Integrali primi.

Una nozione utile in varie applicazioni, specialmente alla fisica, della teoria delle equazioni differenziali è la nozione di *integrale primo* che ora definiremo. Non ci occuperemo più di tale nozione fino al capitolo 4.

Definizione 3.1. **(Integrale primo.)** Se $\mathbf{f} : \Omega \to \mathbb{K}^n$, $\Omega \subset \mathbb{R} \times \mathbb{K}^n \times \cdots \times \mathbb{K}^n$ aperto non vuoto, consideriamo un sistema di equazioni differenziali (3.2). Una funzione $F : \Omega \to \mathbb{K}$ è detta **integrale primo** del sistema, se per ogni soluzione del sistema (3.2) $\mathbf{x} : I \to \mathbb{K}^n$, con $I \subset \mathbb{R}$ vale:

$$F\left(t, \mathbf{x}(t), \frac{d\mathbf{x}}{dt}, \ldots, \frac{d^{k-1}\mathbf{x}}{dt^{k-1}}\right) = c$$

per qualche costante $c \in \mathbb{K}$ dipendente, in generale, dalla soluzione considerata. $\diamondsuit$

3.2 Alcune nozioni e risultati preparatori per il teoremi di esistenza e unicità .

Per provare un teorema di esistenza ed unicità per il problema di Cauchy dato dalla (3.4) con condizione iniziale $\mathbf{x}(t_0) = \mathbf{x}_0$, quando la funzione $\mathbf{f}$ è sufficientemente regolare (localmente lipschitziana nella variabile $\mathbf{x}$), introdurremo lo spazio di Banach delle funzioni continue su un compatto K a valori in $\mathbb{K}^n$, $C^0(K; \mathbb{K}^n)$, e dimostreremo il teorema del punto fisso per le mappe di contrazione in spazi metrici. Da questo teorema, dopo aver trasformato il problema di Cauchy in un equazione integrale coinvolgente una certa mappa di contrazione su un sottoinsieme chiuso di $C^0(K; \mathbb{K}^n)$, proveremo il teorema di esistenza ed unicità per il problema di Cauchy, nelle forme locali e globali.

3.2.1 Lo spazio di Banach $C^0(K; \mathbb{K}^n)$.

Particolari tipi di spazi metrici sono gli *spazi normati* che hanno una struttura molto più ricca degli spazi metrici.
Ricordiamo che uno spazio vettoriale V sul campo $\mathbb{K}$ (che può essere $\mathbb{C}$ o $\mathbb{R}$ indifferentemente come detto all'inizio), si dice **spazio normato** se è dotato di un'applicazione, detta **norma**, $|| \cdot || : V \to \mathbb{R}$ che gode delle seguenti proprietà :

(N1) **positività stretta**. $||\mathbf{v}|| \geq 0$ con $\mathbf{v} = 0$ se $||\mathbf{v}|| = 0$, per ogni $v \in V$,
(N2) **omogeneità** . $||\lambda \mathbf{v}|| = |\lambda| ||\mathbf{v}||$ per ogni $\lambda \in \mathbb{K}$ e $\mathbf{v} \in V$,
(N3) **disuguaglianza triangolare**. $||\mathbf{v} + \mathbf{u}|| \leq ||\mathbf{v}|| + ||\mathbf{u}||$ per ogni $\mathbf{u}, \mathbf{v}$ in V.

Osservazioni 3.2.
(1) Si noti che (N2) implica che $||\mathbf{0}|| = 0$. Quindi ogni norma soddisfa: $||\mathbf{v}|| = 0$ se e solo se $\mathbf{v} = \mathbf{0}$. Se (N1) vale in forma debole (semipositività): $||\mathbf{v}|| \geq 0$ con $\mathbf{v} = \mathbf{0}$, e valgono (N2) e (N3), $|| \cdot ||$ è detta **seminorma** su V.

(2) Ogni spazio normato V con norma $||\cdot||$ ha una naturale struttura di spazio metrico, ed una conseguente topologia metrica, se si definisce la distanza

$$d(\mathbf{u},\mathbf{v}) := ||\mathbf{u}-\mathbf{v}|| \,, \quad \forall \mathbf{u},\mathbf{v} \in V \,.$$

(3) Come gli spazi metrici, gli spazi normati sono spazi di Hausdorff. Pertanto, negli spazi normati vale la proprietà di unicità del limite.

Esempi 3.1.
1. L'esempio più semplice di spazio di spazio normato è quello costruito sullo spazi vettoriale $\mathbb{C}^n$ o $\mathbb{R}^n$ con norma $||\cdot||$ data dalla norma euclidea standard:

$$||(c_1,\cdots,c_n)|| := \sqrt{|c_1|^2+\cdots|c_n|^2}$$

dove $|\cdot|$ è il valore assoluto su $\mathbb{C}$ o $\mathbb{R}$.
2. Se K è uno spazio topologico compatto (per esempio un sottoinsieme di $\mathbb{R}^n$ o $\mathbb{C}^n$ chiuso e limitato) $C^0(K;\mathbb{K})$ indica l'insieme delle funzioni *continue* definite su K a valori in $\mathbb{K}^n$, che al solito può essere $\mathbb{R}^n$ o $\mathbb{C}^n$. Possiamo dotare tale insieme della struttura di spazio vettoriale sul campo $\mathbb{K}$, definendo il prodotto per scalare e la somma di funzioni tramite

$$(\lambda f)(x) := \lambda f(x) \,, \quad \text{per ogni } f \in C^0(K;\mathbb{K}^n),\ \lambda \in \mathbb{K} \text{ e } x \in K$$

e

$$(f+g)(x) := f(x)+g(x) \,, \quad \text{per ogni } f,g \in C^0(K;\mathbb{K}^n) \text{ e } x \in K.$$

Lo spazio $C^0(K;\mathbb{K}^n)$ assume una struttura di spazio normato definendo la norma

$$||f||_\infty := \sup_{x\in K} ||f(x)|| \,.$$

Tale norma è ben definita in quanto, per ogni $f \in C^0(K,\mathbb{K}^n)$, la funzione $K \ni x \mapsto ||f(x)||$ è continua (essendo composizione di funzioni continue) ed è quindi limitata essendo definita su un insieme compatto, mentre le proprietà (N1), (N2) e (N3) per $||\cdot||_\infty$ valgono banalmente come conseguenza delle analoghe proprietà valide per $||\cdot||$ punto per punto, e di note proprietà dell'estremo superiore.

Ricordiamo che se X è uno spazio metrico con distanza d, una successione $\{x_n\}_{n\in\mathbb{N}} \subset X$ è detta **successione di Cauchy** se vale la **condizione di Cauchy**: *per ogni* $\epsilon > 0$ *esiste* $N_\epsilon \in \mathbb{N}$ *tale che* $d(x_n,x_m) < \epsilon$ *quando* $n,m > N_\epsilon$.
Dovrebbe essere ben noto che ogni successione $\{x_n\}_{n\in\mathbb{N}} \subset X$ che converge ad un limite $x \in X$ è automaticamente di Cauchy. Infatti, in tali ipotesi, per ogni $\eta > 0$ deve valere $d(x_n,x_m) \leq d(x_n,x)+d(x_m,x) < 2\eta$ se $n > M_\eta$ e $m > M_\eta$ per qualche $M_\eta \in \mathbb{N}$. Ma allora, per ogni fissato $\epsilon > 0$, definendo $N_\epsilon := M_\eta$ con $\eta = \epsilon/2$, si ha la condizione di Cauchy per la successione $\{x_n\}_{n\in\mathbb{N}}$.
Ricordiamo che uno spazio metrico è detto **completo** se vale anche il viceversa, cioé se ogni

successione di Cauchy converge a qualche elemento dello spazio.

Definizione 3.2. Uno spazio normato V è detto **spazio di Banach** se è completo come spazio metrico. $\diamondsuit$

Ricordiamo che gli spazi $\mathbb{R}$ e $\mathbb{C}$ sono completi [GiustiI] per cui sono esempi elementari di spazi di Banach. Da ciò segue facilmente che $\mathbb{K}^n$ dell'esempio 3.1.**1** è completo. La dimostrazione è molto semplice. Consideriamo una successione $\{\mathbf{u}_m\}_{m\in\mathbb{N}} \subset \mathbb{K}^n$ dove $\mathbf{u}_m = (v_{1m}, \ldots, v_{nm})$. Se $\{\mathbf{u}_m\}_{m\in\mathbb{N}} \subset \mathbb{K}^n$ è di Cauchy, sono di Cauchy le successioni $\{v_{im}\}_{m\in\mathbb{N}} \subset \mathbb{K}$, per ogni fissato $i = 1, 2, \ldots, n$, essendo $|v_{ip} - v_{iq}| \leq ||\mathbf{u}_p - \mathbf{u}_q||$. Se $v_i \in \mathbb{K}$ è il limite a cui tende $\{v_{im}\}_{m\in\mathbb{N}}$, è immediato verificare che $\mathbf{u}_n = (v_{1m}, \ldots, v_{nm}) \to (v_1, \ldots, v_n) =: \mathbf{u}$ per $n \to +\infty$ dato che

$$||\mathbf{u}_m - \mathbf{u}||^2 = |v_{1m} - v_1|^2 + \cdots + |v_{nm} - v_n|^2 \to 0\,, \quad \text{per } m \to +\infty.$$

Il seguente teorema sarà di grande utilità nella teoria delle equazioni differenziali.

Teorema 3.1. *Se K è uno spazio topologico compatto, lo spazio normato $C^0(K;\mathbb{K}^n)$ con norma $||\cdot||_\infty$ è uno spazio di Banach.* $\diamondsuit$

Dimostrazione. Sia $\{f_n\}_{n\in\mathbb{N}} \subset C^0(K;\mathbb{K}^n)$ una successione di Cauchy. Vogliamo dimostrare che esiste $f \in C^0(K)$ tale che $||f_n - f||_\infty \to 0$ per $n \to +\infty$. Dato che $\{f_n\}_{n\in\mathbb{N}}$ è di Cauchy, per ogni fissato $x \in K$, è di Cauchy anche la successione dei vettori di $\mathbb{K}^n$: $\{f_n(x)\}_{n\in\mathbb{N}} \subset \mathbb{K}^n$. Infatti:

$$||f_n(x) - f_m(x)|| \leq \sup_{z\in K} ||f_n(z) - f_m(z)|| = ||f_n - f_m||_\infty < \epsilon\,, \quad \text{se } n, m < N_\epsilon\,.$$

In questo modo, dato che $\mathbb{K}^n$ è completo, punto per punto viene definita la funzione:

$$f(x) := \lim_{n\to+\infty} f_n(x)\,.$$

Vogliamo ora provare che $f \in C^0(K;\mathbb{K}^n)$ e che $||f_n - f||_\infty \to 0$ per $n \to +\infty$.
Dato che la successione $\{f_n\}_{n\in\mathbb{N}}$ è di Cauchy, per ogni $\epsilon > 0$, esiste N_ϵ tale che, se $n, m > N_\epsilon$

$$||f_n(x) - f_m(x)|| < \epsilon\,, \quad \text{per } \textit{tutti} \text{ gli } x \in K.$$

D'altra parte, dalla definizione di f, per un *fissato* $x \in K$ e per ogni $\epsilon'_x > 0$, esisterà N_{x,ϵ'_x} tale che, se $m > N_{x,\epsilon'_x}$ vale $||f_m(x) - f(x)|| < \epsilon'_x$.
Usando entrambi i fatti abbiamo subito che, se $n > N_\epsilon$

$$||f_n(x) - f(x)|| \leq ||f_n(x) - f_m(x)|| + ||f_m(x) - f(x)|| < \epsilon + \epsilon'_x$$

dove abbiamo scelto $m > \max(N_\epsilon, N_{x,\epsilon'_x})$. In definitiva, se $n > N_\epsilon$, vale

$$||f_n(x) - f(x)|| < \epsilon + \epsilon'_x\,, \quad \text{per ogni } \epsilon'_x > 0.$$

Dato che $\epsilon'_x > 0$ è arbitrario, la disuguaglianza di sopra vale anche per $\epsilon'_x = 0$, eventualmente diventando un'uguaglianza. In questo modo si perde la dipendenza da x. In definitiva, per ogni $\epsilon > 0$ esiste $N_\epsilon \in \mathbb{N}$ tale che, se $n > N_\epsilon$

$$||f_n(x) - f(x)|| \leq \epsilon \,, \quad \text{per } \textit{tutti} \text{ gli } x \in K. \tag{3.5}$$

Quindi $\{f_n\}$ converge a f *uniformemente*. Per noti teoremi di analisi [GiustiI], f è allora continua essendo limite uniforme di funzioni continue. Dunque $f \in C^0(K; \mathbb{K}^n)$.
Dato che (3.5) vale per ogni $x \in K$, vale anche prendendo l'estremo superiore degli x in K: per ogni $\epsilon > 0$ esiste $N_\epsilon \in \mathbb{N}$ tale che $\sup_{x \in K} ||f_n(x) - f(x)|| < \epsilon$, se $n > N_\epsilon$. In altre parole

$$||f_n - f||_\infty \to 0 \,, \quad \text{se } n \to +\infty.$$

Questo conclude la dimostrazione. □

Per ultimo dimostriamo la seguente elementare proposizione.

Proposizione 3.1. *Se Y è un sottoinsieme non vuoto e chiuso di uno spazio metrico completo X con distanza d, Y è uno spazio metrico completo rispetto alla restrizione della distanza d a $Y \times Y$.* ♢

Dimostrazione. È ovvio che Y dotato della restrizione di d a $Y \times Y$ è uno spazio metrico. Se $\{y_n\}_{n \in \mathbb{N}} \subset Y$ è una successione di Cauchy, è successione di Cauchy anche nello spazio metrico X. Dato che X è completo, esiste $y \in X$ con $\lim_{n \to +\infty} y_n = y$. Dato che Y è chiuso, esso contiene i suoi punti di accumulazione, pertanto: $y \in Y$. Abbiamo provato che ogni successione di Cauchy in Y ammette un limite in Y. Questo conclude la dimostrazione. □

3.2.2 Teorema del punto fisso in spazi metrici completi.

In uno spazio metrico esistono particolari applicazioni che giocano un ruolo centrale nella dimostrazione del teorema di esistenza ed unicità per il problema di Cauchy. Si tratta delle cosiddette *mappe di contrazione*.

Definizione 3.3. Sia X uno spazio metrico con norma d. Una funzione, $G : X \to X$ è detta **mappa di contrazione**, se esiste un numero reale $\lambda \in [0, 1)$ per cui:

$$d(G(x), G(y)) \leq \lambda d(x, y) \,. \tag{3.6}$$

♢

Osservazioni 3.3.
(1) Notare che è esplicitamente *escluso* il valore $\lambda = 1$.
(2) La (3.6) implica immediatamente che: *ogni mappa di contrazione è continua* nella topologia metrica sullo spazio X. Infatti, dalla (3.6) segue che se $x \to y$ allora $G(x) \to G(y)$.

(3) La definizione vale anche per spazi normati usando la distanza associata alla norma. Si osservi che la funzione G in tal caso *non* è richiesta essere lineare.

Concludiamo con il *teorema del punto fisso* (di Banach-Caccioppoli) che ha molteplici conseguenze in matematica e nelle sue applicazioni (specialmente in fisica).

Teorema 3.2. (**del punto fisso**.) *Sia $G : X \to X$ una mappa di contrazione sullo spazio metrico completo X. Esiste ed, è unico, un elemento $z \in X$, detto* **punto fisso**, *tale che:*

$$G(z) = z\,. \tag{3.7}$$

$\diamondsuit$

Dimostrazione.
Esistenza del punto fisso. Si consideri la successione, dove $x_0 \in X$ è arbitrario, definita per induzione da $x_{n+1} = G(x_n)$. Vogliamo provare che questa successione è di Cauchy. Dato che X è completo esisterà il limite $x \in X$ della successione. Proveremo infine che tale limite è un punto fisso.
Assumiamo, senza perdere generalità che $m \geq n$. Se $m = n$ banalmente $d(x_m, x_n) = 0$, se $m > n$ possiamo usare ripetutamente la disuguaglianza triangolare ottenendo:

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n)\,. \tag{3.8}$$

Consideriamo il generico termine a secondo membro: $d(x_{p+1}, x_p)$. Vale:

$$d(x_{p+1}, x_p) = d(G(x_p), G(x_{p-1})) \leq \lambda d(x_p, x_{p-1}) = \lambda d(G(x_{p-1}), G(x_{p-2})) \leq \lambda^2 d(x_{p-1}, x_{p-2})$$
$$\leq \cdots \leq \lambda^p d(x_1, x_0)\,.$$

Quindi, per ogni $p = 1, 2, \ldots$ vale la disuguaglianza $d(x_{p+1}, x_p) \leq \lambda^p d(x_1, x_0)$. Inserendo questa nel secondo membro di (3.8) otteniamo la stima:

$$d(x_m, x_n) \leq d(x_1, x_0) \sum_{p=n}^{m-1} \lambda^p = d(x_1, x_0)\lambda^n \sum_{p=0}^{m-n-1} \lambda^p \leq \lambda^n d(x_1, x_0) \sum_{p=0}^{+\infty} \lambda^p \leq d(x_1, x_0) \frac{\lambda^n}{1-\lambda}$$

dove abbiamo usato il fatto che, se $0 \leq \lambda < 1$, $\sum_{p=0}^{+\infty} \lambda^p = (1-\lambda)^{-1}$. Concludiamo che:

$$d(x_m, x_n) \leq d(x_1, x_0) \frac{\lambda^n}{1-\lambda}\,. \tag{3.9}$$

Nelle nostre ipotesi $|\lambda| < 1$ per cui $d(x_1, x_0)\lambda^n/(1-\lambda) \to 0$ se $n \to +\infty$. Quindi $d(x_m, x_n)$ può essere reso piccolo a piacere scegliendo il più piccolo tra m ed n sufficientemente grande. Questo risultato implica immediatamente che la successione $\{x_n\}_{n\in\mathbb{N}}$ sia di Cauchy. Dato che

X è completo, sarà $\lim_{n\to+\infty} x_n = x \in X$ per qualche x. Dato che G è continua essendo una mappa di contrazione:

$$G(x) = G\left(\lim_{n\to+\infty} x_n\right) = \lim_{n\to+\infty} G(x_n) = \lim_{n\to+\infty} x_{n+1} = x\ ,$$

per cui $G(x) = x$ come volevamo provare.
Unicità del punto fisso. Supponiamo che x e x' soddisfino entrambi $G(x) = x$ e $G(x') = x'$. Vale allora

$$d(x, x') = d(G(x), G(x')) \leq \lambda d(x, x')\ .$$

Se fosse $d(x, x') \neq 0$, dividendo il primo e l'ultimo membro per $d(x, x')$, otterremmo la disuguaglianza $1 \leq \lambda$ che è impossibile per ipotesi. Quindi deve essere $d(x, x') = 0$ da cui $x = x'$ per la stretta positività della distanza. □

3.2.3 Funzioni lipschitziane.

Come ultimo ingrediente, per applicare la tecnologia matematica introdotta al teorema di esistenza ed unicità del problema di Cauchy, introduciamo la nozione di *funzione (localmente) lipschitziana.* Questa nozione serve a specificare il grado di regolarità della funzione che compare a secondo membro nell'equazione differenziale che definisce il problema di Cauchy.

Definizione 3.4. Siano $p \geq 0$ e $n, m > 0$ naturali fissati e $\Omega \subset \mathbb{K}^p \times \mathbb{K}^n$ aperto non vuoto. $\mathbf{F} : \Omega \to \mathbb{K}^m$ è detta **localmente lipschitziana** (**nella variabile** $\mathbf{x} \in \mathbb{K}^n$ quando $p > 0$), se per ogni $p \in \Omega$ esiste una costante $L_p \geq 0$ tale che, in un aperto $O_p \ni p$, valga:

$$||\mathbf{F}(\mathbf{z}, \mathbf{x}) - \mathbf{F}(\mathbf{z}, \mathbf{x}')|| \leq L_p ||\mathbf{x} - \mathbf{x}'||\ , \quad \text{per ogni coppia } (\mathbf{z}, \mathbf{x}), (\mathbf{z}, \mathbf{x}') \in O_p. \tag{3.10}$$

Se $O_p = \Omega$, la funzione $\mathbf{F}$ è detta **lipschitziana** (**nella variabile** $\mathbf{x} \in \mathbb{K}^n$ quando $p > 0$).
♢

Osservazioni 3.4. Nella (3.10) le coppie $(\mathbf{z}, \mathbf{x}), (\mathbf{z}, \mathbf{x}')$ hanno *lo stesso* primo elemento $\mathbf{z}$.

Esempi 3.2.
1. Usiamo le solite notazioni $\mathbf{x} = (x^1, \ldots, x^n)$, $\mathbf{z} = (z^1, \ldots, z^p)$ e, per una funzione $\mathbf{F} : \Omega \to \mathbb{R}^m$ con $\Omega = A \times B$ con $A \subset \mathbb{R}^p$ e $B \subset \mathbb{R}^n$ aperti non vuoti, usiamo la notazione $\mathbf{F}(\mathbf{z}, \mathbf{x}) = (F^1(\mathbf{z}, \mathbf{x}), \ldots, F^m(\mathbf{z}, \mathbf{x}))$. La classe delle funzioni localmente lipschitziane è grandissima. Per esempio, vale la seguente proposizione.

Proposizione 3.2. *$\mathbf{F} : \Omega \to \mathbb{R}^m$ con Ω suddetto è sicuramente localmente lipschitziana nella variabile $\mathbf{x}$ se, per ogni $\mathbf{z} \in A$, tutte le funzioni $B \ni \mathbf{x} \mapsto F^k(\mathbf{z}, \mathbf{x})$ sono differenziabili e le*

derivate parziali, al variare di $(\mathbf{z}, \mathbf{x})$*, definiscono funzioni continue*[1] *su* Ω.

Dimostrazione. Consideriamo $p = (\mathbf{z}_0, \mathbf{x}_0) \in \Omega$ e siano $B_{\mathbf{x}_0} \subset \mathbb{R}^n$ e $B'_{\mathbf{z}_0} \subset \mathbb{R}^p$ palle aperte di raggio finito (non nullo) centrate rispettivamente in $\mathbf{x}_0$ e $\mathbf{z}_z$ tali che

$$\overline{B_{\mathbf{x}_0}} \times \overline{B'_{\mathbf{z}_0}} \subset \Omega \,.$$

Se $\mathbf{x}_1, \mathbf{x}_2 \in \overline{B_{\mathbf{x}_0}}$, il segmento $\mathbf{x}(t) = \mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)$ è tutto contenuto in $\overline{B_{\mathbf{x}_0}}$ per $t \in [0,1]$, inoltre $\mathbf{x}(0) = \mathbf{x}_1$ e $\mathbf{x}(1) = \mathbf{x}_2$.
Se $\mathbf{z} \in B'_{\mathbf{z}_0}$ e per ogni fissato $k = 1, 2, \ldots, m$, vale di conseguenza

$$F^k(\mathbf{z}, \mathbf{x}_2) - F^k(\mathbf{z}, \mathbf{x}_1) = F^k(\mathbf{z}, \mathbf{x}(1)) - F^k(\mathbf{z}, \mathbf{x}(0)) = \xi \sum_{j=1}^{n} (x_2^j - x_1^j) \left.\frac{\partial F^k}{\partial x^j}\right|_{(\mathbf{z}, \mathbf{x}(\xi))} ,$$

dove abbiamo applicato il teorema di Lagrange alla funzione $[0,1] \ni t \mapsto F^k(\mathbf{z}, \mathbf{x}(t))$, e $\xi \in [0,1]$ (che dipenderà da k e da $\mathbf{z}$). Di conseguenza, applicando la disuguaglianza di Schwartz e tenendo conto che $0 \leq \xi \leq 1$:

$$\left| F^k(\mathbf{z}, \mathbf{x}_2) - F^k(\mathbf{z}, \mathbf{x}_1) \right| \leq \xi \sqrt{\sum_{j=1}^{n} |x_2^j - x_1^j|^2} \sqrt{\sum_{i=1}^{n} \left| \left.\frac{\partial F^k}{\partial x^i}\right|_{(\mathbf{z}, \mathbf{x}(\xi))} \right|^2} \leq ||\mathbf{x}_2 - \mathbf{x}_1|| \sqrt{\sum_{i=1}^{n} \left| \left.\frac{\partial F^k}{\partial x^i}\right|_{(\mathbf{z}, \mathbf{x}(\xi))} \right|^2} .$$

Dato che le derivate nell'ultimo membro sono continue sul compatto $\overline{B_{\mathbf{x}_0}} \times \overline{B'_{\mathbf{z}_0}}$ esisterà $M_k \in [0, +\infty)$ con, su $\overline{B_{\mathbf{x}_0}} \times \overline{B'_{\mathbf{z}_0}}$,

$$\sqrt{\sum_{i=1}^{n} \left| \left.\frac{\partial F^k}{\partial x^i}\right|_{(\mathbf{z}, \mathbf{x})} \right|^2} \leq M_k \,.$$

In definitiva, sull'intorno aperto di p, $B_{\mathbf{x}_0} \times B'_{\mathbf{z}_0}$, vale la stima, per ogni k

$$\left| F^k(\mathbf{z}, \mathbf{x}_2) - F^k(\mathbf{z}, \mathbf{x}_1) \right| \leq M_k ||\mathbf{x}_2 - \mathbf{x}_1|| \,.$$

Se infine $M := \sqrt{\sum_{k=1}^{m} M_k^2}$, si ha immediatamente che sull'intorno aperto di p, $B_{\mathbf{x}_0} \times B'_{\mathbf{z}_0}$:

$$||\mathbf{F}(\mathbf{z}, \mathbf{x}_2) - \mathbf{F}(\mathbf{z}, \mathbf{x}_1)|| \leq M ||\mathbf{x}_2 - \mathbf{x}_1|| \,.$$

Questa disuguaglianza prova quanto volevamo dimostrare. □
2. Dall'esempio precedente segue, in particolare, che ogni $\mathbf{F} \in C^1(\Omega, \mathbb{R}^m)$ è localmente lipschitziana.

[1] Tenendo conto che condizione sufficiente affinché una funzione di più variabili sia differenziabile su un aperto è che esistono e siano continue le derivate parziali su tale aperto, la richiesta completa equivale alla condizione che le funzioni

$$\Omega \ni (\mathbf{z}, \mathbf{x}) \mapsto \left.\frac{\partial F^k}{\partial x^j}\right|_{(\mathbf{z}, \mathbf{x})} , \quad \text{per } k = 1, 2, \ldots, m \text{ e } j = 1, \ldots, n,$$

esistano e siano funzioni continue su Ω.

3. Usando le solite convenzioni $\mathbf{x} = (x^1, \ldots, x^n)$, $\mathbf{z} = (z^1, \ldots, z^p)$ e considerando sempre una funzione $\mathbf{F} : \Omega \to \mathbb{R}^m$ con $\Omega \subset \mathbb{R}^p \times \mathbb{R}^n$, si possono estendere le osservazioni appena fatte al caso di funzione lipschitziana e non solo *localmente* tale. Segue immediatamente dal discorso appena fatto che $\mathbf{F} \in C^1(\Omega, \mathbb{R}^m)$ è lipschitziana nella variabile $\mathbf{x} \in \mathbb{R}^n$ se ristretta ad ogni aperto limitato della forma $A \times C$ con $A \subset \mathbb{R}^p$, $C \subset \mathbb{R}^n$, $\overline{A \times C} \subset \Omega$ e dove C è convesso. In tal caso, infatti, le derivate prime delle varie componenti sono sicuramente limitate sul compatto $\overline{A \times C}$, essendo funzioni continue; inoltre l'ipotesi di convessità permette di scegliere i punti $\mathbf{x}_1$ e $\mathbf{x}_2$ arbitrariamente in C, essendo sicuri che il segmento che unisce tali punti sia tutto contenuto in C. L'argomento usato sopra basato sul teorema di Lagrange prova la lipschitzianità nella variabile $\mathbf{x} \in C$.
Se Ω è convesso (per es. $\Omega = \mathbb{R}$) e tutte le derivate prime di $\mathbf{F} \in C^1(\Omega, \mathbb{R}^n)$ sono limitate su Ω, allora $\mathbf{F}$ è lipschitziana su Ω.
4. La funzione continua $f(x) = |x|$, da $\mathbb{R}$ in $\mathbb{R}$, è localmente lipschitziana nella variabile x, anche se *non è differenziabile per* $x = 0$. Dimostriamolo. Per $x \neq 0$, c'è un intorno aperto di tale punto in cui la funzione è C^∞, per cui vale quanto detto nell'esempio precedente. Consideriamo infine un intorno di $x = 0$, $I := (-\epsilon, \epsilon)$ (con $\epsilon > 0$). Se $x_1, x_2 \in I$ sia, senza perdere generalità $|x_2| \geq |x_1|$, allora $|x_2| = |x_2 - x_1 + x_1| \leq |x_2 - x_1| + |x_1|$ da cui $|x_2| - |x_1| \leq |x_2 - x_1|$. Ossia, nelle ipotesi fatte $|\,|x_2| - |x_1|\,| \leq |x_2 - x_1|$, che significa

$$|f(x_2) - f(x_1)| \leq |x_2 - x_1|\,, \quad \text{per ogni coppia } x_1, x_2 \in I\,.$$

5. La funzione continua da $\mathbb{R}$ in $\mathbb{R}$:

$$f(x) := \begin{cases} \sqrt{x} & \text{se } x \geq 0\,, \\ 0 & \text{se } x < 0\,, \end{cases} \tag{3.11}$$

non è invece localmente lipschitziana (nella variabile x ovviamente). Se escludiamo il punto $x = 0$ dal dominio, ovviamente la funzione è localmente lipschitziana per l'esempio 3, essendo in tale dominio C^∞. Il problema è quindi il comportamento della funzione attorno a $x = 0$. Mostriamo che per ogni fissato intorno aperto I di 0 non esiste alcuna costante $c \in \mathbb{R}$ tale che, se $x_1, x_2 \in I$, valga $|f(x_2) - f(x_1)| \leq c|x_2 - x_1|$. Possiamo assumere $I := (-\epsilon, \epsilon)$, con $\epsilon > 0$, in quanto, se c non esiste per gli intervalli del tipo detto, non può esistere nemmeno per gli intorni[2]. Procediamo *per assurdo*. Assumiamo che esista una costante $c \in \mathbb{R}$ che verifichi $|f(x_2) - f(x_1)| \leq c|x_2 - x_1|$ per ogni coppia $x_1, x_2 \in (-\epsilon, \epsilon)$. Scegliamo $x_2 := z \in (0, \epsilon)$ e $x_1 = -z$. La disuguaglianza $|f(x_2) - f(x_1)| \leq c|x_2 - x_1|$ si scrive allora $\sqrt{z} \leq 2cz$. Di conseguenza deve valere $2c \geq 1/\sqrt{z}$ per ogni $z > 0$ arbitrariamente piccolo. Questo è impossibile comunque fissiamo il valore della costante c.

3.3 Teoremi di esistenza ed unicità per il problema di Cauchy.

Siamo ora in grado di enunciare e provare i teoremi di esistenza ed unicità del problema di Cauchy nella forma locale e globale.

[2] ogni intorno aperto di 0 include un intervallo del tipo detto!

3.3.1 Teorema di esistenza ed unicità locale per il problema di Cauchy.

Cominciamo a dimostrare che in un intorno sufficientemente piccolo delle condizioni inziali, il problema di Cauchy ammette una sola soluzione se la funzione $\mathbf{f}$ che appare nell'equazione differenziale è continua, e localmente lipschitziana nella variabile $\mathbf{x}$.

Teorema 3.3. (**Esistenza ed unicità locale**) *Sia* $\mathbf{f} : \Omega \to \mathbb{K}^n$ *una funzione continua e localmente lipschitziana nella variabile* $\mathbf{x} \in \mathbb{K}^n$ *sull'aperto* $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ *e si consideri il problema di Cauchy per* $(t_0, \mathbf{x}_0) \in \Omega$*:*

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(t, \mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.12}$$

C'è un intervallo aperto $I \ni t_0$ *in cui esiste, ed è unica, la soluzione di (3.12),* $\mathbf{x} : I \to \mathbb{K}^n$*. Inoltre tale funzione è necessariamente in* $C^1(I)$*.* $\diamondsuit$

Dimostrazione. Prima di tutto notiamo che ogni soluzione $\mathbf{x} = \mathbf{x}(t)$ di (3.12) è necessariamente di classe C^1. Infatti è sicuramente continua in quanto differenziabile, ma allora, direttamente da $\frac{d\mathbf{x}}{dt} = \mathbf{f}\left(t, \mathbf{x}(t)\right)$ concludiamo che $\frac{d\mathbf{x}}{dt}$ deve essere continua essendo il secondo membro dell'equazione una funzione continua di t in quanto composizione di funzioni continue.
Passiamo all'esistenza ed all'unicità . Supponiamo che $\mathbf{x} : I \to \mathbb{K}^n$ sia differenziabile e soddisfi (3.12). Tenendo conto del secondo teorema fondamentale del calcolo ed integrando (3.12) (la derivata di $\mathbf{x}(t)$ è continua), $\mathbf{x} : I \to \mathbb{K}^n$ deve soddisfare,

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}\left(\tau, \mathbf{x}(\tau)\right) d\tau\,, \quad \text{per ogni } t \in I. \tag{3.13}$$

Viceversa, se $\mathbf{x} : I \to \mathbb{K}^n$ è continua e soddisfa (3.13) allora, per il primo teorema fondamentale del calcolo ($\mathbf{f}$ è continua), $\mathbf{x} = \mathbf{x}(t)$ è differenziabile e soddisfa (3.12).
Concludiamo che funzioni continue $\mathbf{x} = \mathbf{x}(t)$, definite su un intervallo aperto $I \ni t_0$ e che risolvono l'equazione integrale (3.13), sono tutte e sole le soluzioni di (3.12) definite sullo stesso intervallo I. Invece di risolvere (3.12) risolviamo il problema integrale equivalente (3.13).

Dimostrazione dell'esistenza. Fissiamo, una volta per tutte, un insieme aperto $Q \ni (t_0, \mathbf{x}_0)$ a chiusura compatta e con $\overline{Q} \subset \Omega$. Prendendo $\overline{Q}$ sufficientemente piccolo possiamo sfruttare anche la locale lipschitzianità di $\mathbf{f}$ nella variabile $\mathbf{x}$. Nel seguito:

(i) $0 \leq M := \max\{||\mathbf{f}(t, \mathbf{x})|| \mid (t, \mathbf{x}) \in \overline{Q}\}$;

(ii) $L \geq 0$ è la costante per cui $||\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{x}')|| \leq L||\mathbf{x} - \mathbf{x}'||$ se $(t, \mathbf{x}), (t, \mathbf{x}') \in \overline{Q}$;

(iii) $B_\epsilon(\mathbf{x}_0) := \{\mathbf{x} \in \mathbb{K}^n \mid ||\mathbf{x} - \mathbf{x}_0|| \leq \epsilon\}$ per $\epsilon > 0$.

Consideriamo un intervallo chiuso $J_\delta = [t_0 - \delta, t_0 + \delta]$ con $\delta > 0$ e consideriamo lo spazio di Banach $C^0(J; \mathbb{K}^n)$ delle funzioni continue $\mathbf{X} : J_\delta \to \mathbb{K}^n$. In questo spazio definiamo la funzione G, che associa ad ogni funzione $\mathbf{X}$ la nuova funzione $G(\mathbf{X})$ definita da

$$G(\mathbf{X})(t) := \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}\left(\tau, \mathbf{X}(\tau)\right) d\tau\,, \quad \text{per ogni } t \in J_\delta\,.$$

Si osservi che $G(\mathbf{X}) \in C^0(J_\delta; \mathbb{K}^n)$ se $\mathbf{X} \in C^0(J_\delta; \mathbb{K}^n)$ per le proprietà di continuità della funzione integrale (come funzione dell'estremo superiore quando l'integrando è continuo). Mostriamo che G è una mappa di contrazione su un sottoinsieme chiuso di $C^0(J_\delta; \mathbb{K}^n)$ dato dalla palla chiusa[3] di raggio $\epsilon > 0$:

$$\mathcal{B}_\epsilon^{(\delta)} := \{\mathbf{X} \in C^0(J_\delta; \mathbb{K}^n) \mid \mathbf{X}(t) \in B_\epsilon(\mathbf{x}_0)\,, \forall t \in J_\delta.\}\,.$$

se $0 < \delta < \min\{\epsilon/M, 1/L\}$ e $\delta, \epsilon > 0$ sono piccoli a sufficienza da soddisfare $J_\delta \times B_\epsilon(\mathbf{x}_0) \subset Q$. *In tutto il seguito assumeremo quindi sempre di aver scelto $\epsilon > 0$ e $\delta > 0$ in modo che valga $J_\delta \times B_\epsilon(\mathbf{x}_0) \subset Q$.*
Se $\mathbf{X} \in \mathcal{B}_\epsilon^{(\delta)}$ vale

$$||G(\mathbf{X})(t) - \mathbf{x}_0|| \leq \left|\left|\int_{t_0}^t \mathbf{f}\left(\tau, \mathbf{X}(\tau)\right) d\tau\right|\right| \leq \int_{t_0}^t ||\mathbf{f}\left(\tau, \mathbf{X}(\tau)\right)||\, d\tau \leq \int_{t_0}^t M d\tau \leq \delta M\,.$$

Quindi $G(B_\epsilon^{(\delta)}) \subset \mathcal{B}_\epsilon^{(\delta)}$ se $\delta > 0$ è sufficientemente piccolo da soddisfare $\delta < \epsilon/M$ (rispettando il vincolo $J_\delta \times B_\epsilon \subset Q$).
Se $\mathbf{X}, \mathbf{X}' \in \mathcal{B}_\epsilon^{(\delta)}$ si ha, per ogni $t \in J_\delta$:

$$G(\mathbf{X})(t) - G(\mathbf{X}')(t) = \int_{t_0}^t \left[\mathbf{f}\left(\tau, \mathbf{X}(\tau)\right) - \mathbf{f}\left(\tau, \mathbf{X}'(\tau)\right)\right] d\tau\,,$$

$$\left|\left|G(\mathbf{X})(t) - G(\mathbf{X}')(t)\right|\right| \leq \left|\left|\int_{t_0}^t \left[\mathbf{f}\left(\tau, \mathbf{X}(\tau)\right) - \mathbf{f}\left(\tau, \mathbf{X}'(\tau)\right)\right] d\tau\right|\right| \leq \int_{t_0}^t \left|\left|\mathbf{f}\left(\tau, \mathbf{X}(\tau)\right) - \mathbf{f}\left(\tau, \mathbf{X}'(\tau)\right)\right|\right| d\tau.$$

Dato che nelle nostre ipotesi vale la maggiorazione lipschitziana

$$||\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{x}')|| < L||\mathbf{x} - \mathbf{x}'||\,,$$

si ha ancora:

$$\left|\left|G(\mathbf{X})(t) - G(\mathbf{X}')(t)\right|\right| \leq L \int_{t_0}^t \left|\left|\mathbf{X}(\tau) - \mathbf{X}'(\tau)\right|\right| d\tau \leq \delta L ||\mathbf{X} - \mathbf{X}'||_\infty\,.$$

Infine, prendendo l'estremo superiore anche del primo membro si ha:

$$\left|\left|G(\mathbf{X}) - G(\mathbf{X}')\right|\right|_\infty \leq \delta L ||\mathbf{X} - \mathbf{X}'||_\infty\,.$$

Concludiamo che, se vale (in aggiunta a quanto richiesto sopra) $\delta < 1/L$, $G : \mathcal{B}_\epsilon^{(\delta)} \to \mathcal{B}_\epsilon^{(\delta)}$ è una mappa di contrazione sull'insieme chiuso $\mathcal{B}_\epsilon^{(\delta)}$. Tale insieme è chiuso in uno spazio metrico completo per cui vale la proposizione 3.1 e $\mathcal{B}_\epsilon^{(\delta)}$ è , a sua volta, spazio metrico completo. Per il teorema 3.2 esisterà un punto fisso per G, cioé una funzione continua $\mathbf{x} = \mathbf{x}(t) \in \mathbb{K}^n$ con $t \in J_\delta$,

[3] Si osservi che si può anche scrivere $\mathcal{B}_\epsilon^{(\delta)} = \{\mathbf{X} \in C^0(J_\delta; \mathbb{K}^n) \mid ||\mathbf{X} - \mathbf{X}_0||_\infty \leq \epsilon\}$, dove $\mathbf{X}_0$ indica qui la *funzione* che vale costantemente $\mathbf{x}_0$ su J_δ. Pertanto $\mathcal{B}_\epsilon^{(\delta)}$ è la palla *chiusa* di raggio ϵ centrata in $\mathbf{X}_0$ dello spazio di Banach $C^0(J_\delta; \mathbb{K}^n)$.

che soddisfa la (3.13) per definizione di G. La restrizione di tale funzione all'intervallo aperto $I := (t_0 - \delta, t_0 + \delta)$ è quindi una soluzione del problema di Cauchy (3.12).

Dimostrazione dell'unicità. Consideriamo, nell'intervallo $I := (t_0 - \delta, t_0 + \delta)$ ottenuto sopra, una seconda soluzione $\mathbf{x}' = \mathbf{x}'(t)$ dell'equazione integrale (3.13) a priori differente dalla soluzione $\mathbf{x} = \mathbf{x}(t)$. Per costruzione, per ogni intervallo chiuso $J_{\delta'} := [t_0 - \delta', t_0 + \delta']$ con $0 < \delta' < \delta$ la funzione $G : \mathcal{B}_\epsilon^{(J_{\delta'})} \to \mathcal{B}_\epsilon^{(J_{\delta'})}$ è ancora una mappa di contrazione e $\mathbf{x}' = \mathbf{x}'(t)$ è un suo punto fisso. Un altro punto fisso è la restrizione $\mathbf{x} = \mathbf{x}(t)$ all'intervallo $J_{\delta'}$. Per l'unicità del punto fisso le due soluzioni devono coincidere in $[t_0 - \delta', t_0 + \delta']$. Dato che δ' è fissabile arbitrariamente in $(0, \delta)$, le due soluzioni coincidono su $I = (t_0 - \delta, t_0 + \delta)$. $\square$

Più avanti sarà utile il seguente lemma che è stato provato nella dimostrazione del problema precedente.

Lemma 3.1. *In riferimento alle ipotesi del teorema 3.3, sia $Q \ni (t_0, \mathbf{x}_0)$ un insieme aperto chiusura compatta con $\overline{Q} \subset \Omega$ sul quale $\mathbf{f}$ è lipschitziana nella variabile $\mathbf{x}$ con costante $L \geq 0$. Posto $M := \max\{||\mathbf{f}(t, \mathbf{x})|| \mid (t, \mathbf{x}) \in \overline{Q}\}$ e $B_\epsilon(\mathbf{x}_0) := \{\mathbf{x} \in \mathbb{K}^n \mid ||\mathbf{x} - \mathbf{x}_0|| \leq \epsilon\}$, se $\delta, \epsilon > 0$ sono piccoli a sufficienza da soddisfare:*

(i) $0 < \delta < \min\{\epsilon/M, 1/L\}$,

(ii) $[t_0 - \delta, t_0 + \delta] \times B_\epsilon(\mathbf{x}_0) \subset Q$,

allora $I := (t_0 - \delta, t_0 + \delta)$ è un intervallo su cui vale la tesi del teorema 3.3 e la soluzione soddisfa $(t, \mathbf{x}(t)) \in I \times B_\epsilon(\mathbf{x}_0)$ per ogni $t \in I$. $\diamondsuit$

Osservazioni 3.5.

(1) La richiesta di locale lipschitzianità non è necessaria per l'*esistenza* di una soluzione. Esiste infatti un teorema, dovuto a Peano ed esteso da Picard, che prova l'esistenza di una soluzione per un qualsiasi problema di Cauchy con $\mathbf{f}$ continua, ma non necessariamente lipschitziana. L'esempio classico è il problema di Cauchy con $\Omega = \mathbb{R} \times \mathbb{R}$,

$$\begin{cases} \dfrac{dx}{dt} & = f(x), \\ x(0) & = 0\,. \end{cases}$$

dove f è la funzione che non è localmente lipschitziana in ogni intorno di $(t, x) = (0, 0)$ definita in (3.11). Tuttavia la funzione, per $t \in \mathbb{R}$,

$$x(t) := \begin{cases} t^2/4 & \text{se } x \geq 0\,, \\ 0 & \text{se } t < 0\,, \end{cases} \tag{3.14}$$

risolve il problema di Cauchy.

(2) La richiesta di lipschitzianità è generalmente necessaria per avere l'*unicità* della soluzione. A titolo di esempio, si osservi che, in aggiunta alla soluzione (3.14), il problema di Cauchy nell'esempio precedente ammette anche la soluzione $x(t) = 0$ per ogni $t \in \mathbb{R}$.

3.3.2 Condizione per gli integrali primi.

Mostriamo come il teorema di esistenza ed unicità implichi una condizione necessaria e sufficiente affinché una funzione sia un integrale primo (definizione 3.1). Enunciamo il teorema nel caso di sistemi di grado 1, lasciamo al lettore l'estesione (ovvia) al caso generale.

Teorema 3.4. *Sia* $\mathbf{f} : \Omega \to \mathbb{K}^n$ *una funzione continua e localmente lipschitziana nella variabile* $\mathbf{x} \in \mathbb{K}^n$ *sull'aperto* $\Omega \subset \mathbb{R} \times \mathbb{K}^n$. *Una funzione* $F : \Omega \to \mathbb{K}$ *ovunque differenziabile è un integrale primo per il sistema di equazioni differenziali*

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}\left(t, \mathbf{x}(t)\right)$$

se e solo se vale:

$$\mathbf{f}(t, \mathbf{x}) \cdot \nabla_{\mathbf{x}} F|_{(t,\mathbf{x})} + \left.\frac{\partial F}{\partial t}\right|_{(t,\mathbf{x})} = 0\,, \quad \textit{per ogni } (t, \mathbf{x}) \in \Omega. \tag{3.15}$$

$\diamondsuit$

Dimostrazione. Si supponga che valga la condizione (3.15). Se $\mathbf{x} = \mathbf{x}(t)$ con $t \in I$ intervallo aperto in $\mathbb{R}$ è una soluzione del sistema, allora

$$\frac{d}{dt} F(t, \mathbf{x}(t)) = \mathbf{f}(t, \mathbf{x}) \cdot \nabla_{\mathbf{x}} F|_{(t,\mathbf{x}(t))} + \left.\frac{\partial F}{\partial t}\right|_{(t,\mathbf{x}(t))} = 0\,,$$

per cui $I \ni t \mapsto F(t, \mathbf{x}(t))$ è costante su ogni soluzione.
Viceversa, se F differenziabile è integrale primo, allora vale

$$\frac{d}{dt} F(t, \mathbf{x}(t)) = \mathbf{f}(t, \mathbf{x}) \cdot \nabla_{\mathbf{x}} F|_{(t,\mathbf{x}(t))} + \left.\frac{\partial F}{\partial t}\right|_{(t,\mathbf{x}(t))} = 0\,,$$

su ogni soluzione. D'altra parte per il teorema 3.3, per ogni punto $(t_0, \mathbf{x}_0) \in \Omega$ c'è una soluzione del sistema di equazioni differenziali che soddisfa $\mathbf{x}(t_0) = \mathbf{x}_0$ e pertanto:

$$\mathbf{f}(t_0, \mathbf{x}_0) \cdot \nabla_{\mathbf{x}} F|_{(t_0,\mathbf{x}_0)} + \left.\frac{\partial F}{\partial t}\right|_{(t_0,\mathbf{x}_0)} = 0\,.$$

Questo conclude la dimostrazione in virtù dell'arbitrarietà del punto $(t_0, \mathbf{x}_0) \in \Omega$. $\square$

3.3.3 Teorema di esistenza ed unicità globale per il problema di Cauchy.

Ci occupiamo ora di dimostrare l'esistenza e l'unicità delle soluzioni "in grande". Per prima cosa dimostriamo la seguente proposizione.

Proposizione 3.3. *Sia* $\mathbf{f} : \Omega \to \mathbb{K}^n$ *una funzione continua e localmente lipschitziana nella variabile* $\mathbf{x} \in \mathbb{K}^n$ *sull'aperto* $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ *e si consideri il problema di Cauchy per* $(t_0, \mathbf{x}_0) \in \Omega$*:*

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(t, \mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.16}$$

Se $\mathbf{x}_1 : I_1 \to \mathbb{K}^n$ *e* $\mathbf{x}_2 : I_2 \to \mathbb{K}^n$ *sono due soluzioni del problema di Cauchy (3.16) definite sugli intervalli aperti* $I_1 \ni t_0$ *e* $I_2 \ni t_0$ *rispettivamente, allora*

$$\mathbf{x}_1(t) = \mathbf{x}_2(t), \quad \textit{se } t \in I_1 \cap I_2.$$

$\diamondsuit$

Dimostrazione. Per il teorema 3.3, le due soluzioni coincideranno sicuramente nell'intervallo che contiene t_0, $I \cap I_1 \cap I_2$, dove I è l'intervallo in cui vale il teorema di unicità . Se $I \supset I_1 \cap I_2$ non c'è nulla da provare. Supponiamo quindi che $I \not\supset I_1 \cap I_2$. Per fissare le idee lavoriamo nell'intorno destro di t_0 (la prova per l'intorno sinistro è analoga). Supponiamo dunque che, per un certo $t_1 > t_0$, valga $\mathbf{x}_1(t) = \mathbf{x}_2(t)$ se $t_0 \leq t < t_1$. Supponiamo però che $\mathbf{x}_1$ e $\mathbf{x}_2$ siano *entrambe* definite in un intervallo più grande: per $t_0 \leq t < t_2$ con $t_2 > t_1$. Vogliamo mostrare che esse devono coincidere su tutto questo intervallo: $\mathbf{x}_1(t) = \mathbf{x}_2(t)$ se $t_0 \leq t < t_2$.
Sia

$$s := \sup\{T \in \mathbb{R} \mid t_0 \leq T < t_2, \quad \mathbf{x}_1(t) = \mathbf{x}_2(t) \text{ se } t_0 \leq t < T.\}$$

Tale estremo superiore esiste finito perché l'insieme considerato è non vuoto e superiormente limitato. Se fosse $s < t_2$ avremmo per continuità che $\mathbf{x}_1(s) = \mathbf{x}_2(s)$, dato che le due soluzioni sono continue e coincidono per tutti i punti precedenti a s. Allora impostando il problema di Cauchy con la stessa funzione $\mathbf{f}$ e condizione iniziale $\mathbf{x}(s) = \mathbf{x}_1(s)$, avremmo che esiste un intervallo aperto $I \ni s$ in cui il problema di Cauchy ammette un unica soluzione. Ne conseguirebbe che in un intorno destro di s varrebbe ancora $\mathbf{x}_1(t) = \mathbf{x}_2(t)$. Ma questo sarebbe impossibile per la definizione di s. Concludiamo che $s = t_2$. Pertanto per la definizione di s, $\mathbf{x}_1(t) = \mathbf{x}_2(t)$ se $t_0 \leq t < t_2$ $\square$

Passiamo a considerare le *soluzioni massimali* di un problema di Cauchy. La seguente definizione precisa cosa intendiamo per soluzione massimale. Dal punto di vista intuitivo, intendiamo una soluzione che non possa essere estesa in una soluzione più grande dello stesso problema di Cauchy. Si osservi che una soluzione può essere massimale pur non essendo definita su tutto $(-\infty, +\infty)$.

Definizione 3.5. Sia $\mathbf{f} : \Omega \to \mathbb{K}^n$ una funzione sull'aperto $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ e si consideri il problema di Cauchy per $(t_0, \mathbf{x}_0) \in \Omega$:

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(t, \mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.17}$$

Una soluzione di (3.17) $\mathbf{x} : I \to \mathbb{K}^n$, con $I \ni t_0$ intervallo aperto, è detta **soluzione massimale** se non esiste alcuna altra soluzione di (3.17) $\mathbf{x}' : I' \to \mathbb{K}^n$, con $I' \ni t_0$ intervallo aperto, soddisfacente le condizioni:

(i) $I' \supsetneq I$,

(ii) $\mathbf{x}(t) = \mathbf{x}'(t)$ se $t \in I$.

Una soluzione massimale $\mathbf{x} : I \to \mathbb{K}^n$ è detta:

completa, se $I = (-\infty, +\infty)$;

completa nel futuro, se $I = (a, +\infty)$ con $a \in \mathbb{R} \cup \{-\infty\}$;

completa nel passato, se $I = (-\infty, b)$ con $b \in \mathbb{R} \cup \{+\infty\}$.

Le definizioni date si estendono banalmente al caso di un problema di Cauchy riferito ad un sistema di equazioni differenziali di ordine $n > 1$. $\diamondsuit$

Si osservi che non è affatto ovvio che, se due soluzioni massimali esistono allora devono coincidere[4]. Tuttavia, se $\mathbf{f}$ è tale da rispettare le richieste di continuità e locale lipschitzianità , vale il seguente risultato che prova l'esistenza e l'unicità delle soluzioni massimali come conseguenza della proposizione 3.3.

Teorema 3.5. (**Esistenza ed unicità globale**) *Sia* $\mathbf{f} : \Omega \to \mathbb{K}^n$ *una funzione continua e localmente lipschitziana nella variabile* $\mathbf{x} \in \mathbb{K}^n$ *sull'aperto* $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ *e si consideri il problema di Cauchy per* $(t_0, \mathbf{x}_0) \in \Omega$*:*

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(t, \mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.18}$$

Esiste ed è unica la soluzione massimale $\mathbf{x} : I \to \mathbb{K}^n$*. Tale soluzione è di classe* C^1*.*
Inoltre, se $\mathbf{x}_J : J \to \mathbb{K}^n$*, con* $J \ni t_0$ *intervallo aperto, è un'altra soluzione di (3.18), allora*

(i) $J \subset I$,

(ii) $\mathbf{x}(t) = \mathbf{x}_J(t)$ *se* $t \in J$. $\diamondsuit$

Dimostrazione. L'insieme $\mathfrak{S}$ delle soluzioni del problema di Cauchy (3.18) è non vuoto (per il teorema 3.3) ed ogni soluzione $\mathbf{x}_J : J \to \mathbb{K}^n$, con $J \ni t_0$ intervallo aperto, è univocamente determinata dal suo dominio per la proposizione 3.3. Se $I := \cup_{J \in \mathfrak{S}} J$, definiamo

$$\mathbf{x}_I(t) := \mathbf{x}_J(t)\,, \quad \text{se } t \in J \text{ e per } J \in \mathfrak{S}.$$

La definizione è ben posta perché , dalla proposizione 3.3 segue facilmente che se $t \in J \cap J'$ con $J, J' \in \mathfrak{S}$, allora $\mathbf{x}_J(t) = \mathbf{x}_{J'}(t)$. Pertanto si ha anche che $\mathbf{x}_I \restriction_J = \mathbf{x}_J$ per ogni $J \in \mathfrak{S}$ e perciò $\mathbf{x}_I$ risolve il problema di Cauchy se ristretta ad ogni $J \in \mathfrak{S}$. Infine, per costruzione I è un intervallo aperto e quindi $\mathbf{x}_I \in \mathfrak{S}$. Per costruzione $\mathbf{x} := \mathbf{x}_I$ è una soluzione massimale del nostro problema di Cauchy.

Due soluzioni massimali del problema di Cauchy (3.18), se esistono, devono coincidere sull'intersezione dei domini per la proposizione 3.3. Infine i due domini devono coincidere altrimenti una

[4] L'esistenza di soluzioni massimali è invece una diretta conseguenza del Lemma di Zorn.

delle due soluzioni (o entrambe) non sarebbe soluzione massimale. Ogni soluzione (massimale o no) deve essere di classe C^1 dato che $\mathbf{f}$ è continua. $\square$

Il teorema appena provato ha una conseguenza immediata riguardante i problemi di Cauchy per equazioni differenziali di ordine superiore. Il risultato è espresso nel seguente corollario che, in particolare, prova anche il teorema 2.2.

Corollario al teorema 3.5. *Sia* $\mathbf{f} : \Omega \to \mathbb{K}^m$ *dove* $\Omega = I \times D$ *con* $I \subset \mathbb{R}$ *intervallo aperto e* $D \subset \mathbb{K}^m \times \cdots \times \mathbb{K}^m$ *insieme aperto non vuoto e si consideri il problema di Cauchy*

$$\begin{cases} \dfrac{d^n\mathbf{x}}{dt^n} &= \mathbf{f}\left(t, \mathbf{x}(t), \frac{d\mathbf{x}}{dt}, \ldots, \frac{d^{n-1}\mathbf{x}}{dt^{n-1}}\right), \\ \mathbf{x}(t_0) &= \mathbf{x}_0\,, \frac{d\mathbf{x}}{dt}(t_0) = \mathbf{x}_0^{(1)}\,, \cdots, \frac{d^{n-1}\mathbf{x}}{dt^{n-1}}(t_0) = \mathbf{x}_0^{(n-1)} \end{cases} \tag{3.19}$$

dove $(t_0, \mathbf{x}_0, \mathbf{x}_0^{(1)}, \ldots, \mathbf{x}_0^{(n-1)}) \in \Omega$ *è fissato arbitrariamente. Se* $\mathbf{f}$ *soddisfa:*

(i) è continua su Ω*,*

(ii) per ogni $t \in I$*,* $D \ni (\mathbf{x}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n-1)}) \mapsto \mathbf{f}(t, \mathbf{x}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n-1)})$ *è differenziabile,*

(iii) le derivate delle componenti di $\mathbf{f}$ *nelle variabili in* D *sono funzioni continue su* Ω*,*

allora esiste ed è unica una soluzione massimale del problema di Cauchy (3.19). Tale soluzione è di classe C^k*.*

Dimostrazione. Passiamo al sistema del prim'ordine equivalente a (3.19)

$$\begin{cases} \dfrac{d\mathbf{X}}{dt} &= \mathbf{F}\left(t, \mathbf{X}(t)\right), \\ \mathbf{X}(t_0) &= \mathbf{X}_0. \end{cases}$$

come spiegato nella sezione 3.1.1. Nelle ipotesi fatte si ha immediatamente che $\mathbf{F}(t, \mathbf{X})$ è continua nelle sue variabili congiuntamente, è differenziabile nella variabile $\mathbf{X}$ e le derivate definiscono funzioni continue di $(t, \mathbf{X})$ congiuntamente. Per l'esempio 3.2.**1**, la funzione $\mathbf{F}(t, \mathbf{X})$ risulta essere localmente lipschitziana nella variabile $\mathbf{X}$. Possiamo applicare tale teorema 3.5 concludendo la dimostrazione. $\square$

Esempi 3.3.

1. Consideriamo un'equazione differenziale di ordine n su $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$ indifferentemente,

$$p_0(t)\frac{d^n x}{dt^n} + p_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + p_n(t)x = q(t)\,, \tag{3.20}$$

dove le funzioni $p_k : \mathbb{R} \to \mathbb{K}$ e $q : \mathbb{R} \to \mathbb{C}$ sono tutte continue e la funzione, incognita dell'equazione, $x = x(t)$ assume valori su $\mathbb{K}$ ed è supposta differenziabile n volte. L'equazione scritta sopra è detta **equazione differenziale lineare non omogenea**. La non omogeneità è dovuta alla presenza del cosiddetto **termine noto**: la funzione q. Nell'insieme aperto $U \subset \mathbb{R}$ in cui $p_0(t) \neq 0$, si può riportare l'equazione di sopra in forma normale

$$\frac{d^n x}{dt^n} + P_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + P_n(t)x = Q(t)\,, \tag{3.21}$$

avendo definito $P_k(t) := p_k(t)/p_1(t)$ e $Q(t) := q(t)/p(t)$ su U. I punti $t \in \mathbb{R}$ per cui $p_1(t) = 0$ si dicono **punti singolari** e devono essere studiati a parte.
Se $J \subset U$ è un intervallo aperto, consideriamo il problema di Cauchy di ordine n su $\Omega = J \times \mathbb{K}$ dato dall'equazione (3.21), con l'aggiunta delle condizioni iniziali, se $t_0 \in J$,

$$x(t_0) = x_0 \,, \quad \frac{dx}{dt}(t_0) = x_0^{(1)} \,, \cdots \,, \frac{d^{n-1}x}{dt^{n-1}}(t_0) = x_0^{(n-1)} \,, \tag{3.22}$$

dove $x_0, x_0^{(1)}, \dots x_0^{(n-1)} \in \mathbb{K}$.
Con la tecnica esposta nella sezione 3.1.1, il problema di Cauchy suddetto si può sempre ridurre ad un problema del primo'ordine su $\Omega := J \times \mathbb{K}^n$ della forma:

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = -A(t)\mathbf{x}(t) + \mathbf{b}(t), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.23}$$

dove: $J \ni t_0$, $\mathbf{x}(t) := (x(t), dx/dt, \cdots, d^{n-1}x/dt^{n-1})^t$. Inoltre $\mathbf{x}_0 := (x_0, x_0^{(1)}, \cdots, x_0^{(n-1)})^t$ e, per ogni $t \in I$, $A(t) : \mathbb{K}^n \to \mathbb{K}^n$ è una matrice $n \times n$ a coefficienti in $\mathbb{K}$ la cui prima riga è data dal vettore riga $(P_1(t), \dots, P_0(t))$ e i coefficienti delle successive righe valgono 0 oppure 1 in conformità con le $n-1$ relazioni:

$$\begin{array}{rcl} \dfrac{dx_{k-2}}{dt} & = & x_{k-1} \,, \\ \cdots & = & \cdots \\ \dfrac{dx_1}{dt} & = & x_2 \,, \\ \dfrac{dx}{dt} & = & x_1 \,. \end{array}$$

Infine il vettore colonna $\mathbf{b}(t)$ ha coefficienti tutti nulli, eccettuato il primo che vale $Q(t)$.
Il problema di Cauchy (3.23) è ben posto in quanto la funzione a secondo membro

$$\mathbf{f}(t, \mathbf{x}) := -A(t)\mathbf{x} + \mathbf{b}(t)$$

è continua ed è sicuramente localmente lipschitziana nelle ipotesi fatte: la costante di Lipschits può essere presa come $\max_{t \in J'} ||A(t)||$, su ogni insieme $J' \times B_\epsilon(\mathbf{x}_0)$, dove $J' \ni t_0$ è un intervallo aperto la cui chiusura è contenuta in J e $B_\epsilon(\mathbf{x}_0) \subset \mathbb{K}^n$ una palla aperta centrata in $\mathbf{x}_0$ di raggio $\epsilon > 0$ arbitario. Di conseguenza c'è sempre un'unica soluzione massimale del sistema considerato.
2. In riferimento al problema di Cauchy del prim'ordine su $\Omega = J \times \mathbb{K}^n$ (3.23) dell'esempio 1 (e quindi all'equivalente problema di Cauchy di ordine n), dimostriamo che la soluzione massimale ha come dominio in t tutto l'intervallo J. *Pertanto, se $J = \mathbb{R}$, la soluzione massimale è anche completa.*
Consideriamo inizialmente il caso di sistema omogeneo, per cui il problema di Cauchy è

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}(t, \mathbf{x}(t)), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.24}$$

dove

$$\mathbf{f}(t,\mathbf{x}) := -A(t)\mathbf{x}\,, \quad \text{con } (t,\mathbf{x}) \in \Omega := J \times \mathbb{K}^n,$$

e $J \ni t \mapsto A(t)$ è una funzione a valori matriciali continua.
Sia $(\alpha,\omega) \subset J$ il dominio della soluzione massimale. Definiamo, in riferimento al lemma 3.1, $Q = (t_0 - \Delta, t_0 + \Delta) \times B_E(\mathbf{x}_0)$ con $E, \Delta > 0$ in modo che

$$J \supset [t_0 - \Delta, t_0 + \Delta] \supset (t_0 - \Delta, t_0 + \Delta)\,.$$

la costante di Lipschitz su Q può essere fissata come $L = \max_{t\in[t_0-\Delta,t_0+\Delta]} ||A(t)||$ (se $L = 0$ tutta la discussione diventa banale per cui assumeremo d'ora in poi che $L > 0$), mentre il massimo di $||\mathbf{f}||$ è , evidentemente, $M = LE$.
Se ω non è l'estremo destro di J, rimanendo in J, possiamo scegliere Δ sufficientemente grande in modo che $J \supset [t_0 - \Delta, t_0 + \Delta]$ e $t_0 + \Delta > \omega$. Consideriamo quindi la soluzione massimale $\mathbf{x} : (\alpha,\omega) \to \mathbb{K}^n$. Per il lemma 3.1, se $t_1 \in (\alpha,\omega)$, la soluzione massimale deve essere comunque ben definita nell'intervallo $(t_1 - \delta, t_1 + \delta)$ ed assumere valori in $B_\epsilon(\mathbf{x}(t_1))$ purché , $0 < \epsilon < E$ e $\delta < \min(1/L, \epsilon/M)$ cioé $\delta < (\epsilon/E)(1/L)$. Fissiamo $t_1 < \omega$ in modo tale che $|\omega - t_1| < 1/(2L)$. Allora, fissando ϵ abbastanza vicino a E, si può prendere δ sufficientemente piccolo da soddisfare $\delta < (\epsilon/E)(1/L)$, ma anche $|t_1 - \omega| < \delta$. Questo significa che $t_1 + \delta > \omega$, per cui la soluzione massimale, che è definita sicuramente in $(t_1 - \delta, t_1 + \delta)$ è anche definita nell'*intorno destro* di ω. ω non può dunque essere l'estremo destro dell'intervallo massimale su cui è definita la soluzione. L'unica possibilità è che ω coincida con l'estremo destro di J. Con la stessa procedura si prova che α coincide con l'estremo sinistro di J. Pertanto la soluzione massimale è definita su tutto J.
Consideriamo infine il problema non omogeneo

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}(t,\mathbf{x}(t)), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases} \tag{3.25}$$

dove

$$\mathbf{f}(t,\mathbf{x}) := -A(t)\mathbf{x} + \mathbf{b}(t)\,, \quad \text{con } (t,\mathbf{x}) \in \Omega := J \times \mathbb{K}^n,$$

con $J \ni t \mapsto A(t)$ funzione a valori matriciali e $J \ni t \mapsto \mathbf{b}(t) \in \mathbb{K}^n$ funzione a valori vettoriali. Entrambe le funzioni sono continue.
Per computo diretto (lasciato per esercizio) si vede che, per $t \in J$,

$$\mathbf{x}(t) = G(t)^{-1}\left(\mathbf{x}(t_0) + \int_{t_0}^{t} G(\tau)\mathbf{b}(\tau)\,d\tau\right) \tag{3.26}$$

è la soluzione (su tutto J) del problema (3.25). La matrice $G(t)$, $n \times n$ a coefficienti in $\mathbb{K}$, è la trasposta della matrice la cui k-esima colonna è , al variare di t, la soluzione massimale su J del problema omogeneo associato

$$\begin{cases} \dfrac{d\,\mathbf{g}_{(k)}}{dt} & = -A(t)^t\,\mathbf{g}_{(k)}(t), \\ \mathbf{g}_{(k)}(t_0) & = \boldsymbol{\delta}_{(k)}. \end{cases} \tag{3.27}$$

dove $\boldsymbol{\delta}_{(k)}$ è la k-esima colonna della matrice identità .
3. Consideriamo ancora il problema di Cauchy di ordine n su $\mathbb{R} \times \mathbb{K}^n$, con equazione lineare omogenea:

$$\frac{d^n x}{dt^n} + P_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + P_n(t)x = 0 \,, \tag{3.28}$$

dove le funzioni P_k sono continue su J, e con condizioni iniziali, se $t_0 \in J$,

$$x(t_0) = x_0 \,, \quad \frac{dx}{dt}(t_0) = x_0^{(1)} \,, \cdots \,, \frac{d^{n-1}x}{dt^{n-1}}(t_0) = x_0^{(n-1)} \,, \tag{3.29}$$

dove $x_0, x_0^{(1)}, \ldots x_0^{(n-1)} \in \mathbb{K}$.
Consideriamo due soluzioni definite su tutto J dell'equazione (3.28): x e x'. Data la linearità dell'equazione, è chiaro che se $a, b \in \mathbb{K}$, la funzione $J \ni t \mapsto ax(t) + bx'(t)$ è ancora soluzione della stessa equazione differenziale. Da ciò si conclude immediatamente che l'insieme delle soluzioni dell'equazione (3.28) definite su tutto J costituisce uno *spazio vettoriale* sul campo $\mathbb{K}$, con operazioni di spazio vettoriale definite "punto per punto" come segue. Se $a, b \in \mathbb{K}$ e $x, x' : J \to \mathbb{K}$ risolvono (3.28):

$$(ax + bx')(t) := ax(t) + bx'(t) \,, \text{per ogni } t \in J.$$

Vogliamo determinare la dimensione di tale spazio vettoriale. Fissiamo una volta per tutte un punto $t_0 \in J$. Dal teorema di unicità , ogni soluzione della (3.28) è biunivocamente determinata dai suoi dati iniziali (3.29). Inoltre l'insieme dei possibili dati iniziali $(x_0, x_0^{(1)}, \ldots, x_0^{(n-1)})$ ricopre tutto $\mathbb{K}^n$. Se x e x' soddisfano le condizioni iniziali rispettivamente:

$$x(t_0) = x_0 \,, \quad \frac{dx}{dt}(t_0) = x_0^{(1)} \,, \cdots \,, \frac{d^{n-1}x}{dt^{n-1}}(t_0) = x_0^{(n-1)}$$

e

$$x'(t_0) = x_0' \,, \quad \frac{dx'}{dt}(t_0) = x_0^{'(1)} \,, \cdots \,, \frac{d^{n-1}x'}{dt^{n-1}}(t_0) = x_0^{'(n-1)} \,,$$

la soluzione $x'' := ax + bx'$ soddisfa banalmente le condizioni iniziali (ed è l'unica a fare ciò):

$$\mathbf{x}''(t_0) = ax_0 + bx_0' \,, \quad \frac{dx''}{dt}(t_0) = ax_0^{(1)} + bx_0^{'(1)} \,, \cdots \,, \frac{d^{n-1}x''}{dt^{n-1}}(t_0) = ax_0^{(n-1)} + bx_0^{'(n-1)} \,.$$

Questo dimostra immediatamente che l'applicazione che associa $(x_0, x_0^{(1)}, \ldots, x_0^{(n-1)}) \in \mathbb{K}^n$ all'unica soluzione massimale del problema di Cauchy con equazione (3.28) con dati iniziali (3.29), è un omomorfismo di spazi vettoriali. Dato che la corrispondenza tra dati iniziali e soluzioni è biunivoca, concludiamo che tale omomorfismo è in isomorfismo e pertanto conserva le dimensioni degli spazi vettoriali identificati. Siamo giunti a poter affermare che: *lo spazio vettoriale su $\mathbb{K}$ delle soluzioni dell'equazione differenziale lineare omogenea a coefficienti continui (3.28) di ordine n su J ha dimensione pari all'ordine stesso n.*
4. Consideriamo ancora su $\mathbb{R} \times \mathbb{K}^n$ l'equazione lineare *non* omogenea:

$$\frac{d^n x}{dt^n} + P_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + P_n(t)x = Q(t) \,, \tag{3.30}$$

dove le funzioni P_k e Q sono continue su J. L'insieme di tutte le soluzioni $s : J \to \mathbb{K}$ di questa equazione si può scrivere come

$$\{s = s_0 + x \mid x : J \to \mathbb{K} \text{ è soluzione di (3.28)}\} ,$$

dove $s_0 : J \to \mathbb{K}$ è una qualsiasi fissata soluzione di (3.30). (Abbiamo ancora una volta usato la definizione di somma di funzioni "punto per punto" $(s_0 + x)(t) := s_0(t) + x(t)$ per ogni $t \in J$.) Infatti, se s e s_0 risolvono (3.30), per linearità dell'equazione differenziale, la loro differenza risolve (3.28). Viceversa, se s_0 risolve (3.30) e x risolve (3.28), ancora una volta per linearità, $s_0 + x$ risolve (3.30).

5. Il contenuto dei quattro esempi precedenti si generalizza facilmente al caso in cui si considera il problema di Cauchy di ordine n che usa l'equazione, nell'incognita $\mathbf{x} = \mathbf{x}(t) \in \mathbb{K}^m$, per ogni $t \in U \subset \mathbb{R}$ insieme aperto,

$$p_0(t)\frac{d^n\mathbf{x}}{dt^n} + p_1(t)\frac{d^{n-1}\mathbf{x}}{dt^{n-1}} + \cdots + p_n(t)\mathbf{x} = \mathbf{q}(t) . \tag{3.31}$$

Sopra, per ogni $t \in U$, le $p_k(t)$ sono *matrici* $m \times m$ con elementi in $\mathbb{K}$ e $\mathbf{q}(t)$ è un *vettore* colonna in $\mathbb{K}^m$ ed, al variare di $t \in U$, le funzioni suddette sono continue. Se su U vale $detp_0(t) \neq 0$, si può riportare l'equazione di sopra in forma normale

$$\frac{d^n\mathbf{x}}{dt^n} + P_1(t)\frac{d^{n-1}\mathbf{x}}{dt^{n-1}} + \cdots + P_n(t)\mathbf{x} = \mathbf{Q}(t) , \tag{3.32}$$

avendo definito $P_k(t) := p_k(t)(p_1(t))^{-1}$ e $\mathbf{Q}(t) := \mathbf{q}(t)(p(t))^{-1}$. Il problema di Cauchy si può impostare su $\Omega = J \times \mathbb{K}^m$ dove $J \subset U$ è un intervallo aperto. Ci si riduce ad un problema di Cauchy di ordine 1 in $J \times \mathbb{K}^{n \cdot m}$ con la solita procedura e pertanto valgono i risultati esposti prima.

6. Consideriamo ora le **equazioni differenziali lineari omogenee a coefficienti costanti di grado** n:

$$\frac{d^n x}{dt^n} + a_1\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_n x(t) = 0 . \tag{3.33}$$

Quest'equazione la pensiamo come definita su $(t, x) \in \mathbb{R} \times \mathbb{C}$. Il caso $(t, x) \in \mathbb{R} \times \mathbb{R}$ si ottiene per restrizione. I coefficienti $a_1, a_2, \ldots a_{n-1} \in \mathbb{C}$ sono dunque dei numeri fissati. Vogliamo scrivere esplicitamente la classe di tutte le soluzioni massimali dell'equazione suddetta. Dato che l'insieme delle soluzioni è uno spazio vettoriale di dimensione n, la generica soluzione sarà una combinazione lineare di n soluzioni particolari e linearmente indipendenti di (3.33). La strategia che seguiremo è la seguente.

(i) Per prima cosa cerchiamo le soluzioni massimale, $x : \mathbb{R} \to \mathbb{C}$, dell'equazione di grado 1 e non omogenea

$$\left(\frac{d}{dt} - b\right) x(t) = f(t) , \tag{3.34}$$

dove $f : \mathbb{R} \to \mathbb{C}$ è una funzione continua assegnata e $a \in \mathbb{C}$ una costante, e dove abbiamo usato la notazione:

$$\left(\frac{d}{dt} + a\right) x(t) := \frac{dx}{dt} + ax(t) .$$

(ii) Poi mostriamo che l'equazione (3.33) si può riscrivere:

$$\left(\frac{d}{dt} - b_1\right)\left(\frac{d}{dt} - b_2\right)\cdots\left(\frac{d}{dt} - b_n\right)x(t) = 0\,, \tag{3.35}$$

dove $b_1, b_2, \ldots, b_n$ sono le soluzioni in $\mathbb{C}$, eventualmente alcune di esse coincidenti, dell'equazione algebrica in $\chi \in \mathbb{C}$:

$$\chi^n + a_1\chi^{n-1} + \cdots + a_{n-1}\chi + a_n = 0\,. \tag{3.36}$$

(iii) Conoscendo la soluzione generale di (3.34) ed applicando la formula ricorrentemente all'espressione di sopra, si ricava la classe completa delle soluzioni di (3.34).
Partiamo dal punto (i). Sappiamo, dall'esempio 4, che le soluzioni di (3.34) sono scrivibili come $s(t) = s_0(t) + x(t)$ con $t \in \mathbb{R}$, dove s_0 è una soluzione particolare fissata arbitrariamente di (3.34) e x spazia nell'insieme delle soluzioni dell'equazione omogenea:

$$\left(\frac{d}{dt} - b\right)x(t) = 0\,. \tag{3.37}$$

Le soluzioni di quest'ultima si ottengono immediatamente come, per $c \in \mathbb{C}$ fissato arbitrariamente,

$$x(t) = ce^{bt}\,, \quad t \in \mathbb{R}\,.$$

Si vede che, effettivamente si ha uno spazio vettoriale di dimensione 1 (sul campo $\mathbb{C}$). Per computo diretto, si vede che una soluzione particolare della (3.34) è

$$s_0(t) = e^{bt}\int_{t_0}^{t} f(\tau)e^{-b\tau}\,d\tau\,, \quad t \in \mathbb{R}\,,$$

dove $t_0 \in \mathbb{R}$ è un punto fissato arbitrariamente. In definitiva, fissato $t_0 \in \mathbb{R}$ arbitrariamente, tutte le soluzioni di (3.34) si ottengono al variare di $c \in \mathbb{C}$ come:

$$x_c(t) = ce^{bt} + e^{bt}\int_{t_0}^{t} f(\tau)e^{-b\tau}\,d\tau\,, \quad t \in \mathbb{R}\,, \tag{3.38}$$

Passiamo al punto (ii). Definendo $\frac{d^0}{dt^0}$ come l'operatore identità , consideriamo l'insieme $\mathcal{D}$ degli operatori differenziali

$$a_0\frac{d^n}{dt^n} + a_1\frac{d^{n-1}}{dt^{n-1}} + \cdots + a_n\frac{d^0}{dt^0}\ :\ C^\infty(\mathbb{R};\mathbb{C}) \to C^\infty(\mathbb{R};\mathbb{C})$$

al variare dei numeri $a_0, \ldots, a_n \in \mathbb{C}$ ed infine al variare dell'ordine $n = 0, 1, 2, \ldots$. Si verifica facilmente che $\mathcal{D}$ è un *anello commutativo* rispetto alla somma definita come, per ogni $D, D' \in \mathcal{D}$

$$(D + D')(f) := D(f) + D'(f)\,, \quad \text{per ogni } f \in C^\infty(\mathbb{R};\mathbb{C})$$

ed al prodotto definito come la composizione operatoriale: per ogni $D, D' \in \mathcal{D}$

$$(DD')(f) := D(D'(f))\,.$$

Se $\mathcal{P}(\mathbb{C})$ indica l'anello dei polinomi, nell'indeterminata χ, di ordine arbitrario e sul campo $\mathbb{C}$, l'applicazione

$$F:\ a_0\chi^n + a_1\chi^{n-1} + \cdots + a_n \ \mapsto\ a_0\frac{d^n}{dt^n} + a_1\frac{d^{n-1}}{dt^{n-1}} + \cdots + a_n\frac{d^0}{dt^0}x(t)$$

è biettiva: la suriettività è di immediata verifica, per l'iniettività è sufficiente dimostrare che se, per $p \in \mathcal{P}(\mathbb{C})$, $F(p)$ è l'operatore nullo di $\mathcal{D}$ (cioé quello che trasforma ogni funzione $f \in C^\infty(\mathbb{R};\mathbb{C})$ nella funzione che vale costantemente zero), allora p è il polinomio nullo. In effetti se $p(\chi) = a_0\chi^n + a_1\chi^{n-1} + \cdots + a_n$ e $F(p) = a_0\frac{d^n}{dt^n} + a_1\frac{d^{n-1}}{dt^{n-1}} + \cdots + a_n\frac{d^0}{dt^0}$ è l'operatore nullo, dovendo essere $(a_0\frac{d^n}{dt^n} + a_1\frac{d^{n-1}}{dt^{n-1}} + \cdots + a_n\frac{d^0}{dt^0})(f) = 0$ e scegliendo per f la funzione che vale costantemente a_n, si ha $a_n^2 = 0$ per cui $a_n = 0$. Scegliendo f come la funzione che vale $a_{n-1}t$ e reiterando la procedura, si ha che $a_{n-1} = 0$ e via di seguito, trovando alla fine che p è il polinomio nullo come richiesto. Infine, come è immediato provare, usando in particolare il fatto che $\frac{d^n}{dt^n}\frac{d^m}{dt^m} = \frac{d^{n+m}}{dt^{n+m}}$, si verifica che F e F^{-1} conservano i prodotti e le somme. In definitiva F è un *isomorfismo di anelli.*

Dato che, per il *teorema fondamentale dell'algebra*, ogni polinomio di $p \in \mathcal{P}(\mathbb{C})$,

$$p(\chi) = \chi^n + a_1\chi^{n-1} + \cdots + a_{n-1}\chi + a_n$$

è decomponibile come prodotto di polinomi elementari:

$$p(\chi) = (\chi - b_1)\cdots(\chi - b_n)$$

dove $b_1, b_2, \ldots, b_n$ sono le n soluzioni di $p(\chi) = 0$, attraverso l'isomorfismo F giungiamo a dimostrare che vale la decomposizione (3.35).

Passiamo al punto (iii) L'equazione differenziale iniziale (3.33) è equivalente all'equazione (3.35):

$$\left(\frac{d}{dt} - b_1\right)\left(\frac{d}{dt} - b_2\right)\cdots\left(\frac{d}{dt} - b_n\right)x(t) = 0\,,$$

cioé

$$\left(\frac{d}{dt} - b_1\right)\left[\left(\frac{d}{dt} - b_2\right)\cdots\left(\frac{d}{dt} - b_n\right)x(t)\right] = 0\,.$$

Tenendo conto di della soluzione generale (3.38) dell'equazione (3.37), concludiamo che deve essere, per $c_1 \in \mathbb{C}$:

$$\left(\frac{d}{dt} - b_2\right)\cdots\left(\frac{d}{dt} - b_n\right)x(t) = c_1 e^{b_1 t}\,.$$

Possiamo iterare la procedura ottenendo

$$\left(\frac{d}{dt} - b_3\right)\cdots\left(\frac{d}{dt} - b_n\right)x(t) = c_2 e^{b_2 t} + c_1 e^{b_2 t}\int_{t_0}^{t} e^{b_1\tau}e^{-b_2\tau}d\tau\,.$$

Notiamo che, se $b_1 \neq b_2$, si trova, raggruppando le costanti a fattore delle stesse funzioni esponenziali,

$$\left(\frac{d}{dt} - b_3\right) \cdots \left(\frac{d}{dt} - b_n\right) x(t) = k_2 e^{b_2 t} + k_1 e^{b_1 t} \, .$$

Se invece $b_1 = b_2$, il calcolo diretto fornisce

$$\left(\frac{d}{dt} - b_3\right) \cdots \left(\frac{d}{dt} - b_n\right) x(t) = k_2 e^{b_1 t} + k_1 t \, e^{b_1 t} \, .$$

In entrambi i casi k_1 e k_2 sono costanti arbitrarie in $\mathbb{C}$. Per esercizio, si può provare, iterando la procedura fino ad avere a primo membro solo la funzione $x = x(t)$, la ricetta che fornisce la soluzione generale dell'equazione (3.33)

$$\frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \cdots + a_n x(t) = 0 \, .$$

Tale ricetta è la seguente. Si cercano le soluzioni in $\mathbb{C}$ del polinomio, detto **polinomio caratteristico** associato

$$\chi^n + a_1 \chi^{n-1} + \cdots + a_{n-1} \chi + a_n \, .$$

Si raggruppano le soluzioni per soluzioni coincidenti:
b_1 con molteplicità m_1,
b_2 con molteplicità m_2,
...,
b_k con molteplicità m_k,
in modo che $1 \leq m_j \leq n$ e $m_1 + \cdots + m_k = n$. La soluzione generale si scrive allora, al variare delle n costanti $C_{j,r_j} \in \mathbb{C}$ con $j = 1, 2, \ldots, k$ e $r_j = 1, \ldots, m_j$:

$$x(t) = \sum_{j=1}^{k} \sum_{r_j=1}^{m_j} C_{j,r_j} t^{j-1} e^{b_j t} \, , \quad \text{con } t \in \mathbb{R} \, . \tag{3.39}$$

Nella pratica: se la soluzione $b \in \mathbb{C}$ del polinomio caratteristico ha molteplicità 1 (cioé non ci sono radici coincidenti con quel valore) essa contribuisce alla soluzione generale con la sola funzione esponenziale e^{bt}. Se ha molteplicità 2, contribuisce con due funzioni: e^{bt} e $t\, e^{bt}$. Se ha molteplicità m, contribuisce con m funzioni e^{bt}, te^{bt}, $t^2 e^{bt}$,...,$t^{m-1} e^{bt}$. Ogni soluzione contribuisce in questo modo ed infine, tutti i singoli contributi vanno sommati avendoli preventivamente moltiplicati per costanti complesse arbitrarie.
È chiaro che il metodo si può applicare anche per l'equazione (3.33) nel caso in cui i coefficienti $a_0, a_1, \ldots, a_n$ siano reali e si cerchino soluzioni reali: è sufficiente determinare prima tutte le soluzioni complesse con la procedura appena descritta e quindi, tenendo conto dell'identità di Eulero:

$$e^{(\alpha + i\beta)} = e^{\alpha} \left(\cos\beta + i \sin\beta\right) \, ,$$

scegliere le costanti C_{j,r_j} (generalmente complesse) in modo tale che l'espressione finale di $x(t)$ risulti essere reale.

Esercizi 3.1.

1. Sia $\mathbf{f} : \mathbb{R} \times D \to \mathbb{K}^n$ una funzione continua e localmente lipschitziana nella variabile $\mathbf{x} \in \mathbb{K}^n$, con $D \subset \mathbb{K}^n$ aperto non vuoto. Si supponga infine che valga, per un fissato $T \in \mathbb{R}$,

$$\mathbf{f}(t - T, \mathbf{x}) = \mathbf{f}(t - T)\,, \quad \forall (t, \mathbf{x}) \in \mathbb{R} \times D\,.$$

Si scelga $k \in \mathbb{Z}$ arbitrariamente. Provare che $\mathbf{x} : (\alpha, \omega) \to \mathbb{K}^n$ è soluzione massimale del problema di Cauchy per $(t_0, \mathbf{x}_0) \in \Omega$:

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(t, \mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases},$$

se e solo se $\mathbf{x}'(t) := \mathbf{x}(t - kT)$ con $t \in (\alpha + kT, \omega + kT)$ è una soluzione massimale del problema di Cauchy

$$\begin{cases} \dfrac{d\mathbf{x}'}{dt} & = \mathbf{f}\left(t, \mathbf{x}'(t)\right), \\ \mathbf{x}'(t_0 - kT) & = \mathbf{x}_0. \end{cases},$$

2. Sia $\mathbf{f} : D \to \mathbb{K}^n$ una funzione localmente lipschitziana, con $D \subset \mathbb{K}^n$ aperto non vuoto. Si scelga $T \in \mathbb{R}$ arbitrariamente. Provare che $\mathbf{x} : (\alpha, \omega) \to \mathbb{K}^n$ è soluzione massimale del problema di Cauchy autonomo per $(t_0, \mathbf{x}_0) \in \Omega$:

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(\mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0. \end{cases},$$

se e solo se $\mathbf{x}'(t) := \mathbf{x}(t - T)$ con $t \in (\alpha + T, \omega + T)$ è una soluzione massimale del problema di Cauchy autonomo

$$\begin{cases} \dfrac{d\mathbf{x}'}{dt} & = \mathbf{f}\left(\mathbf{x}'(t)\right), \\ \mathbf{x}'(t_0 - T) & = \mathbf{x}_0. \end{cases},$$

3. Si consideri un punto materiale di massa $m > 0$ vincolato a stare su un piano orizzontale liscio in quiete in un riferimento inerziale. Si supponga che il punto sia sottoposto alla forza, con $g > 0$:

$$\mathbf{F}(\mathbf{x}) := \kappa r\, \mathbf{e}_r - mg\, \mathbf{e}_z\,,$$

dove è stato introdotto un sistema di coordinate cilindriche solidali con il piano r, φ, z rispetto all'origine O ed il piano ha equazione $z = 0$. Studiare il moto del punto nei due casi $\kappa > 0$ e $\kappa < 0$. Nel caso $\kappa < 0$ si determini la legge oraria del punto (soluzione massimale) con condizioni iniziali $\mathbf{x}(0) = \mathbf{0}$, $\mathbf{v}(0) = v(\,\mathbf{e}_x - \,\mathbf{e}_y)$.

4. Si consideri un punto materiale di massa $m > 0$ vincolato a muoversi su un piano orizzontale liscio in quiete in un riferimento inerziale. Si supponga che il punto sia sottoposto alla forza, con $\kappa < 0, \beta > 0, g > 0$,

$$\mathbf{F}(\mathbf{x}, \mathbf{v}) := \kappa r\, \mathbf{e}_r + \beta\, \mathbf{e}_z \wedge \mathbf{v} - mg\, \mathbf{e}_z\ ,$$

dove è stato introdotto un sistema di coordinate cilindirche solidali con il piano r, φ, z rispetto all'origine O ed il piano ha equazione $z = 0$.

5. Si consideri un punto materiale di massa $m > 0$ posto su un piano orizzontale liscio in quiete in un riferimento inerziale. Si supponga che il punto sia sottoposto alla forza, con $\kappa, \gamma > 0, g > 0$,

$$\mathbf{F}(\mathbf{x}, \mathbf{v}) := \kappa r\, \mathbf{e}_r - \gamma \mathbf{v} - mg\, \mathbf{e}_z\ ,$$

dove è stato introdotto un sistema di coordinate cilindirche solidali con il piano r, φ, z rispetto all'origine O ed il piano ha equazione $z = 0$. Studiare il moto del punto nei due casi $\kappa > 0$ e $\kappa < 0$.

3.3.4 Completezza di soluzioni massimali.

Ci occupiamo ora di alcune questioni tecniche riguardanti la completezza di soluzioni massimali nelle ipotesi di validità del teorema di esistenza ed unicità globale 3.5. Per prima cosa proviamo una proposizione che spiega cosa succede ad una soluzione massimale non completa. L'idea intuitiva è che una soluzione massimale $\mathbf{x} : I \to \mathbb{K}^n$ può non essere completa se, per grandi valori del parametro t, avvicinandosi al bordo di I, la curva integrale fatta dai punti $\{(t, \mathbf{x}(t)) \mid t \in I\}$ tenda ad "uscire" dal dominio Ω, cioé tenda a qualche punto su $\partial\Omega$, dove Ω è il dominio della funzione $\mathbf{f}$ che appare in (3.18). In realtà può anche accadere che per grandi t, avvicinandosi al bordo di I, la curva integrale non tenda a nulla.

Esempi 3.4.
1. Il primo esempio è banale. Consideriamo il problema di Cauchy su $\Omega = (-1, 1) \times \mathbb{R}$:

$$\begin{cases} \dfrac{dx}{dt} & = 1, \\ x(0) & = 0. \end{cases}$$

La soluzione massimale è ovviamente $x : (-1, 1) \ni t \mapsto t$. Questa non è completa e, per $t \to +1$, $(t, x(t)) \to (1, 1) \in \partial\Omega$. L'incompletezza di questo esempio è in qualche modo artificiale.
2. Questo secondo caso è molto meno artificiale ed il risultato é in qualche modo inaspettato. Si consideri $\Omega = \mathbb{R}^2$ e su di esso il problema di Cauchy

$$\begin{cases} \dfrac{dx}{dt} & = x^2, \\ x(0) & = 1. \end{cases}$$

Separando le variabili si ottiene la soluzione:

$$x(t) = \frac{1}{1-t}\ , \quad t \in (-\infty, 1)\ .$$

Questa è l'unica soluzione, evidentemente massimale, del problema di Cauchy. L'unicità è assicurata dal teorema di unicità globale dato che siamo banalmente nelle ipotesi. La soluzione non è completa nel futuro e, per $t \to 1^-$, accade che $(t, x(t))$ non tende a nulla, in quanto diverge.
3. Consideriamo infine un caso ancora più complicato sull'insieme $\Omega = (0, +\infty) \times \mathbb{R}^2$

$$\begin{cases} \dfrac{dx}{dt} & = -\dfrac{y}{t^2}, \\ \dfrac{dy}{dt} & = \dfrac{x}{t^2}, \\ x(2/\pi) & = 1, \quad y(2/\pi) = 0\,. \end{cases} .$$

Sull'insieme Ω, il secondo membro del sistema scritto sopra definisce una funzione di classe C^1 per cui siamo nelle ipotesi del teorema di esistenza ed unicità globale. Una soluzione del problema di Cauchy scritto è

$$\begin{cases} x(t) & = \sin(1/t), \\ y(t) & = \cos(1/t), \end{cases} \quad t \in (0, +\infty)\,. \tag{3.40}$$

La soluzione, che è evidentemente massimale, è unica per il teorema suddetto. In questo caso però il limite per $t \to 0^+$ della soluzione non tende ad alcun valore e nemmeno diverge. Tuttavia possiamo trovare una successione di punti t_k in $(0, +\infty)$ che tende a 0^+ per $k \to +\infty$ e tale che $(t_k, \mathbf{x}(t_k))$ tende ad un punto su $\partial\Omega$ per $k \to +\infty$. Per esempio, è sufficiente scegliere $t_0 \in (0, +\infty)$ arbitrariamente e definire $t_k = t_0/(1 + 2\pi k t_0)$ (in modo che $1/t_k = 2\pi k + 1/t_0$), $k = 1, 2, \ldots$. In tal modo $\mathbf{x}(t_k) = \mathbf{x}(t_0)$ *costantemente* per ogni $k = 1, 2, \ldots$ e pertanto esiste il limite di $(t_k, \mathbf{x}(t_k))$ e converge al punto $(0, \mathbf{x}(t_0)) \in \partial\Omega$. L'esistenza di una simile successione era preclusa nell'esempio precedente.

Possiamo provare la seguente proposizione generale.

Proposizione 3.4. *Nelle ipotesi del teorema 3.5 per* $\mathbf{f} : \Omega \to \mathbb{K}^n$, *sia* $\mathbf{x} : I \to \mathbb{K}^n$, *con* $I = (\alpha, \omega)$, *una soluzione massimale del problema di Cauchy (3.18).*
(a) *Se* $\mathbf{x}$ *non è completa nel futuro e* $\{t_k\}_{k\in\mathbb{N}} \subset I$ *è tale che* $t_k \to \omega$ *per* $k \to +\infty$ *ed esiste* $\lim_{k\to+\infty} \mathbf{x}(t_k) = \mathbf{x}_\omega \in \mathbb{K}^n$, *allora deve essere:*

$$(\omega, \mathbf{x}_\omega) \in \partial\Omega\,.$$

(b) *Se* $\mathbf{x}$ *non è completa nel passato e* $\{t_k\}_{k\in\mathbb{N}} \subset I$ *è tale che* $t_k \to \alpha$ *per* $k \to +\infty$ *ed esiste* $\lim_{k\to+\infty} \mathbf{x}(t_k) = \mathbf{x}_\alpha \in \mathbb{K}^n$, *allora deve essere:*

$$(\alpha, \mathbf{x}_\alpha) \in \partial\Omega\,.$$

$\diamondsuit$

Dimostrazione. Diamo la dimostrazione per il caso (a), quella per il caso (b) è del tutto analoga. Useremo il lemma 3.1. Supponiamo, per assurdo, che la soluzione non sia completa

nel futuro e che, per il limite della successione considerata, valga $(\omega, \mathbf{x}_\omega) \in \Omega$. Consideriamo un aperto $Q \subset \Omega$ che soddisfi $(\omega, \mathbf{x}_\omega) \in Q$ e $\overline{Q} \subset \Omega$. Scegliamo Q in modo che su $\overline{Q}$ valga la condizione di lipschitzianità locale $||\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{x}')|| \leq L||\mathbf{x} - \mathbf{x}'||$ se $(t, \mathbf{x}), (t, \mathbf{x}') \in \overline{Q}$; con costante $L \geq 0$. Quindi definiamo $M := \max\{||\mathbf{f}(t, \mathbf{x})|| \,|\, (t, \mathbf{x}) \in \overline{Q}\}$. Per fissare le idee prendiamo $\overline{Q}$ della forma

$$[\omega - \Delta, \omega + \Delta] \times B_E(\omega) \,,$$

dove $\Delta > 0$ e $B_E(\omega)$ è la palla chiusa in $\mathbb{K}^n$ centrata in ω e con raggio $E > 0$.
Dato che $t_k \to \omega$ e $\mathbf{x}(t_k) \to \mathbf{x}_\omega$, per $k \to +\infty$, scegliendo k sufficientemente grande possiamo trovare $\delta > 0$ e $\epsilon > 0$ sufficientemente piccoli da soddisfare contemporaneamente le tre condizioni ($B_\epsilon(\mathbf{x}(t_k))$ è la palla chiusa di raggio ϵ centrata in t_k):

(1) $|t_k - \omega| < \delta$,
(2) $[t_k - \delta, t_k + \delta] \times B_\epsilon(\mathbf{x}(t_k)) \subset [\omega - \Delta, \omega + \Delta] \times B_E(\mathbf{x}_\omega)$,
(3) $\delta < \min\{\epsilon/M, 1/L\}$.

Se $M = 0$, $\mathbf{f} \equiv 0$ su Q e, a partire dal primo t_k per cui $(t_k, \mathbf{x}(t_k)) \in Q$ l'orbita diventa costante per cui $\omega = +\infty$ e la soluzione è completa nel futuro contrariamente alle nostre ipotesi. Quindi deve essere $M > 0$. In questa situazione, per soddisfare (3), è sufficiente prendere $\epsilon = 2\delta M$ e quindi scegliere $0 < \delta < 1/(2L)$. Si può scegliere $\delta > 0$ piccolo a sufficienza in modo che valgano anche (1) e (2), questo è sempre possibile visto che $(t_k, \mathbf{x}(t_k))$ può essere reso arbitrariamente vicino a $(\omega, \mathbf{x}_\omega)$ per ipotesi.
La prima condizione assicura che $\omega \in (t_k - \delta, t_k + \delta)$. Le rimanenti due condizioni, in base al lemma 3.1, assicurano che esista una soluzione del problema di Cauchy

$$\begin{cases} \dfrac{d\mathbf{x}'}{dt} & = \mathbf{f}\left(t, \mathbf{x}'(t)\right), \\ \mathbf{x}'(t_k) & = \mathbf{x}(t_k). \end{cases} \tag{3.41}$$

definita in tutto l'intervallo $(t_k - \delta, t_k + \delta)$. Tale soluzione, per costruzione è un'estensione dell'originale soluzione $\mathbf{x} = \mathbf{x}(t)$. Infatti soddisfa la stessa equazione differenziale e si raccorda con la soluzione $\mathbf{x} = \mathbf{x}(t)$ per $t = t_k$ ed è definita anche nell'intorno destro di ω, dato che $\omega \in (t_k - \delta, t_k + \delta)$. Questo è impossibile visto che $\mathbf{x} : (\alpha, \omega) \to \mathbb{K}^n$ è una soluzione massimale. □

La proposizione 3.4 ha delle importanti conseguenze sulla completezza delle soluzioni massimali di un problema di Cauchy quando valgono alcune condizioni aggiuntive. Elenchiamo tali conseguenze in un'unica proposizione.

Proposizione 3.5. *Nelle ipotesi su $\mathbf{f} : \Omega \to \mathbb{K}^n$ date nel teorema 3.5, si supponga che $\Omega = \mathbb{R} \times D$ con $D \subset \mathbb{K}^n$ aperto non vuoto.*
Se $\mathbf{x} : I \to \mathbb{K}^n$, con $I = (\alpha, \omega)$, una soluzione massimale del problema di Cauchy (3.18) valgono i seguenti fatti.
(a) *Se l'orbita di $\mathbf{x} : I \to \mathbb{K}^n$ è contenuta in un compatto $K \subset D$ per $t \geq T \in I$, allora $\mathbf{x} : I \to \mathbb{K}^n$ è completa nel futuro.*
(b) *Se l'orbita di $\mathbf{x} : I \to \mathbb{K}^n$ è contenuta in un compatto $K \subset D$ per $t \leq T \in I$, allora*

$\mathbf{x} : I \to \mathbb{K}^n$ *è completa nel passato.*
(c) *Se* $D = \mathbb{K}^n$ *e* $\sup_{(t,\mathbf{x})\in\Omega} ||\mathbf{f}(t,\mathbf{x})|| < +\infty$ *allora* $\mathbf{x} : I \to \mathbb{K}^n$ *è completa.*
(d) *Se* $\mathbf{f}$ *non dipende dalla variabile* t *(cioé il sistema di equazioni differenziali associato è autonomo) e il supporto di* $\mathbf{f}$ *(definito rispetto alla topologia di* D*) è compatto, allora* $\mathbf{x} : I \to \mathbb{K}^n$ *è completa.*
$\diamondsuit$

Dimostrazione.
(a) Supponiamo che la soluzione non sia completa nel futuro, per cui $I = (\alpha, \omega)$ con $\omega \in \mathbb{R}$. Consideriamo una successione $\{t_k\}_{k\in\mathbb{N}} \subset (T, \omega)$ con $t_k \to \omega$ per $k \to +\infty$. Dato che $K \supset \{\mathbf{x}(t_k)\}_{k\in\mathbb{N}}$ è compatto in uno spazio metrico, è anche compatto per successioni, pertanto esisterà una sottosuccessione $\{\mathbf{x}(t_{k_p})\}_{p\in\mathbb{N}}$ convergente a qualche $\mathbf{x}_\omega \in K$. Per la proposizione 3.4 deve essere $(\omega, \mathbf{x}_\omega) \in \partial\Omega$. Essendo $\Omega = \mathbb{R} \times D$ deve necessariamente essere $\mathbf{x}_\omega \in \partial D$, che è impossibile visto che $\mathbf{x}_\omega \in K \subset D$ e $\partial D \cap D = \emptyset$ essendo D aperto. Quindi $\omega = +\infty$.
(b) La prova è strettamente analoga a quella data per il caso (a).
(c) Supponiamo che la soluzione non sia completa nel futuro, per cui $I = (\alpha, \omega)$ con $\omega \in \mathbb{R}$. Consideriamo le palle chiuse $B_k \subset \mathbb{K}^n$ centrate nell'origine e di raggio $k = 1, 2, \ldots$. Deve accadere che, per ogni $k \in \mathbb{N}$ esiste $t_k \in (\alpha, \omega)$, con $t_k > t_{k-1}$ e $\mathbf{x}(t_k) \notin B_k$. Altrimenti, a partire da qualche $T \in (\alpha, \omega)$ l'orbita sarebbe confinata in uno dei compatti B_k e per (a) dovrebbe essere $\omega = +\infty$. Allora, dato che $\{||\mathbf{x}(t_k)||\}_{k\in\mathbb{N}}$ non è limitato e tenendo conto del fatto che $t_k - t_0 < \omega - t_0 < +\infty$, concludiamo che il rapporto incrementale della prima componente di $\mathbf{x} = \mathbf{x}(t)$ (della parte reale di essa se $\mathbb{K} = \mathbb{C}$), $(x^1(t_k) - x^1(t_0))/(t_k - t_0)$ deve divergere per $k \to +\infty$. Per il teorema di Lagrange esiste una successione di punti τ_k, con $t_0 \le \tau_k \le t_k$, per cui l'insieme $\{||\frac{d\mathbf{x}}{dt}(\tau_k)||\}_{k\in\mathbb{N}}$ non è limitato. Direttamente dall'equazione differenziale, $d\mathbf{x}/dt = \mathbf{f}(t, \mathbf{x}(t))$, concludiamo che $\mathbf{f}$ non può essere limitata. Questo contraddice le ipotesi, per cui $\omega = +\infty$. Con una dimostrazione analoga, usando (b), si prova che $\alpha = -\infty$.
(d) Indichiamo con $supp_D\mathbf{f}$ il supporto di $\mathbf{f}$ nella topologia di D, cioé la chiusura dell'insieme $\{\mathbf{x} \in D \mid \mathbf{f}(\mathbf{x}) \neq 0\}$, rispetto alla topologia relativa di D. Per definizione quindi $supp_D\mathbf{f} \subset D$ e $supp_D\mathbf{f}$ è compatto[5]. Se $\mathbf{f}(\mathbf{x}_0) = 0$, una soluzione, *e quindi l'unica*, del sistema considerato è $\mathbf{x}(t) = \mathbf{x}_0$ per $t \in \mathbb{R}$. Questa soluzione è completa per costruzione. Se $\mathbf{f}(\mathbf{x}_0) \neq 0$, l'orbita non può comunque entrare, per qualche punto $t_1 \neq t_0$, nell'insieme $\{\mathbf{x} \in D \mid \mathbf{f}(\mathbf{x}) = 0\}$ per lo stesso motivo. In questo caso l'orbita è sicuramente tutta confinata in $\{\mathbf{x} \in D \mid \mathbf{f}(\mathbf{x}) \neq 0\}$ e quindi nel compatto $supp_D\mathbf{f} \subset D$. Allora per (a) e (b) la soluzione è completa. $\square$

[5] La compattezza non dipende dalla topologia relativa, per cui se $supp_D\mathbf{f}$ è compatto rispetto a D lo è rispetto a $\mathbb{K}^n$.

3.4 *Confronto tra equazioni differenziali, dipendenza dalle condizioni iniziali e da parametri.

In questa sezione ci occuperemo di alcune questioni tecniche riguardanti l'andamento qualitativo delle soluzioni dei problemi di Cauchy. Essenzialmente ci occupiamo di capire cosa succede, dal punto di vista matematico, alle soluzioni quando si cambia "qualcosa" nel problema di Cauchy in funzione del cambiamento apportato. Possiamo cambiare due cose: la funzione $\mathbf{f}$ oppure la condizione iniziale (eventualmente entrambe).

3.4.1 Lemma di Gronwall e sue conseguenze.

Il seguente risultato molto tecnico, benché elementare ed apparentemente inutile, è invece di fondamentale importanza nella teoria dei sistemi di equazioni differenziali. Il lemma di Gronwall ha infatti importanti conseguenze nello studio dell'andamento *nel tempo* delle soluzioni di un problema di Cauchy (con funzione $\mathbf{f}$ lipschitziana), quando si parte da condizioni inziali differenti, ma "vicine". Si vede allora che le due orbite si allontanano secondo una certa legge esponenziale il cui andamento è fissato dalla costante di Lipschitz. Una seconda conseguenza si ha invece quando si considerano due soluzioni di due *differenti* problemi di Cauchy imponendo la *stessa* condizione iniziale. Anche in questo caso, al variare del tempo le orbite si allontanano seguendo una legge esponenziale regolata dalla costante di Lipschitz.

Teorema 3.6. (***Lemma di Gronwall**) *Siano* $h : [t_0, t_1] \to [0, +\infty)$ *e* $u : [t_0, t_1] \to [0, +\infty)$ *(con* $t_1 > t_0$*) due funzioni continue. Se, per qualche scelta di* $a, b \in (0, +\infty)$*, vale*

$$h(t) \leq a + b \int_{t_0}^{t} h(s)u(s)ds \,, \quad \textit{per ogni } t \in [t_0, t_1] \tag{3.42}$$

allora vale anche

$$h(t) \leq a e^{b \int_{t_0}^{t} u(s)ds} \,, \quad \textit{per ogni } t \in [t_0, t_1] \,. \tag{3.43}$$

$\diamondsuit$

***Dimostrazione**. Posto

$$V(t) := \int_{t_0}^{t} h(s)u(s)ds \,,$$

la (3.42) implica che

$$\frac{dV}{dt} \leq u(t)(a + bV(t)) \,,$$

ovvero, equivalentemente

$$\frac{d}{dt}\left(\ln\left|\frac{a + bV(t)}{a}\right| - b\int_{t_0}^{t} u(s)ds\right) \leq 0 \,.$$

La funzione tra parentesi vale 0 per $t = t_0$, per cui è non positiva se $t > t_0$. Di conseguenza

$$\ln\left|\frac{a+bV(t)}{a}\right| \leq b\int_{t_0}^{t} u(s)ds\ ,$$

e quindi (tenendo conto che l'argomento del logaritmo è non negativo per costruzione)

$$a + bV(t) \leq ae^{b\int_{t_0}^{t} u(s)ds}\ ,$$

ossia, da (3.42),

$$h(t) \leq ae^{b\int_{t_0}^{t} u(s)ds}\ .$$

Questo completa la dimostrazione. □

Il lemma di Gronwall produce il seguente notevole risultato.

Teorema 3.7. *Sia $\Omega \subset \mathbb{R}\times\mathbb{K}^n$ un aperto non vuoto e $\mathbf{f},\mathbf{g} : \Omega \to \mathbb{K}^n$ due funzioni continue e localmente lipschitziane rispetto alla variabile $\mathbf{x} \in \mathbb{K}^n$. Si supponga che ulteriormente:*

(i) $\mathbf{f}$ sia lipschitziana su tutto Ω rispetto a $\mathbf{x}$ ed alla costante K;

(ii) valga $||\mathbf{f}(t,\mathbf{x}) - \mathbf{g}(t,\mathbf{x})|| \leq \epsilon$, per qualche $\epsilon \geq 0$, per ogni $(t,\mathbf{x}) \in \Omega$.

Se $\mathbf{x} : I \to \mathbb{K}^n$ e $\mathbf{y} : J \to \mathbb{K}^n$ sono soluzioni di problemi di Cauchy rispettivamente individuati da $\mathbf{f}$, con condizione iniziale $\mathbf{x}(t_0) = \mathbf{x}_0$, e $\mathbf{g}$, con condizione iniziale $\mathbf{y}(\tau_0) = \mathbf{y}_0$, e si restringono sullo stesso intervallo $(\alpha,\beta) \ni t_0, \tau_0$, allora vale la stima:

$$||\mathbf{x}(t) - \mathbf{y}(t)|| \leq (||\mathbf{x}_0 - \mathbf{y}_0|| + M|t_0 - \tau_0| + \epsilon|\alpha - \beta|)\, e^{K|t-t_0|}\ , \quad \textit{per ogni } t \in (\alpha,\beta), \tag{3.44}$$

avendo definito

$$M := \max\{||\mathbf{g}(t,\mathbf{y}(t))|| \ \mid\ t \in [\alpha,\beta]\}\ .$$

♢

***Dimostrazione**. Dimostriamo la tesi per $t \in [t_0,\beta)$, nell'altro semi intervallo la dimostrazione è del tutto analoga. Vale

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s,\mathbf{x}(s))\, ds \quad \text{e} \quad \mathbf{y}(t) = \mathbf{y}_0 + \int_{\tau_0}^{t} \mathbf{g}(s,\mathbf{y}(s))\, ds\ .$$

Sottranedo membro a membro si trova:

$$\mathbf{x}(t) - \mathbf{y}(t) = \mathbf{x}_0 - \mathbf{y}_0 + \int_{t_0}^{t} [\mathbf{f}(s,\mathbf{x}(s)) - \mathbf{g}(s,\mathbf{y}(s))]\, ds + \int_{t_0}^{\tau_0} \mathbf{g}(s,\mathbf{y}(s))\, ds\ ,$$

che si può riscrivere, sottraendo e aggiungendo uno stesso integrale:

$$\mathbf{x}(t) - \mathbf{y}(t) = \mathbf{x}_0 - \mathbf{y}_0 + \int_{t_0}^{t} [\mathbf{f}(s,\mathbf{x}(s)) - \mathbf{f}(s,\mathbf{y}(s))]\, ds$$

$$+\int_{t_0}^{t}[\mathbf{f}(s,\mathbf{y}(s))-\mathbf{g}(s,\mathbf{y}(s))]\,ds+\int_{t_0}^{\tau_0}\mathbf{g}(s,\mathbf{y}(s))\,ds\,.$$

Nelle nostre ipotesi, sfruttando la disuguaglianza triangolare, note stime integrali e la disuguaglianza che definisce la lipschitzianità , si ha di conseguenza che

$$||\mathbf{x}(t)-\mathbf{y}(t)||\leq||\mathbf{x}_0-\mathbf{y}_0||+K\int_{t_0}^{t}||\mathbf{x}(s)-\mathbf{y}(s)||+\epsilon|\beta-\alpha|+M|\tau_0-t_0|\,.$$

Consideriamo ora che, se $\epsilon>0$, vale:

$$||\mathbf{x}_0-\mathbf{y}_0||+M|\tau_0-t_0|+\epsilon|\beta-\alpha|>0\,. \tag{3.45}$$

In tal caso (3.44) è immediata conseguenza del lemma di Gronwell usando $a:=||\mathbf{x}_0-\mathbf{y}_0||+M|\tau_0-t_0|+\epsilon|\beta-\alpha|$, $t_1:=\beta$, $h(t):=||\mathbf{x}(t)-\mathbf{y}(t)||$ e $u(t)=1$ costantemente ed infine, $b:=K$. Se nelle ipotesi del nostro teorema $\epsilon=0$, cioè $||\mathbf{f}(t,\mathbf{x})-\mathbf{g}(t,\mathbf{x})||=0$ per ogni $(t,\mathbf{x})\in\Omega$, allora vale anche $||\mathbf{f}(t,\mathbf{x})-\mathbf{g}(t,\mathbf{x})||\leq\epsilon$ per ogni $\epsilon>0$. In tal caso possiamo sfruttare la dimostrazione appena fatta ottenendo che, per ogni t fissato:

$$||\mathbf{x}(t)-\mathbf{y}(t)||\leq(||\mathbf{x}_0-\mathbf{y}_0||+M|t_0-\tau_0|+\epsilon|\alpha-\beta|)\,e^{K|t-t_0|}$$

per ogni $\epsilon>0$. Dato che questo vale, per ogni fissato t, per ogni $\epsilon>0$, deve anche essere, per tale valore di t:

$$||\mathbf{x}(t)-\mathbf{y}(t)||\leq(||\mathbf{x}_0-\mathbf{y}_0||+M|t_0-\tau_0|)\,e^{K|t-t_0|}\,,$$

per cui abbiamo mostrato la tesi anche per $\epsilon=0$. □

Ci sono diversi casi possibili in cui si può applicare il teorema provato. Esaminiamone due in particolare.
(1) Si considera un unico sistema di equazioni differenziali ($\mathbf{f}=\mathbf{g}$), ma si considerano due distinte condizioni iniziali $\mathbf{x}_0$, $\mathbf{y}_0$ allo stesso $t_0=\tau_0$. Il teorema dice che in tal caso, per $t\in(\alpha,\beta)$, vale la stima:

$$||\mathbf{x}(t)-\mathbf{y}(t)||\leq||\mathbf{x}_0-\mathbf{y}_0||e^{K|t-t_0|}\,.$$

(2) Si considera la stessa condizione iniziale, ma due distinti sistemi di equazioni differenziali, uno relativo a $\mathbf{f}$ e l'altro a $\mathbf{g}$, con $||\mathbf{f}-\mathbf{g}||_\infty\leq\epsilon$ su Ω. Il teorema dice che in tal caso, per $t\in(\alpha,\beta)$ vale la stima:

$$||\mathbf{x}(t)-\mathbf{y}(t)||\leq\epsilon e^{K|t-t_0|}\,.$$

3.4.2 Regolarità della dipendenza dai dati di Cauchy e questioni connesse.

Un problema importante, specie nelle applicazioni fisiche è la regolarità con cui dipendono, dalle condizioni iniziali e da altri eventuali parametri, le soluzioni massimali di un'equazione differenziale. Tale problema è stato studiato a lungo a partire da Poincaré . Citiamo senza dimostrazione il seguente teorema onnicomprensivo.

Teorema 3.8. *Si consideri una funzione* $\mathbf{f} : \Lambda \times \Omega \to \mathbb{K}^n$ *dove:*

(i) $\Lambda \subset \mathbb{K}^l$ *è un aperto non vuoto i cui elementi sono indicati con* λ*;*

(ii) $\Omega \subset \mathbb{R} \times \mathbb{K}^n$ *è un aperto non vuoto i cui elementi sono indicati con* $(t, \mathbf{x})$*;*

(iii) $\mathbf{f}$ *è continua su* $\Lambda \times \Omega$ *e localmente lipschitziana nella variabile* $\mathbf{x} \in \mathbb{K}^n$*.*

Se

$$(\lambda, t_0, \mathbf{x}_0, t) \mapsto \mathbf{x}_{\lambda,t_0,\mathbf{x}_0}(t)\,, \quad \text{con } (\lambda, t_0, \mathbf{x}_0) \in \Lambda \times \Omega, \tag{3.46}$$

indica la soluzione massimale del problema di Cauchy:

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} & = \mathbf{f}\left(\lambda, t, \mathbf{x}(t)\right), \\ \mathbf{x}(t_0) & = \mathbf{x}_0\,, \end{cases} \tag{3.47}$$

dove $t \in I_{\lambda,t_0,\mathbf{x}_0} \subset \mathbb{R}$*, allora valgono i seguenti fatti.*

(a) *L'insieme*

$$\Gamma := \{(t, \lambda, t_0, \mathbf{x}_0) \mid (\lambda, t_0, \mathbf{x}_0) \in \Lambda \times \Omega\,,\ t \in I_{\lambda,t_0,\mathbf{x}_0}\}$$

è un insieme aperto in $\mathbb{R} \times \Lambda \times \Omega$*.*

(b) *La funzione (3.46) da* Γ *in* $\mathbb{K}^n$ *è una funzione continua congiuntamente in tutte le sue variabili.*

(c) *Se* $\mathbf{f}$ *è di classe* $C^k(\Lambda \times \Omega)$*, allora la funzione (3.46), da* Γ *in* $\mathbb{K}^n$*, e la sua derivata prima in* t *per* $(\lambda, t_0, \mathbf{x}_0)$ *fissati, sono funzioni di classe* $C^k(\Gamma)$*.*

(d) *Se* $\mathbf{f}$ *è analitica in tutte le variabili congiuntamente allora è tale la funzione (3.46) da* Γ *in* $\mathbb{K}^n$*.*

3.5 *Problema di Cauchy su varietà differenziabili.

Il problema di Cauchy può essere studiato anche su varietà differenziabili M di classe C^k con $k > 1$. Esso però deve essere riformulato nel linguaggio dei campi vettoriali.

Definizione 3.6. Se $\mathbf{X}$ è un campo vettoriale di classe C^0 (almeno) sulla varietà differenziabile M, una **curva integrale** $\gamma : I \to M$ del campo $\mathbf{X}$, dove $I \subset \mathbb{R}$ è un intervallo aperto, è una curva di classe C^1 (almeno) tale che

$$\dot{\gamma}(t) = \mathbf{X}(\gamma(t))\,, \quad \text{per ogni } t \in I. \tag{3.48}$$

$\diamondsuit$

Al solito $\dot{\gamma}(t)$ denota il vettore tangente a γ nel punto $\gamma(t)$ in cui il parametro vale t (vedi Appendice A). L'identità (3.48) in coordinate locali (U, ϕ) in cui

$$\mathbf{X}(x^1, \ldots, x^n) = \sum_{i=1}^{n} X^i(x^1, \ldots, x^n) \frac{\partial}{\partial x^i}$$

ed in cui la curva γ è parametrizzata come $x^i = x^i(t)$ per $i = 1, 2, \ldots, n$, si esprime come

$$\frac{dx^i}{dt} = X^i(x^1(t), \ldots, x^n(t)) \,, \quad \text{per } i = 1, 2, \ldots, n. \tag{3.49}$$

Questa non è altro che un'equazione differenziale autonoma su $\Omega = \mathbb{R} \times U$.
Il caso in cui il campo $\mathbf{X}$ dipende parametricamente da t può essere inglobato nel caso precedente in vari modi. La via più diretta è quella di passare alla varietà prodotto $M' := \mathbb{R} \times M$, e definendo su essa il campo

$$\mathbf{X}' := \frac{\partial}{\partial u} + \mathbf{X} \,,$$

dove $u \in \mathbb{R}$ è la coordinata naturale di $\mathbb{R}$. Il problema di determinare le curve integrali a $\mathbf{X}'$:

$$\dot{\gamma}'(t) = \mathbf{X}'(\gamma'(t)) \,, \quad \text{per ogni } t \in I, \tag{3.50}$$

si riduce in coordinate locali $u, x^1, \ldots, x^n$ al sistema di equazioni:

$$\frac{du}{dt} = 1 \,, \tag{3.51}$$

$$\frac{dx^i}{dt} = X^i(u(t), x^1(t), \ldots, x^n(t)) \quad \text{per } i = 1, 2, \ldots, n. \tag{3.52}$$

Una maniera meno rigida è quella di fare uso di una struttura di varietà fibrata (vedi l'appendice A) E, in cui la fibra è diffeomorfa a M e la base è l'asse $\mathbb{R}$. Tale struttura, in coordinate locali naturali di E adattate alla fibrazione riproduce il sistema di sopra. La scelta viene solitamente suggerita dal problema fisico concreto. Nel seguito di questa sezione ci restringeremo al caso autonomo.

3.5.1 Problema di Cauchy, esistenza ed unicità globali.

In base a quanto detto precedentemente, il problema di Cauchy, per un campo vettoriale $\mathbf{X}$ almeno continuo sulla varietà differenziabile M di classe C^1 (o superiore), si può pensare come il problema di determinare le curve integrali $\gamma : I \to M$ del campo $\mathbf{X}$, dove $I \subset \mathbb{R}$ è un intervallo aperto del campo $\mathbf{X}$ che soddisfino un'assegnata condizione iniziale $\gamma(t_0) = p_0$:

$$\begin{cases} \dot{\gamma}(t) &= \mathbf{X}(\gamma(t)), \\ \gamma(t_0) &= p_0. \end{cases} \tag{3.53}$$

Definizione 3.7. Sia $\mathbf{X}$ un campo vettoriale almeno continuo sulla varietà differenziabile M. si consideri il problema di Cauchy per $(t_0, p_0) \in \mathbb{R} \times M$:

$$\begin{cases} \dot{\gamma}(t) &= \mathbf{X}(\gamma(t)) \,, \\ \gamma(t_0) &= p_0. \end{cases} \tag{3.54}$$

Una soluzione di (3.54) $\gamma : I \to M$, con $I \ni t_0$ intervallo aperto, è detta **soluzione massimale** se non esiste alcuna altra soluzione di (3.54) $\gamma' : I' \to M$, con $I' \ni t_0$ intervallo aperto, soddisfacente le condizioni:

(i) $I' \supsetneq I$,

(ii) $\gamma(t) = \gamma'(t)$ se $t \in I$.

Una soluzione massimale $\gamma : I \to M$ è detta:
completa, se $I = (-\infty, +\infty)$;
completa nel futuro, se $I = (a, +\infty)$ con $a \in \mathbb{R} \cup \{-\infty\}$;
completa nel passato, se $I = (-\infty, b)$ con $b \in \mathbb{R} \cup \{+\infty\}$.
Il campo $\mathbf{X}$ è detto **completo** quando tutte le soluzioni massimali di (3.54) sono complete. $\diamondsuit$

Analogamente al caso in $\mathbb{K}^n$, vale il seguente teorema di esistenza ed unicità globale.

Teorema 3.9. (**Esistenza ed unicità globale**) *Sia* $\mathbf{X}$ *un campo vettoriale di classe* C^1 *sulla varietà differenziabile* M*. Si consideri il problema di Cauchy per* $(t_0, p_0) \in \mathbb{R} \times M$*:*

$$\begin{cases} \dot{\gamma}(t) &= \mathbf{X}(\gamma(t)) \,, \\ \gamma(t_0) &= p_0. \end{cases} \tag{3.55}$$

Esiste ed è unica la soluzione massimale $\gamma : I \to M$*.*
Inoltre, se $\gamma_J : J \to M$*, con* $J \ni t_0$ *intervallo aperto, è un'altra soluzione di (3.55), allora*

(i) $J \subset I$*,*

(ii) $\gamma(t) = \gamma_J(t)$ *se* $t \in J$*.* $\diamondsuit$

Dimostrazione. Per prima cosa si osservi che in ogni sistema di coordinate locali, il problema di Cauchy scritto sopra si riduce ad un problema di Cauchy autonomo su un aperto di $\mathbb{R} \times \mathbb{R}^n$. Essendo il campo di classe C^1, siamo nelle ipotesi di locale lipschitzianità e pertanto, in ogni sistema locale di coordinate la soluzione locale esiste ed è unica. Ogni punto raggiunto da γ cade in un sistema di coordinate locali per definizione di varietà stessa. Per questo motivo il lemma 3.1 continua a valere, con la stessa dimostrazione, anche rimpiazzando Ω con $\mathbb{R} \times M$. Tenuto conto di ciò , la dimostrazione è identica a quella data per il teorema 3.5. $\square$

3.5.2 Completezza delle soluzioni.

Possiamo anche dare qualche risultato sulla completezza delle soluzioni. Per prima cosa enunciamo un risultato preliminare.

Proposizione 3.6. *Nelle ipotesi su* $\mathbf{X}$ *del teorema 3.9, sia* $\gamma : I \to M$*, con* $I = (\alpha, \omega)$*, una soluzione massimale del problema di Cauchy (3.55) .*
(a) *Se* γ *non è completa nel futuro, non può esistere alcuna successione* $\{t_k\}_{k \in \mathbb{N}} \subset I$ *tale che* $t_k \to \omega$ *per* $k \to +\infty$ *ed esiste* $\lim_{k \to +\infty} \gamma(t_k) \in M$*.*

(b) *Se γ non è completa nel passato, non può esistere alcuna successione $\{t_k\}_{k\in\mathbb{N}} \subset I$ tale che $t_k \to \alpha$ per $k \to +\infty$ ed esiste $\lim_{k\to+\infty} \gamma(t_k) \in M$.*
$\diamondsuit$

Dimostrazione. La dimostrazione è la stessa data per la proposizione 3.4, riducendosi ad un sistema di coordinate locali. $\square$

La proposizione 3.6 ha delle importanti conseguenze sulla completezza delle soluzioni massimali di un problema di Cauchy quando valgono alcune condizioni aggiuntive. Elenchiamo tali conseguenze in un'unica proposizione.

Proposizione 3.7. *Nelle ipotesi su $\mathbf{X}$ date nel teorema 3.9, se $\gamma : I \to M$, con $I = (\alpha, \omega)$, una soluzione massimale del problema di Cauchy (3.55) valgono i seguenti fatti.*
(a) *Se l'orbita di $\gamma : I \to M$ è contenuta in un compatto $K \subset M$ per $t \geq T \in I$, allora $\gamma : I \to M$ è completa nel futuro.*
(b) *Se l'orbita di $\mathbf{x} : I \to M$ è contenuta in un compatto $K \subset M$ per $t \leq T \in I$, allora $\mathbf{x} : I \to M$ è completa nel passato.*
(c) *Se M è compatta o, più debolmente, $\mathbf{X}$ ha supporto compatto, allora ogni soluzione massimale $\mathbf{x} : I \to M$ è completa, cioé $\mathbf{X}$ è completo.* $\diamondsuit$

Dimostrazione.
(a) Supponiamo che la soluzione non sia completa nel futuro, per cui $I = (\alpha, \omega)$ con $\omega \in \mathbb{R}$. Consideriamo una successione $\{t_k\}_{k\in\mathbb{N}} \subset (T, \omega)$ con $t_k \to \omega$ per $k \to +\infty$. Per ogni $p \in M$, sia (U, ϕ) una carta locale con $U \ni p$. Per ogni U prendiamo un aperto $V \ni p$ tale che $\overline{V}$ sia compatto e sia incluso completamente in U. La classe delle carte locali $(V, \phi \restriction_V)$ definisce, per costruzione, un atlante di M. Dato che K è compatto ed è ricoperto dalla classe degli aperti V suddetti, possiamo estrarre, da tale classe, un sottoricoprimento finito. Ci sarà quindi una classe finita carte locali $\{(U_i, \phi_i)\}_{i=1,\ldots,N}$ ed una corrispondente classe finita di compatti $\overline{V_i}$, tali che $\cup_{i=1}^N \overline{V_i} \supset K$ e $\overline{V_i} \subset U_i$. Dato che i punti $\gamma(t_k)$ sono infiniti, almeno uno dei compatti, diciamo $\overline{V_j}$, ne dovrà contenere una quantità infinita. Ci sarà dunque una sottosuccessione $\{\gamma(t_{k_r})\}_{r\in\mathbb{N}} \subset \overline{V_j}$. $\overline{V_j}$ si può identificare con un compatto di $\mathbb{R}^n$ dato che $\overline{V_j} \subset U_j$ e U_j è identificato con un aperto di $\mathbb{R}^n$ da un omeomorfismo ϕ_j. Per la proprietà di sequenziale compattezza dei compatti di $\mathbb{R}^n$, deve esistere una sotto successione di $\{\gamma(t_{k_r})\}_{r\in\mathbb{N}}$, che indichiamo con $\{\gamma(t_{k_{r_s}})\}_{s\in\mathbb{N}}$, che ammette limite nel compatto $\overline{V_j}$. Ma questo significa che, posto $\tau_s := t_{k_{r_s}}$, la successione $\{\tau_s\}_{s\in\mathbb{N}}$ soddisfa $\tau_s \to \omega$ per $s \to +\infty$ ed esiste il limite, in M, di $\gamma(\tau_s)$. Questo è vietato dalla proposizione 3.6. Di conseguenza γ deve essere completa nel futuro. La dimostrazione di (a) è essenzialmente identica con ovvie modifiche.
(c) Se $\mathbf{X}$ ha supporto compatto, ogni curva integrale massimale di $\mathbf{X}$ che ha almeno un punto fuori dal supporto di $\mathbf{X}$ deve essere una curva costante per il teorema di unicità. In tal caso è completa per costruzione. Le rimanenti curve integrali hanno, per definizione, orbita inclusa in un compatto: il supporto di $\mathbf{X}$. Per (a) e (b) anche tali soluzioni sono complete. $\square$

Lavorando in carte locali, il teorema 3.8 può essere esteso facilmente anche nel caso di equazioni differenziali su varietà giungendo al seguente risultato:

Teorema 3.10. *Si consideri una classe di campi vettoriali* $\{\mathbf{X}_\lambda\}_{\lambda\in\Lambda}$ *sulla varietà differenziabile* M*, dove* $\Lambda \subset \mathbb{R}^l$ *è un aperto non vuoto e la funzione* $\mathbf{X} : (\lambda, x) \mapsto \mathbf{X}_\lambda(x)$ *è congiuntamente continua su* $\Lambda \times M$ *e di classe* $C^1(M)$ *per ogni fissato* $\lambda \in \Lambda$*. Se*

$$(\lambda, x_0, t) \mapsto \mathbf{x}_{\lambda,x_0}(t)\,, \quad \text{con } (\lambda, x_0) \in \Lambda \times M, \tag{3.56}$$

indica la soluzione massimale del problema di Cauchy:

$$\begin{cases} \dot\gamma(t) &= \mathbf{X}_\lambda(\gamma(t))\,, \\ \gamma(0) &= x_0. \end{cases} \tag{3.57}$$

dove $t \in I_{\lambda,x} \subset \mathbb{R}$*, allora valgono i seguenti fatti.*
(a) *L'insieme*

$$\Gamma := \{(t, \lambda, x_0) \mid (\lambda, x) \in \Lambda \times M\,,\ t \in I_{\lambda,x_0}\}$$

è un insieme aperto in $\mathbb{R} \times \Lambda \times M$*.*
(b) *La funzione (3.56) da* Γ *in* M *è una funzione continua congiuntamente in tutte le sue variabili (in riferimento a un atlante su* M*).*
(c) *Se* $\mathbf{X}$ *è di classe* $C^k(\Lambda \times M)$*, allora la funzione (3.56), e la sua derivata in* t *a* (λ, x_0) *fissati sono funzioni di classe* $C^k(\Gamma)$ *congiuntamente in tutte le variabili.* $\diamondsuit$

3.5.3 Gruppi di diffeomorfismi locali ad un parametro.

Consideriamo una varietà differenziabile M (per esempio $\mathbb{R}^n$) ed un campo vettoriale $\mathbf{X}$ di classe C^k, con $k \geq 1$, su di essa. Consideriamo infine, per ogni $p \in M$ la soluzione massimale $I_p \ni t \mapsto \gamma_p(t)$ del problema di Cauchy (3.55) dove $p_0 = p$ e $t_0 = 0$. Consideriamo l'applicazione

$$(t, p) \mapsto \gamma_p(t) =: \phi_t^{(\mathbf{X})}(p)$$

dove $p \in M$ e $t \in I_p$. In base al teorema 3.10 l'insieme in cui variano le coppie (t, p) è aperto, inoltre in base al teorema 3.10 e per definizione di soluzione massimale, valgono le due seguenti proprietà di immediata verifica:

(i) l'applicazione $(t, p) \mapsto \phi_t^{(\mathbf{X})}(p)$ e la sua derivata in t (a p fissato) sono ciascuna di classe C^k congiuntamente nelle due variabili t e p,

(ii) $\phi_0^{(\mathbf{X})}(p) = p$ per ogni $p \in M$,

(iii) $\phi_t^{(\mathbf{X})}(\phi_{t'}^{(\mathbf{X})}(p)) = \phi_{t+t'}^{(\mathbf{X})}(p)$ per ogni $p \in M$ e per $t, t' \in I_p$ tali che $t + t' \in I_p$.

Osservazioni 3.6.
(1) Per un *fissato* $p \in M$, la proprietà (iii) vale sicuramente, se $|t|$ e $|t'|$ sono sufficientemente piccoli dato che la soluzione massimale dell'equazione differenziale definente $\phi^{(\mathbf{X})}$ con condizione

iniziale p a $t = 0$ è sicuramente definita in un intervallo aperto I_p non vuoto.
(2) La condizione (iii) vale per *ogni* coppia $t, t' \in \mathbb{R}$ quando le soluzioni massimali di (3.55) sono *complete*, e ciò accade in particolare se $\mathbf{X}$ ha supporto compatto e quindi tutte le volte in cui M è compatta per (c) del teorema 3.6.
(3) Se $t, -t \in I_p$, (i) e (ii) implicano che $\phi_t^{(\mathbf{X})}(\phi_{-t}^{(\mathbf{X})}(p)) = p$. Non possiamo però concludere immediatamente che $\phi_{-t}^{(\mathbf{X})} = (\phi_t^{(\mathbf{X})})^{-1}$, dato che non è detto che variando l'argomento p, il valore di t possa tenersi costante: t deve essere preso all'interno del dominio I_p della soluzione massimale dell'equazione differenziale definente $\phi^{(\mathbf{X})}$ con condizione iniziale p a $t = 0$ e I_p può dipendere da p. Approfondiamo la situazione. Per un fissato $q \in M$, dato che la funzione $(t, p) \mapsto \phi_t^{(\mathbf{X})}(p)$ è continua congiuntamente nelle due variabili che variano in un aperto, la controimmagine di un intorno aperto V_q di $q = \phi_0^{(\mathbf{X})}(q)$ deve essere un intorno aperto F di $(0, q)$. F conterrà un aperto di forma $(-\epsilon, \epsilon) \times U_q$, dove U_q è un secondo intorno aperto di q, se scegliamo $\epsilon > 0$ e U_q sufficientemente piccoli. Per questa via concludiamo che, posto $J_q := (-\epsilon/2, \epsilon/2)$, se $t, t' \in J_q$ allora $t, t', t + t' \in (-\epsilon, \epsilon)$ e quindi $\phi_t^{(\mathbf{X})}(p)$, $\phi_{t'}^{(\mathbf{X})}(p)$ $\phi_{t+t'}^{(\mathbf{X})}(p)$ sono ben definiti se $p \in U_q$. Riassumendo abbiamo ottenuto che ogni punto $q \in M$ ammette un suo intorno aperto $U_q \subset M$ ed un intervallo aperto $J_q \ni 0$ tali che, la proprietà (iii) è valida per ogni $p \in U_q$ ed ogni coppia $t, t' \in J_q$. Si osservi che, con la stessa scelta di J_q e U_q, se $t \in J_q$ sia $\phi_t^{(\mathbf{X})}$ che $\phi_{-t}^{(\mathbf{X})}$ (e ovviamente $\phi_{t-t}^{(\mathbf{X})}$)) sono ben definiti. Possiamo finalmente concludere che le funzioni $\phi_t^{(\mathbf{X})}$ definiscono dei *diffeomorfismi locali* su M: per ogni $q \in M$ e prendendo un qualsiasi $t \in J_q$, essi identificano diffeomorficamente un intorno aperto di q, $U_q \subset M$, con un corrispondente aperto, dipendente da t, $V_{q,t} := \phi_t^{(\mathbf{X})}(U_q)$. Restringendo le applicazioni a questa coppia di intorni aperti vale effettivamente $\phi_{-t}^{(\mathbf{X})} \restriction_{V_{q,t}} = (\phi_t^{(\mathbf{X})} \restriction_{U_q})^{-1}$. Inoltre, scegliendo $t, t' \in J_q$ la composizione $\phi_t^{(\mathbf{X})} \circ \phi_{t'}^{(\mathbf{X})}$ è ben definita su U_q e produce $\phi_{t+t'}^{(\mathbf{X})}$.

Mostriamo che la procedura per costruire l'insieme di diffeomorfismi locali $\phi_t^{(\mathbf{X})}$ a partire da un campo vettoriale $\mathbf{X}$ può essere invertita.

Teorema 3.11. *Si consideri una varietà differenziabile M dotata di un'applicazione della forma $A \ni (t, x) \mapsto \phi_t(x) \subset M$, dove $A \subset \mathbb{R} \times M$ è un aperto non vuoto tale che, se $(t_x, x) \in A$, allora l'insieme dei numeri t con $(t, x) \in A$ è un intervallo aperto che include lo zero. Si assuma che ϕ soddisfi le seguenti proprietà :*

(i) l'applicazione $(t, p) \mapsto \phi_t(p)$ e la sua derivata in t (a p fissato) sono ciascuna di classe C^k ($k \geq 1$) congiuntamente nelle due variabili t e p,

(ii) $\phi_0(p) = p$ per ogni $p \in M$,

(iii) $\phi_t(\phi_{t'}(p)) = \phi_{t+t'}(p)$ per ogni $p \in M$ e per $t, t' \in \mathbb{R}$ tali che entrambi i membri siano definiti.

In questo caso esiste un unico campo vettoriale $\mathbf{X}$ su M tale che, sul dominio di ϕ vale: $\phi = \phi^{(\mathbf{X})}$. $\mathbf{X}$ è di classe C^k $\diamondsuit$

Dimostrazione. Si consideri il campo vettoriale $\mathbf{X}(x) := \frac{d}{dt}|_{t=0}\phi_t(x)$. Questo è C^k da (i).

Dalle proprietà (ii) e (iii) segue subito che, dove il primo membro è definito, $\frac{d}{dt}\phi_t(x) = \mathbf{X}(\phi_t(x))$, ma allora, dal teorema di unicità globale 3.9, $t \mapsto \phi_t(x)$ è una restrizione della soluzione massimale dell'equazione differenziale $\frac{d}{dt}\gamma(t) = \mathbf{X}(\gamma(t))$ con condizione iniziale $\gamma(0) = x$. Pertanto $\phi_t(x) = \phi_t^{(\mathbf{X})}(x)$ ogni qual volta il primo membro è definito. Il fatto che $\mathbf{X}$ sia unico segue immediatamente dal fatto che se $\phi_t(p) = \phi_t^{(\mathbf{X})}(p)$ per t definito in un intorno di 0, allora $\mathbf{X}(x) := \frac{d}{dt}|_{t=0}\phi_t(x)$. $\square$

Definizione 3.8. Una classe di di trasformazioni ϕ_t sulla varietà differenziabile M, che soddisfa le proprietà (i), (ii) e (iii) del teorema 3.11 è detta **gruppo ad un parametro di diffeomorfismi locali** (di classe k) su M. L'unico campo vettoriale $\mathbf{X}$ su M tale che, sul dominio di ϕ vale $\phi = \phi^{(\mathbf{X})}$, è detto **generatore** del gruppo ad un parametro di diffeomorfismi locali.
Un gruppo ad un parametro di diffeomorfismi locali è detto **gruppo ad un parametro di diffeomorfismi** se la variabile t è definita su tutto $\mathbb{R}$ indipendentemente dall'argomento p di $\phi_t(p)$. $\diamondsuit$

Osservazioni 3.7.
(1) Ogni applicazione $\phi_t : M \to M$ di un *gruppo ad un parametro di diffeomorfismi* definisce un diffeomorfismo (globale!) di M in M, inoltre l'insieme delle funzioni $\{\phi_t\}_{t\in\mathbb{R}}$ costituisce effettivamente un gruppo commutativo (ossia abeliano), usando come legge di composizione l'ordinaria composizione di funzioni.
(2) Evidentemente un gruppo ad un parametro di diffeomorfismi *locali* sulla varietà differenziabile M si può estendere in un gruppo ad un parametro di diffeomorfismi (globali!) se e solo se il generatore $\mathbf{X}$ è completo. E ciò accade in particolare se il supporto di $\mathbf{X}$ è compatto, oppure più fortemente, se M è compatta.

3.5.4 Esistenza di integrali primi funzionalmente indipendenti.

La nozione di *integrale primo* $F : M \to \mathbb{R}$ rispetto ad un campo vettoriale differenziabile $\mathbf{X}$ sulla varietà differenziabile M è la banale generalizzazione della definizione data su $\mathbb{K}^n$:

Definizione 3.9. **(Integrale primo su M.)** Sia $\mathbf{X}$ un campo vettoriale di classe C^1 sulla varietà differenziabile M di classe C^k, con $k \geq 1$. Una funzione $F \in C^1(M)$ è detta integrale primo di $\mathbf{X}$ quando F è costante se ristretta alle curve integrali del campo $\mathbf{X}$ (il valore della costante dipende in generale dalla curva integrale). $\diamondsuit$

Si ha immediatamente il seguente risultato.

Proposizione 3.8. *Sia $\mathbf{X}$ un campo vettoriale di classe C^1 sulla varietà differenziabile M di classe C^k, con $k \geq 1$. Una funzione $F \in C^1(M)$ è un integrale primo di $\mathbf{X}$ se e solo se*

$$\mathbf{X}(F) = 0\,, \quad \textit{ovunque su } M\,.$$

Se $F \in C^1(M)$ è integrale primo per $\mathbf{X}$, allora ogni insieme

$$E_{\alpha,\beta} := \{x \in M \mid \alpha \leq F \leq \beta\}$$

(ed in particolare ogni insieme di livello di F) è **invariante** *sotto $\mathbf{X}$, cioé :*

$$\phi_t^{(\mathbf{X})}(\mathbf{x}) \in E_{\alpha,\beta} \quad \textit{se } x \in E_{\alpha,\beta}, \textit{ ogni volta che } \phi_t^{(\mathbf{X})}(\mathbf{x}) \textit{ è definito.}$$

$\diamondsuit$

Dimostrazione. La dimostrazione della prima parte, lavorando in coordinate locali, è la stessa data per il teorema 3.4. La seconda parte è ovvia per definizione di integrale primo. $\square$

La seguente proposizione tecnica è di grande utilità in varie costruzioni.

Proposizione 3.9. *Sia $\mathbf{X}$ un campo vettoriale di classe C^k sulla varietà differenziabile M di ordine di differenziabilità k. Sia $p \in M$ tale che $\mathbf{X}(p) \neq 0$, allora esiste una carta locale (U, ϕ) con $p \in U$ e dove $\phi : U \ni q \mapsto (y^1(q), \ldots, y^n(q))$, in cui ogni curva integrale di $\mathbf{X}$ espressa in coordinate è del tipo: $I \ni t \mapsto (y^1 + t, y^2, \ldots, y^n)$ per un corrispondente intervallo aperto $I \subset \mathbb{R}$ e pertanto:*

$$\mathbf{X}\restriction_U = \frac{\partial}{\partial y^1} \, .$$

$\diamondsuit$

Dimostrazione. Consideriamo una carta locale attorno a p con coordinate $(x^1, \ldots, x^n) \in V \subset \mathbb{R}^n$ in modo tale che p corrisponda all'origine di $\mathbb{R}^n$. Dato che $\mathbf{X}(p) \neq 0$, almeno una componente di

$$\mathbf{X}(p) = \sum_{i=1}^{n} X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p$$

deve essere non nulla. Senza perdere generalità , eventualmente cambiando i nomi delle coordinate, assumaremo che $X^1(p) \neq 0$. Consideriamo le curve integrali di $\mathbf{X}$ espresse da, nelle coordinate dette, in un intorno sufficientemente piccolo di p: $x^i = x^i(t, x_0^1, \ldots, x_0^n)$ dove $(x_0^1, \ldots, x_0^n)$ sono le coordinate del punto iniziale da cui esce la curva a $t = 0$ e $t \in (-\delta, \delta)$ con $\delta > 0$ sufficientemente piccolo. Questa funzione è ben definita ed è di classe C^1 per il teorema 3.10. Consideriamo infine la trasformazione definita nell'intorno dell'origine di $\mathbb{R}^n$:

$$(y^1, y^2, \ldots y^n) \mapsto x^i(y^1, 0, y^2, \ldots y^n) \, , \quad \text{per } i = 1, \ldots, n.$$

Tenendo conto del fatto che, per costruzione:

$$x^i(t = 0, x_0^1, \ldots, x_0^n) = x_0^i \, , \quad \text{per } i = 1, \ldots, n,$$

mentre

$$\left.\frac{\partial x^i(t, x_0^1, \ldots, x_0^n)}{\partial t}\right|_{t=0} = X^i(p) \, , \quad \text{per } i = 1, \ldots, n,$$

concludiamo che la matrice jacobiana di questa trasformazione calcolata in $(t, y^2, \dots y^n) = (0, 0, \dots, 0)$ è costituita come segue:

(i) sulla prima riga appare il vettore riga $(X^1(p), 0, \dots, 0)$,

(ii) sulla prima colonna appare il vettore colonna $(X^1(p), X^2(p), \dots, X^n(p))^t$,

(iii) nella rimanente matrice quadrata $(n-1) \times (n-1)$, ottenuta sopprimendo la prima riga e la prima colonna, appare la matrice identità I. Sviluppando il determinante rispetto alla prima riga, concludiamo che esso vale $X^1(p) \neq 0$. Ma allora, per il teorema del Dini [GiustiII], la trasformazione

$$(y^1, y^2, \dots y^n) \mapsto x^i(y^1, 0, y^2, \dots y^n)\,, \quad \text{per } i = 1, \dots, n.$$

è un diffeomorfismo locale di classe C^k nell'intorno dell'origine di $\mathbb{R}^n$. In definitiva, in un intorno aperto U di p sufficientemente piccolo, le funzioni $y^i = y^i(x^1, \dots, x^n)$ definiscono un sistema di coordinate locali ammissibili su M attorno a p. In questo sistema di coordinate, per costruzione, le curve integrali di $\mathbf{X}$ hanno la forma $I \ni t \mapsto (y^1 + t, y^2, \dots, y^n)$. $\square$

La proposizione provata ha l'immediata conseguenza concernente l'esistenza locale di $(dim M) - 1$ integrali primi funzionalmente indipendenti per ogni campo vettoriale $C^1(M)$ non nullo.

Teorema 3.12. **(Esistenza locale di $n-1$ integrali primi indipendenti.)** *Sia* $\mathbf{X}$ *un campo vettoriale di classe* C^1 *sulla varietà differenziabile* M *di dimensione* n *e ordine di differenziabilità* k*, con* $k \geq 1$*. Sia* $p \in M$ *tale che* $\mathbf{X}(p) \neq 0$*, allora esistono una carta locale* (U, ϕ) *con* $p \in U$ *e* $n-1$ *funzioni* $F_k : U \to \mathbb{R}$ *di classe* $C^1(M)$ *che soddisfano:*

(a) *sono integrali primi per* $\mathbf{X}$*,*

(b) *sono* **funzionalmente indipendenti***, cioé la matrice jacobiana delle* $n-1$ *funzioni* F_k *calcolata rispetto alle coordinate locali di* U *ha rango massimo possibile:* $n-1$*.* $\diamondsuit$

Dimostrazione. Usando il sistema di coordinate (U, ϕ) della proposizione 3.9 (eventualmente passando ad un atlante di classe C^1 se $k > 1$), le $n-1$ funzioni $F_k : U \ni q \mapsto y^k(q)$ soddisfano banalmente i due requisiti detti. $\square$

Osservazioni 3.8.

(1) La richiesta di funzionale indipendenza non dipende dalle coordinate usate e può per esempio esprimersi dicendo che, per ogni $q \in U$, gli $n-1$ vettori $dF_k(q)$ sono linearmente indipendenti. Lasciamo la prova per esercizio al lettore.

(2) Il risultato sull'esistenza degli $n-1$ integrali primi ha validità locale e si basa fortemente sull'ipotesi che $\mathbf{X}$ non si annulli. L'insieme di tutti i punti $p \in M$ in cui $\mathbf{X}(p) = 0$ è invariante sotto il gruppo ad un parametro di diffeomorfismi locali associato a $\mathbf{X}$. Per alcuni punti isolati p di tale insieme accade che verso di essi convergono o divergono contemporaneamente più linee integrale di $\mathbf{X}$ con apparente violazione del teorema di unicità delle soluzioni delle linee integrali di $\mathbf{X}$. Per esempio ciò accade per p dato dall'origine di $M := \mathbb{R}^2$, rispetto al campo $\mathbf{X}(x, y) = x\,\mathbf{e}_x + y\,\mathbf{e}_y$. Tuttavia nessuna di queste linee raggiunge p dato che ciò accadrebbe "per valori infiniti del parametro".

Capitolo 4

Leggi di bilancio ed integrali primi in Meccanica.

Il fine di questo capitolo è quello di introdurre alcune quantità fisiche relative a sistemi di punti materiali che entrano in gioco in particolari "leggi di bilancio" e leggi di "conservazione". Tali leggi sono in realtà teoremi che seguono dai principi della meccanica classica enunciati nel capitolo 3. Essi riguardano in particolare: la quantità di moto, il momento angolare e l'energia meccanica. Daremo una trattazione comune al caso di punto materiale e di sistemi di punti materiali.
Solitamente, ma ci sono alcune eccezioni, la forma generale dell'enunciato di questi teoremi è un'uguaglianza tra la derivata temporale di una certa grandezza G, funzione delle posizioni e delle velocità dei punti costituenti un sistema fisico ed una seconda grandezza X, normalmente riguardate l'"esterno" del sistema:

$$\frac{dG(P_1(t),\ldots,P_N(t),\mathbf{v}_{P_1}(t),\ldots,\mathbf{v}_{P_N}(t))}{dt} = X(t,P_1(t),\ldots,P_N(t),\mathbf{v}_{P_1}(t),\ldots\mathbf{v}_{P_N}(t),\text{"var. est."})\,.$$

L'uguaglianza sussiste quando il sistema evolve in conformità delle leggi della dinamica C1-C4 enunciate nel capitolo 3. Quando la grandezza a secondo membro è nulla (almeno sul moto considerato), la legge di bilancio diventa una legge di *conservazione* nel tempo della quantità a primo membro *valutata su un fissato moto del sistema fisico.* Dal punto di vista della teoria dei sistemi di equazioni differenziali, le grandezze conservate nel tempo vengono chiamate *integrali primi* (del moto) come stabilito nella definizione 3.1.

4.1 Equazioni cardinali per i sistemi di punti materiali, conservazione dell'impulso e del momento angolare

Introduciamo ora delle equazioni, valide per tutti i sistemi fisici meccanici soddisfacenti i principi C1-C4 (anche nelle versioni generalizzate valide in presenza di reazioni vincolari e forze inerziali) enunciati al cap. 3, costituiti da punti materiali, dette *Equazioni Cardinali* della dinamica.

4.1.1 Massa totale, impulso totale, momento angolare totale, energia cinetica totale

Definizione 4.1. Si consideri un sistema S di N punti materiali P_i, $i = 1, \ldots, N$ con masse rispettivamente m_i, $i = 1, \ldots, N$. Si danno le seguenti definizioni.
(a) La **massa totale del sistema** è:

$$M := \sum_{k=1}^{N} m_k \, .$$

(b) Il **centro di massa** del sistema S al tempo $t \in \mathbb{R}$, $G(t)$ è il punto (non necessariamente un punto materiale del sistema) individuato su ogni Σ_t dall'equazione:

$$M(G(t) - O) = \sum_{k=1}^{N} m_k (P_k(t) - O) \, ,$$

dove $O \in \Sigma_t$ è un punto qualsiasi.
(c) L'**impulso (totale)** o **quantità di moto (totale)** del sistema S rispetto ad un riferimento $\mathscr{I}$ al tempo t è il vettore di V_t:

$$\mathbf{P}|_{\mathscr{I}}(t) := \sum_{k=1}^{N} m_k \mathbf{v}_{P_k}|_{\mathscr{I}}(t) \, .$$

(d) Se $O = O(t)$ è una qualsiasi linea di universo (non necessariamente quella del punto (b)) e $\mathscr{I}$ è un sistema di riferimento, il **momento angolare (totale)** o **momento della quantità di moto (totale)** del sistema rispetto al polo O ed ad un riferimento $\mathscr{I}$ al tempo t è il vettore di V_t:

$$\mathbf{\Gamma}_O|_{\mathscr{I}}(t) := \sum_{k=1}^{N} m_k (P_k(t) - O(t)) \wedge \mathbf{v}_{P_k}|_{\mathscr{I}}(t) \, .$$

$\diamondsuit$

L'importanza delle quantità definite sopra è essenzialmente legata al fatto che, sotto determinate ipotesi e per un fissato sistema meccanica, tali quantità si conservano nel tempo o appaiono nelle espressioni definitorie di quantità che si conservano nel tempo. In molti casi, la conoscenza dei valori di grandezze conservate nel tempo fornisce importanti informazioni sul moto del sistema, anche se non si riesce a risolvere esplicitamente l'equazione del moto.

Osservazioni 4.1.
(1) *D'ora in poi parlando di un sistema di punti materiali, assumeremo sempre che il loro numero N sia finito.*
(2) Ovviamente le definizioni date sopra sono valide anche per $N = 1$. In tal caso la massa

totale coincide con la massa del punto, il centro di massa coincide con la posizione del punto, l'impulso totale coincide con l'impulso del punto, il momento angolare totale e l'energia cinetica totale coincidono , rispettivamente e per definizione, con il momento angolare e l'energia cinetica del punto materiale.
(3) La definizione di G è ben posta, nel senso che G è univocamente determinato, una volta fissato O, da:

$$G := O + \frac{1}{M}\sum_{k=1}^{N} m_k(P_k - O)\,,$$

inoltre G *non dipende dalla scelta di* O. Infatti, se definiamo G_O tramite:

$$M(G_O - O) = +\sum_{k=1}^{N} m_k(P_k - O)\,,$$

e $G_{O'}$ tramite:

$$M(G_{O'} - O') = +\sum_{k=1}^{N} m_k(P_k - O')\,,$$

allora:

$$G_O - G_{O'} = (O - O') + \frac{1}{M}\sum_{k=1}^{N} m_k(P_k - O) - \frac{1}{M}\sum_{k=1}^{N} m_k(P_k - O') = (O - O') - (O - O') = \mathbf{0}\,.$$

(4) Se $\mathbf{v}_G|_{\mathscr{I}}$ è la velocità del centro di massa nel riferimento $\mathscr{I}$ per un sistema di punti materiali di massa totale M, allora vale la relazione

$$\mathbf{P}|_{\mathscr{I}} = M\mathbf{v}_G|_{\mathscr{I}}\,. \tag{4.1}$$

In altre parole: *l'impulso totale del sistema è quello che avrebbe un singolo punto materiale di massa M concentrata nel centro di massa del sistema.*
La verifica di ciò è immediata, scegliendo una linea di universo $O = O(t)$, derivando membro a membro nel tempo l'identità $M(G(t) - O(t)) = \sum_{k=1}^{N} m_k(P_k(t) - O(t))$ e tenendo conto di $M = \sum_k m_k$.

Esercizi 4.1.

1. Dimostrare che, passando dal riferimento $\mathscr{I}$ al riferimento $\mathscr{I}'$, vale la legge di trasformazione del momento della quantità di moto (mantenendo lo stesso polo O):

$$\mathbf{\Gamma}_O|_{\mathscr{I}}(t) = \mathbf{\Gamma}_O|_{\mathscr{I}'}(t) + \mathbf{I}_{t,O(t)}(\boldsymbol{\omega}_{\mathscr{I}'}|\mathscr{I})\,, \tag{4.2}$$

dove abbiamo introdotto il **tensore d'inerzia** al tempo t e rispetto al polo O, dato dalla trasformazione lineare

$$V_t \ni \mathbf{u} \mapsto \mathbf{I}_{t,O(t)}(\mathbf{u}) := \sum_{k=1}^{N} m_k(P_k(t) - O(t)) \wedge (\mathbf{u} \wedge (P_k(t) - O(t)))\,. \tag{4.3}$$

2. Dimostrare che, passando dal polo O al polo O', ma rimanendo nello stesso riferimento $\mathscr{I}$, vale la legge di trasformazione del momento della quantità di moto (al tempo $t \in \mathbb{R}$ fissato):

$$\mathbf{\Gamma}_O|_{\mathscr{I}} = \mathbf{\Gamma}_{O'}|_{\mathscr{I}} + (O' - O) \wedge \mathbf{P}|_{\mathscr{I}} \,. \tag{4.4}$$

In particolare, scegliendo $O' = G$, $\mathbf{\Gamma}_O|_{\mathscr{I}}$ si può sempre scrivere la somma del momento angolare totale in $\mathscr{I}$ rispetto a G e del momento angolare di un unico punto materiale di posizione G avente massa pari alla massa totale del sistema:

$$\mathbf{\Gamma}_O|_{\mathscr{I}} = \mathbf{\Gamma}_G|_{\mathscr{I}} + (G - O) \wedge \mathbf{P}|_{\mathscr{I}} \,. \tag{4.5}$$

4.1.2 Equazioni cardinali.

Possiamo ora enunciare e provare le cosiddette *equazioni cardinali della dinamica* e ne deduciamo le corrispondenti leggi di bilancio/conservazione. Per enunciarli definiamo preventivamente la nozione di *forza interna* e di *forza esterna*. Se abbiamo un sistema S di punti materiali, una forza $\mathbf{F}$ agente $P \in S$ è detta **interna** se la corrispondente reazione agisce su un punto $P' \in S$. Una forza che non è interna è detta forza **esterna**. Le forze inerziali sono sempre considerate forze esterne.

Teorema 4.1. (**Equazioni cardinali della dinamica dei sistemi di punti materiali**.) *Si consideri un sistema di N punti materiali che soddisfa i principi C1-C4 (includendo il caso di forze inerziali e reazioni vincolari). Se $\mathscr{I}$ è un sistema di riferimento, per ogni fissato istante del tempo assoluto t valgono le seguenti, rispettivamente,* **prima e seconda equazione cardinale della dinamica dei sistemi di punti materiali**.
(a) *Se all'istante considerato $\mathbf{F}_i^{(e)}$ è la somma delle forze esterne agenti sull'i-esimo punto (includendo forze inerziali e reazioni vincolari se necessario):*

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{P}|_{\mathscr{I}} = \sum_{i=1}^{N} \mathbf{F}_i^{(e)} \,. \tag{4.6}$$

(b) *Con la stessa notazione di sopra:*

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{\Gamma}_O|_{\mathscr{I}} + \mathbf{v}_O|_{\mathscr{I}} \wedge \mathbf{P}|_{\mathscr{I}} = \sum_{i=1}^{N} (P_i - O) \wedge \mathbf{F}_i^{(e)} \,. \tag{4.7}$$

$\diamondsuit$

Dimostrazione. (a) Il secondo principio della dinamica per il punto P_i si scrive:

$$\sum_j \mathbf{F}_{ij}^{(i)} + \mathbf{F}_i^{(e)} = m_i \frac{d^2 P_i}{dt^2}|_{\mathscr{I}} \,,$$

dove $\mathbf{F}_{ij}$ è la forza agente su P_i dovuta al punto P_j, mentre $\mathbf{F}_i^{(e)}$ è la somma delle forze esterne agenti su P_i (includendo le forze inerziali se $\mathscr{I}$ non è inerziale). Da ciò segue che:

$$\sum_i m_i \frac{d^2 P_i}{dt^2}|_{\mathscr{I}} = \sum_{i,j} \mathbf{F}_{ij} + \sum_i \mathbf{F}_i^{(e)} .$$

Cambiando nome agli indici:

$$\sum_{i,j} \mathbf{F}_{ij}^{(i)} = \sum_{j,i} \mathbf{F}_{ji}^{(i)} .$$

D'altra parte, per il principio di azione e razione: $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$, per cui:

$$\sum_{i,j} \mathbf{F}_{ij}^{(i)} = -\sum_{j,i} \mathbf{F}_{ij}^{(i)} = -\sum_{i,j} \mathbf{F}_{ij} ,$$

dove abbiamo tenuto conto del fatto che l'ordine con cui si esegue la somma (prima per i e poi per j oppure viceversa) è irrilevante. Concludiamo che $\sum_{i,j} \mathbf{F}_{ij}^{(i)} = \mathbf{0}$ e pertanto

$$\sum_i m_i \frac{d^2 P_i}{dt^2}|_{\mathscr{I}} = \sum_i \mathbf{F}_i^{(e)} .$$

cioè :

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \sum_i m_i \mathbf{v}_{P_i}|_{\mathscr{I}} = \sum_i \mathbf{F}_i^{(e)} .$$

Tenendo conto della definizione impulso totale si ha subito la (4.6).
(b) Con la stessa notazione di sopra, dalla definizione di $\mathbf{\Gamma}_O|_{\mathscr{I}}$, si ha:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{\Gamma}_O|_{\mathscr{I}} = \sum_i m_i (\mathbf{v}_{P_i}|_{\mathscr{I}} - \mathbf{v}_O|_{\mathscr{I}}) \wedge \mathbf{v}_{P_i}|_{\mathscr{I}} + \sum_{i,j} (P_i - O) \wedge \mathbf{F}_{ij} + \sum_i (P_i - O) \wedge \mathbf{F}_i^{(e)} .$$

Ovvero:

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} \mathbf{\Gamma}_O|_{\mathscr{I}} + \sum_i m_i \mathbf{v}_O|_{\mathscr{I}} \wedge \mathbf{v}_{P_i}|_{\mathscr{I}} = \sum_i (P_i - O) \wedge \mathbf{F}_i^{(e)} + \sum_{i,j} (P_i - O) \wedge \mathbf{F}_{ij} .$$

Dato che $\sum_i m_i \mathbf{v}_O|_{\mathscr{I}} \wedge \mathbf{v}_{P_i}|_{\mathscr{I}} = \mathbf{v}_O|_{\mathscr{I}} \wedge \mathbf{P}|_{\mathscr{I}}$, per concludere è sufficiente provare che

$$\sum_{i,j} (P_i - O) \wedge \mathbf{F}_{ij} = \mathbf{0} .$$

In effetti, cambiando nomi agli indici, si ha che:

$$\sum_{i,j} (P_i - O) \wedge \mathbf{F}_{ij} = \frac{1}{2} \left(\sum_{i,j} (P_i - O) \wedge \mathbf{F}_{ij} + \sum_{j,i} (P_j - O) \wedge \mathbf{F}_{ji} \right) .$$

Tenendo ora conto del fatto che l'ordine in cui si somma è irrilevante, si trova che:

$$\sum_{i,j}(P_i - O) \wedge \mathbf{F}_{ij} = \frac{1}{2}\sum_{i,j}\left((P_i - O) \wedge \mathbf{F}_{ij} + (P_j - O) \wedge \mathbf{F}_{ji}\right) .$$

Usando il principio di azione e reazione:

$$\sum_{i,j}(P_i - O) \wedge \mathbf{F}_{ij} = \frac{1}{2}\sum_{i,j}\left((P_i - O) \wedge \mathbf{F}_{ij} - (P_j - O) \wedge \mathbf{F}_{ij}\right) = \frac{1}{2}\sum_{i,j}(P_i - P_j) \wedge \mathbf{F}_{ij} .$$

Il secondo principio della dinamica *in forma forte* (**C3**) assicura che $\mathbf{F}_{ij}$ è parallela a $P_i - P_j$, pertanto $(P_i - P_j) \wedge \mathbf{F}_{ij} = \mathbf{0}$ e questo conclude la dimostrazione. □

Osservazioni 4.2.
(1) I vettori di V_t

$$\mathbf{R}^{(e)} := \sum_{i=1}^{N} \mathbf{F}_i^{(e)}$$

e

$$\mathbf{M}_O^{(e)} := \sum_{i=1}^{N}(P_i - O) \wedge \mathbf{F}_i^{(e)} ,$$

vengono chiamati, rispettivamente, **risultante delle forze esterne** e **risultante dei momenti delle forze esterne** rispetto al polo O. Più in generale, se O è un punto arbitrario e $\mathbf{F}$ è una forza agente sul punto materiale P, $(P - O) \wedge \mathbf{F}$ è detto **momento della forza F rispetto al polo** O.
(2) A causa della (4.1) la prima equazione cardinale può essere equivalentemente riformulata come:

$$M\mathbf{a}_G|_{\mathscr{I}} = \sum_{i=1}^{N} \mathbf{F}_i^{(e)} . \tag{4.8}$$

In questa forma l'equazione afferma che il centro di massa del sistema evolve, in accordo con il secondo principio della dinamica, come se fosse un unico punto materiale in cui è concentrata tutta la massa del sistema e che è sottoposto alla somma di tutte le forze *esterne* agenti sul sistema.
(3) Si definiscono analogamente la **risultante delle forze interne**, $\mathbf{R}^{(i)}$, e la **risultante dei momenti delle forze interne**, $\mathbf{M}_O^{(i)}$. Nel corso della dimostrazione delle equazioni cardinali, è stato provato che

$$\mathbf{R}^{(i)} = \mathbf{0} \quad \text{e} \quad \mathbf{M}_O^{(i)} = \mathbf{0} ,$$

per ogni sistema di punti materiali che soddisfa i principi della dinamica C1-C4 (anche nella versione estesa) e, per quanto riguarda la risultate dei momenti, indipendentemente dalla scelta del polo O.
(4) In generale, le due equazioni cardinali non sono in grado di determinare il moto del sistema

una volta assegnate condizioni iniziali. Tuttavia nel caso in cui il sistema soddisfi il *vincolo di rigidità* (cioè le distanze tra i punti del sistema sono costanti indipendentemente dalle forze esterne e dalle condizioni iniziali) le due equazioni determinano, se assegnate le condizioni iniziali, il moto del sistema fisico.

4.1.3 Leggi di bilancio/conservazione di impulso e momento angolare.

In base alla prima equazione cardinale, se il sistema è isolato oppure, più debolmente, se la somma delle forze esterne è nulla (in un intervallo di tempo), allora l'impulso totale del sistema si conserva nel tempo (in quell'intervallo di tempo).
In base alla seconda equazione cardinale, se il sistema è isolato oppure, più debolmente, se la somma dei momenti delle forze esterne rispetto a O è nulla (in un intervallo di tempo), e in entrambi i casi si è scelto $O \equiv G$ oppure O con accelerazione nulla in $\mathscr{I}$, allora il momento angolare totale del sistema è un *integrale primo del moto*, cioé si conserva nel tempo (in quell'intervallo di tempo).
I risultati appena citati valgono più debolmente lungo una fissata direzione.

Proposizione 4.1. *Nelle ipotesi del teorema 4.1 valgono i seguenti fatti in riferimento ad un arbitrario versore* $\mathbf{n}$ *costante nel tempo nel riferimento* $\mathscr{I}$.
(a) *Se* $\mathbf{R}^{(e)} \cdot \mathbf{n} = \mathbf{0}$ *in un intervallo temporale, allora* $\mathbf{P}|_{\mathscr{I}} \cdot \mathbf{n}$ *è costante in quell'intervallo temporale su ogni moto del sistema (il valore della costante dipende dal particolare moto).*
(b) *Se* $O = G$ *oppure* O *è in quiete in* $\mathscr{I}$, *e* $\mathbf{M}_O^{(e)} \cdot \mathbf{n} = 0$ *in un intervallo temporale, allora* $\mathbf{\Gamma}_O|_{\mathscr{I}} \cdot \mathbf{n}$ *è costante in quell'intervallo temporale su ogni moto del sistema (il valore della costante dipende dal particolare moto).*
(a) e (b) valgono in particolare in ogni sistema di riferimento inerziale e per ogni direzione $\mathbf{n}$, *quando il sistema di punti materiali è* **isolato**, *cioé non è sottoposto a forze esterne.*

Dimostrazione. La prova di (a) è immediata dalla prima equazione cardinale. La prova di (b) è immediata dalla seconda equazione cardinale tenendo conto che, se $O = G$ oppure O è in quiete in $\mathscr{I}$ allora $\mathbf{v}_O|_{\mathscr{I}} \wedge \mathbf{P}|_{\mathscr{I}} = \mathbf{0}$ (nel primo caso perché risulta $\mathbf{v}_O|_{\mathscr{I}} = \mathbf{v}_G|_{\mathscr{I}}$ e vale la (4.1)). □

Osservazioni 4.3.
(1) Nella storia della fisica, estendendo la classe di sistemi fisici studiati, si è visto che è stato sempre possibile definire un contributo all'impulso totale del sistema (che può non essere totalmente meccanico, per esempio può contenere il campo elettromagnetico) in modo tale che l'impulso totale di un sistema isolato venga ancora conservato. Questo principio, cioè che si possa sempre estendere la nozione di impulso in modo tale da avere alla fine una legge di conservazione, prende il nome di **principio di conservazione dell'impulso**. La storia della fisica mostra che tale principio è molto più importante del teorema corrispondente in meccanica, ed ha validità sia nelle teorie relativistiche che in quelle quantistiche ed è legato alle proprietà di *omogeneità* dello spazio nei riferimenti inerziali.

(2) Nello stesso modo, si è visto che estendendo la classe di sistemi fisici studiati, è stato sempre possibile definire un contributo al momento angolare totale del sistema (che può non essere totalmente meccanico, per esempio può contenere il campo elettromagnetico) in modo tale che il momento angolare totale di un sistema isolato venga ancora conservato. Questo principio, cioè che si possa sempre estendere la nozione di momento angolare in modo tale da avere alla fine una legge di conservazione, prende il nome di **principio di conservazione del momento angolare**. La storia della fisica mostra che tale principio è molto più importante del teorema corrispondente in meccanica, ed ha validità sia nelle teorie relativistiche che in quelle quantistiche ed è legato alle proprietà di *isotropia* dello spazio nei riferimenti inerziali.

4.2 Energia meccanica

Introduciamo ora le nozioni fondamentali che si riferiscono all'energia meccanica per punti materiali e sistemi di punti materiali.

Definizione 4.2. Sia P un punto materiale di massa m, soggetto alla forza $\mathbf{F}$ all'istante t.
(a) L'**energia cinetica** di P rispetto ad un riferimento $\mathscr{I}$ al tempo t è il numero:

$$\tau|_{\mathscr{I}}(t) := \frac{1}{2} m_k \mathbf{v}_P^2|_{\mathscr{I}}(t) \,,$$

dove abbiamo usato la notazione $\mathbf{u}^2 := \mathbf{u} \cdot \mathbf{u}$ per $\mathbf{u} \in V_t$.
(b) La **potenza** dissipata dalla forza all'istante t è

$$\pi|_{\mathscr{I}}(t) := \mathbf{v}_P|_{\mathscr{I}}(t) \cdot \mathbf{F} \,.$$

Se è dato un sistema S punti materiali P_k con masse m_k, $k = 1, 2, \ldots, N$,
(c) l'**energia cinetica totale** del sistema S rispetto ad un riferimento $\mathscr{I}$ al tempo t è il numero:

$$\mathscr{T}|_{\mathscr{I}}(t) := \sum_{k=1}^{N} \frac{1}{2} m_k \mathbf{v}_{P_k}^2|_{\mathscr{I}}(t) \,.$$

$\diamondsuit$

Esempi 4.1.
1. Esistono forze che vengono dette **forze resistenti passive** o **forze dissipative** rispetto ad un riferimento $\mathscr{I}$. Queste forze hanno la caratteristica di avere potenza dissipata strettamente negativa rispetto a $\mathscr{I}$ per velocità non nulle:

$$\mathbf{F}(t, P, \mathbf{v}_P|_{\mathscr{I}}) \cdot \mathbf{v}_P|_{\mathscr{I}} < 0 \,, \quad \text{per ogni } \mathbf{v}_P|_{\mathscr{I}} \in V_{\mathscr{I}},\ \mathbf{v}_P|_{\mathscr{I}} \neq \mathbf{0}.$$

L'esempio tipico è quello della *forza viscosa*. Un fluido in quiete in un riferimento $\mathscr{I}$ esercita su un punto materiale P una forza che ha la forma funzionale

$$\mathbf{F} = -g(||\mathbf{v}_P|_{\mathscr{I}}||) \frac{\mathbf{v}_P|_{\mathscr{I}}}{||\mathbf{v}_P|_{\mathscr{I}}||} \,.$$

La funzione g è non negativa. Per piccole velocità (es. fino a $2m/s$ in aria) la funzione g è costante. Per alte velocità l'andamento è più complicato e segue una legge di potenza come $k||\mathbf{v}_P|_{\mathscr{I}}||^n$ con $k > 0$.
2. Esistono forze che vengono dette **forze girostatiche** rispetto ad un riferimento $\mathscr{I}$. Esse sono caratterizzate dal fatto che la potenza dissipata rispetto a $\mathscr{I}$ è sempre nulla:

$$\mathbf{F}(t, P, \mathbf{v}_P|_{\mathscr{I}}) \cdot \mathbf{v}_P|_{\mathscr{I}} = 0 \,, \quad \text{per ogni } \mathbf{v}_P|_{\mathscr{I}} \in V_{\mathscr{I}}.$$

L'esempio tipico è la *forza di Lorentz* che agisce su un punto materiale di carica e, quando è immerso nel campo magnetico $\mathbf{B}(t, P)$ assegnato nel riferimento $\mathscr{I}$ (c è la velocità della luce):

$$\mathbf{F} = \frac{e}{c}\mathbf{v}_P|_{\mathscr{I}} \wedge \mathbf{B}(t, P) \,.$$

Un'altro esempio, nel qual caso si tratta di una forza inerziale, è dato dalla forza di Coriolis discussa nel capitolo 3:

$$\mathbf{F}^{(Coriolis)} = -2m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{v}_P|_{\mathscr{I}'} \,.$$

L'energia cinetica di un sistema di punti si può decomporre in modo canonico come energia del centro di massa ed energia "attorno al centro di massa". Questa decomposizione è utile in vari campi. Essa è introdotta dal seguente elementare ma popolare teorema.

Teorema 4.2. (**Teorema di König**.) *Per un sistema di punti materiali P_k, $k = 1, \ldots, N$, l'energia cinetica rispetto al riferimento $\mathscr{I}$ è pari all'energia cinetica nel riferimento $\mathscr{I}_G$, in cui G è in quiete e $\boldsymbol{\omega}_{\mathscr{I}_G}|_{\mathscr{I}} = \mathbf{0}$, sommata all'energia cinetica di un punto materiale che occupa la posizione di G istante per istante ed ha massa pari alla massa totale del sistema M: In formule (tralasciando la dipendenza temporale per semplicità)*

$$\mathscr{T}|_{\mathscr{I}} = \frac{1}{2}M\mathbf{v}_G^2|_{\mathscr{I}} + K|_{\mathscr{I}_G} \,. \tag{4.9}$$

$\diamondsuit$

Dimostrazione. Nelle ipotesi fatte, dalla definizione di energia cinetica ed usando (1.68)

$$\mathscr{T}|_{\mathscr{I}} = \sum_k \frac{1}{2}m_k\mathbf{v}_{P_k}|_{\mathscr{I}} \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} = \sum_k \frac{1}{2}m_k\left(\mathbf{v}_{P_k}|_{\mathscr{I}_G} + \mathbf{v}_G|_{\mathscr{I}}\right) \cdot \left(\mathbf{v}_{P_k}|_{\mathscr{I}_G} + \mathbf{v}_G|_{\mathscr{I}}\right)$$

$$= \sum_k \frac{1}{2}m_k\mathbf{v}_{P_k}^2|_{\mathscr{I}_G} + \sum_k \frac{1}{2}m_k\mathbf{v}_G^2|_{\mathscr{I}} + \sum_k m_k\mathbf{v}_{P_k}|_{\mathscr{I}_G} \cdot \mathbf{v}_G|_{\mathscr{I}}$$

$$= \sum_k \frac{1}{2}m_k\mathbf{v}_{P_k}^2|_{\mathscr{I}_G} + \frac{1}{2}M\mathbf{v}_G^2|_{\mathscr{I}} + \sum_k m_k\mathbf{v}_{P_k}|_{\mathscr{I}_G} \cdot \mathbf{v}_G|_{\mathscr{I}} = \sum_k \frac{1}{2}m_k\mathbf{v}_{P_k}^2|_{\mathscr{I}_G} + \frac{1}{2}M\mathbf{v}_G^2|_{\mathscr{I}} + M\mathbf{v}_G|_{\mathscr{I}_G} \cdot \mathbf{v}_G|_{\mathscr{I}}.$$

Dove abbiamo usato la definizione di centro di massa. L'ultimo addendo è nullo essendo, nelle nostre ipotesi, $\mathbf{v}_G|_{\mathscr{I}_G} = \mathbf{0}$ e quindi:

$$\mathscr{T}|_{\mathscr{I}} = \sum_k \frac{1}{2}m_k\mathbf{v}_{P_k}^2|_{\mathscr{I}_G} + \frac{1}{2}M\mathbf{v}_G^2|_{\mathscr{I}} \,.$$

$\square$

4.2.1 Teorema delle forze vive.

Il primo teorema di bilancio che riguarda l'energia cinetica è il cosiddetto *teorema delle forze vive.*

Teorema 4.3. (**Teorema delle forze vive**.) *Si consideri un sistema S di punti materiali P_k con masse m_k, $k = 1, 2, \ldots, N$, che soddisfa i principi C1-C4 (includendo il caso di forze inerziali e reazioni vincolari). Nel riferimento $\mathscr{I}$ e per ogni fissato istante del tempo assoluto t vale* **l'equazione delle forze vive***:*

$$\Pi^{(e)}|_{\mathscr{I}} + \Pi^{(i)}|_{\mathscr{I}} = \frac{d\mathscr{T}|_{\mathscr{I}}}{dt} \,. \tag{4.10}$$

Dove, all'istante considerato, $\Pi^{(e)}|_{\mathscr{I}}$ e $\Pi^{(i)}|_{\mathscr{I}}$ sono rispettivamente la **potenza totale delle forze esterne***, cioé la somma delle potenze dissipate dalle forze esterne su ciascun punto del sistema, e la* **potenza totale dissipata dalle forze interne***, cioé la somma delle potenze dissipate dalle forze interne su ciascun punto del sistema.*

Dimostrazione. Per il punto materiale k-esimo, se $\mathbf{F}_k^{(i)}$ e $\mathbf{F}_k^{(e)}$ sono rispettivamente la somma delle forze interne ed esterne agente su di esso, tenendo conto del secondo principio della dinamica abbiamo:

$$\mathbf{v}_{P_k}|_{\mathscr{I}} \cdot (\mathbf{F}_k^{(i)} + \mathbf{F}_k^{(e)}) = m_k \mathbf{v}_{P_k}|_{\mathscr{I}} \cdot \frac{d\mathbf{v}_{P_k}|_{\mathscr{I}}}{dt} = \frac{d}{dt}\left(\frac{1}{2} m_k \mathbf{v}_{P_k}^2|_{\mathscr{I}}\right) \,.$$

Prendendo la somma su k si ottiene l'equazione (4.3). □

Osservazioni 4.4. Abbiamo precisato il riferimento anche in relazione alla potenza totale dissipata dalle forze interne. In realtà si può provare che tale potenza totale *non dipende dal riferimento.* Questo risultato notevole cade sotto il nome di **principio di indifferenza meccanica** ed ha una notevole importanza nei successivi sviluppi della fisica, passando dalla meccanica alla termodinamica, nella definizione di energia termodinamica interna del sistema come funzione indipendente dal riferimento. Dimostriamo l'indipendenza citata.
Le forze interne sono forze "vere" quindi, per definizione, non dipendono dal sistema di riferimento $\Pi^{(i)}|_{\mathscr{I}} = \Pi^{(i)}|_{\mathscr{I}'}$. La relazione tra le differenti velocità di ciascun punto del sistema al variare del riferimento è data dalle (1.68).

$$\mathbf{v}_{P_k}|_{\mathscr{I}} = \mathbf{v}_{P_k}|_{\mathscr{I}'} + \mathbf{v}_{O'}|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P_k - O') \,.$$

Da ciò, con ovvie notazioni si ha:

$$\Pi^{(i)}|_{\mathscr{I}} - \Pi^{(i)}|_{\mathscr{I}'} = \mathbf{v}_{O'}|_{\mathscr{I}} \cdot \sum_k \mathbf{F}_k^{(i)} + \sum_k \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P_k - O') \cdot \mathbf{F}_k^{(i)} \,.$$

Per note proprietà del prodotto vettore: $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P_k - O') \cdot \mathbf{F}_k^{(i)} = (P_k - O') \wedge \mathbf{F}_k^{(i)} \cdot \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}$, per cui:

$$\Pi^{(i)}|_{\mathscr{I}} - \Pi^{(i)}|_{\mathscr{I}'} = \mathbf{v}_{O'}|_{\mathscr{I}} \cdot \sum_k \mathbf{F}_k^{(i)} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \cdot \sum_k (P_k - O') \wedge \mathbf{F}_k^{(i)}$$

Cioé

$$\Pi^{(i)}|_{\mathscr{I}} - \Pi^{(i)}_{\mathscr{I}'} = \mathbf{v}_{O'}|_{\mathscr{I}} \cdot \mathbf{R}^{(i)} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \cdot \mathbf{M}^{(i)}_{O'}$$

Da (3) in osservazioni 4.2, concludiamo che $\mathbf{R}^{(i)} = \mathbf{0}$ e $\mathbf{M}^{(i)}_{O'} = \mathbf{0}$ per cui:

$$\Pi^{(i)}|_{\mathscr{I}} - \Pi^{(i)}|_{\mathscr{I}'} = 0\,.$$

4.2.2 Forze conservative.

Definizione 4.3. Dato un sistema S di punti materiali P_k, con $k = 1, 2, \ldots, N$, si dice che è stata assegnato un sistema di **forze posizionali** rispetto al riferimento $\mathscr{I}$, se per ogni punto P_k è assegnata una funzione di forza (escludiamo il caso di forze *inerziali*) $\mathbf{F}_k$, che nel riferimento $\mathscr{I}$ assume forma $\mathbf{F}_k = \mathbf{F}_k(P_1, \ldots, P_k)$. Si dice ulteriormente che il sistema di forze posizionali è **conservativo** nel sistema $\mathscr{I}$ se esiste una funzione $\mathscr{U}|_{\mathscr{I}} \in C^1(E_{\mathscr{I}} \times \cdots \times E_{\mathscr{I}})$, detta **energia potenziale** associata al sistema di forze in $\mathscr{I}$, tale che in $\mathscr{I}$:

$$\mathbf{F}_k(P_1, \ldots, P_N) = -\nabla_{P_k}\mathscr{U}|_{\mathscr{I}}(P_1, \ldots, P_N)\,, \quad \text{per ogni } P_k \in E_{\mathscr{I}} \text{ e } k = 1, 2, \ldots, N\,. \tag{4.11}$$

$\diamondsuit$

Osservazioni 4.5.
(1) La funzione $-\mathscr{U}|_{\mathscr{I}}$ viene chiamata **potenziale** del sistema di forze. Ovviamente sia $\mathscr{U}|_{\mathscr{I}}$ che il potenziale sono definiti a meno di costanti additive.
(2) Esistono sistemi di forze le cui forze possono essere scritte come in (4.11) dove però , $\mathscr{U}|_{\mathscr{I}}$ dipende anche dal tempo. In tal caso $\mathscr{U}|_{\mathscr{I}}$ *non* prende più il nome di energia potenziale, ma $-\mathscr{U}|_{\mathscr{I}}$ viene ancora chiamato potenziale della forza. La situazione descritta si ha immediatamente quando, data un sistema di forze conservative rispetto al riferimento $\mathscr{I}$, si passa ad un nuovo riferimento con una trasformazione di Galileo che coinvolge una velocità di trascinamento.
(3) Particolari casi di forze conservative, per sistemi composti da un'unica particella, sono le cosiddette *forze centrali*: una forza $\mathbf{F}$ è detta **centrale** nel riferimento $\mathscr{I}$ con **centro** $O \in E_{\mathscr{I}}$, se è posizionale e soddisfa la due condizioni:
(i) $\mathbf{F}(P)$ è parallela a $P - O$, per ogni $P \in E_{\mathscr{I}}$,
(ii) $\mathbf{F}(P)$ è funzione solamente di $||P - O||$.
La dimostrazione del fatto che una forza centrale è conservativa è lasciata negli esercizi. È invece facile provare che ogni forza conservativa $\mathbf{F}$ rispetto al riferimento $\mathscr{I}$, che soddisfa (i) rispetto a $O \in E_{\mathscr{I}}$, è centrale rispetto ad O. Lavorando in coordinate polari sferiche solidali con $\mathscr{I}$ e con origine in O, deve essere

$$\mathbf{F}(P) = -\frac{\partial\mathscr{U}|_{\mathscr{I}}}{\partial r}\,\mathbf{e}_r - \frac{1}{r}\frac{\partial\mathscr{U}|_{\mathscr{I}}}{\partial\theta}\,\mathbf{e}_\theta - \frac{1}{r\sin\theta}\frac{\partial\mathscr{U}|_{\mathscr{I}}}{\partial\varphi}\,\mathbf{e}_\varphi\,.$$

Per soddisfare il vincolo (i), le due ultime derivate devono essere nulle e, di conseguenza, l'energia potenziale U e quindi la forza stessa, non possono dipende da θ e φ, ma solo da $r = ||P - O||$.

Usando teoremi ben noti dai corsi di Analisi [GiustiI] si può provare facilmente il seguente teorema, che caratterizza la forze conservative nel caso di un sistema composto da un unico punto. Il teorema può essere facilmente generalizzato al caso di sistemi di più punti materiali, lasciamo al lettore tale generalizzazione.

Teorema 4.4. *Si consideri la forza posizionale rispetto al riferimento* $\mathscr{I}$, $\mathbf{F} : \Omega \to \mathbb{R}$, *con* $\Omega \subset E_{\mathscr{I}}$ *aperto e si supponga che* $\mathbf{F}$ *sia continua.*
(a) *le seguenti condizioni sono equivalenti:*
(i) $\mathbf{F}$ *è conservativa;*
(ii) per ogni coppia di punti $P, Q \in E_{\mathscr{I}}$, *l'integrale* $\int_{P\ \Gamma}^{Q} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}$, *non dipende dalla curva regolare* $\Gamma : \mathbf{x} = \mathbf{x}(s)$ *di estremi* P *e* Q *purché sia completamente inclusa in* Ω*;*
(iii) per ogni curva chiusa regolare Γ *tutta contenuta in* Ω *vale*

$$\oint_{\Gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = 0 \,.$$

(b) *Se* $\mathbf{F}$ *è conservativa l'energia potenziale* U *è di classe* $C^1(\Omega)$ *e può essere definita come:*

$$\mathscr{U}|_{\mathscr{I}}(P) := -\int_{O\ \Gamma}^{P} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} \,,$$

dove $O \in \Omega$ *è un punto arbitrario fissato una volta per tutte e* Γ *è una qualsiasi curva regolare da* O *a* P*, tutta contenuta in* Ω*.*
(c) *Se* $\mathbf{F} \in C^1(\Omega)$ *ed è conservativa allora*

$$\nabla \wedge \mathbf{F}(P) = \mathbf{0} \,, \quad \textit{per ogni } P \in E_{\mathscr{I}} \,. \tag{4.12}$$

(d) *Se* $\mathbf{F} \in C^1(\Omega)$, Ω *è semplicemente connesso e vale (4.12), allora* $\mathbf{F}$ *è conservativa.* $\diamondsuit$

Esercizi 4.2.

1. Dimostrare che se $\mathbf{F} : \Omega \to V_{\mathscr{I}}$ è continua e centrale rispetto al punto $O \in E_{\mathscr{I}}$, con Ω aperto, allora $\mathbf{F}$ è conservativa.

2. Si consideri un punto materiale P sottoposto alla sola forza centrale $\mathbf{F}$ con centro O nel riferimento inerziale $\mathscr{I}$. Mostrare che, nel riferimento $\mathscr{I}$, il moto del punto avviene in un piano che è perpendicolare al momento angolare del punto $\mathbf{\Gamma}_O|_{\mathscr{I}}$ che è costante nel tempo.

Osservazioni 4.6.
(1) Si possono considerare sistemi di forze conservative che sono tali *indipendentemente dal riferimento*. Questa situazione si per sistemi di *due* punti materiali, P e Q quando essi interagiscono con una coppia di forze (costituente una *coppia azione-reazione*), $\mathbf{F}(P, Q)$ e $-\mathbf{F}(P, Q)$, rispettivamente agente su P e su Q per la quale esista una funzione $U = U(r)$ di classe $C^1(\mathbb{R})$ (o su qualche sottoinsieme aperto di $\mathbb{R}$) con, se $\mathscr{U}(P, Q) := U(||P - Q||)$

$$\mathbf{F}(P, Q) = -\nabla_P \mathscr{U}(P, Q) \,. \tag{4.13}$$

Si osservi che, in tal caso, $-\nabla_Q U(||P-Q||) = -\mathbf{F}(P,Q)$, per cui il ruolo dei due punti è simmetrico. La richiesta che U sia funzione di $||P-Q||$ e non di $P-Q$ implica immediatamente che $\mathbf{F}$ sia diretta lungo la congiungente P e Q come richiesto dal principio di azione e reazione in forma forte. Si osservi infine che la relazione (4.13) vale in ogni riferimento (le forze considerate sono forze vere e quindi invarianti al variare del riferimento) assumendo, come è tacito dalla notazione, che forma funzionale di U non dipenda del riferimento. Ciò è coerente visto che la distanza $d_t(P,Q) = ||P-Q||$ è assoluta e non dipende dal riferimento. Anche in questo caso la funzione U viene detta *energia potenziale* ora associata alla *coppia* di forze $\mathbf{F}(P,Q)$ e $-\mathbf{F}(P,Q)$.
La situazione si estende al caso di più punti materiali quando, per ogni coppia di punti P_i, P_j con $i \neq j$, la corrispondente coppia azione-reazione del sistema di forze ammette un'energia potenziale $\mathscr{U}_{ij}(P_i,P_j) = U_{ij}(||P_i - P_j||)$ del tipo detto sopra. L'energia potenziale totale, per la quale vale (4.13), si ottiene sommando quella di tutte le coppie possibili:

$$\mathscr{U}(P_1, P_2, \ldots, P_N) = \sum_{i<j} \mathscr{U}_{ij}(P_i, P_j)\,.$$

Si osservi che la coppia P_i, P_j e la coppia P_j, P_i contribuiscono con un'unica funzione energia potenziale $\mathscr{U}_{ij}$.
(2) Tornando al caso di due punti materiali P, Q con energia potenziale $\mathscr{U}(P,Q) = U(||P-Q||)$, se la posizione di Q è tenuta fissa, in quiete, in un riferimento $\mathscr{I}$, mediate l'azione di forze supplettive agenti su Q, la forza agente sul P è di fatto una forza centrale con centro dato da Q, quando la si descrive nel riferimento $\mathscr{I}$.

4.2.3 Bilancio e conservazione dell'energia meccanica.

Da tutto quanto visto si ha il seguente fondamentale teorema, che enuncia l'equazione di bilancio per l'energia meccanica e la sua proprietà di conservazione nel caso in cui tutte le forze in gioco siano conservative.

Teorema 4.5. (**Bilancio e conservazione dell'energia meccanica**.) *Si consideri un sistema S di punti materiali P_k, con $k = 1, 2, \ldots, N$, di masse m_k rispettivamente che soddisfi i principi C1-C4 (anche nelle forme generalizzate). Si supponga che S sia sottoposto, in addizione ad altre eventuali forze* non conservative*, ad un sistema di forze conservative rispetto al riferimento $\mathscr{I}$ con energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(P_1, \ldots, P_N)$.*
*Se si definisce l'***energia meccanica totale del sistema** *nel riferimento $\mathscr{I}$:*

$$\mathscr{E}|_{\mathscr{I}} := \mathscr{T}|_{\mathscr{I}} + \mathscr{U}|_{\mathscr{I}}\,.$$

Per ogni istante di tempo assoluto t vale l'equazione di bilancio, su ogni moto del sistema

$$\Pi|_{\mathscr{I}}^{(noncons.)} = \frac{d\mathscr{E}|_{\mathscr{I}}}{dt}\,,$$

dove $\Pi|_{\mathscr{I}}^{(noncons)}$ *è la potenza totale delle forze non conservative nel riferimento* $\mathscr{I}$*.*
Nel caso in cui tutte le forze agenti sul sistema siano conservative, l'energia meccanica è un integrale primo del moto*, cioé è conservata al variare del tempo lungo i moti del sistema.*

Dimostrazione. Dall'equazione (4.10) è sufficiente dimostrare che la potenza delle forze conservative $\Pi|_{\mathscr{I}}^{(cons)}$ soddisfa

$$\Pi|_{\mathscr{I}}^{(cons)} = -\frac{\mathscr{U}|_{\mathscr{I}}(P_1(t), \ldots, P_N(t))}{dt} ,$$

dove $P_k = P_k(t)$ risolve l'equazione del moto per il sistema di punti. In effetti, dalla (4.11), si ha che:

$$\Pi|_{\mathscr{I}}^{(cons)} = \sum_k \pi^{(cons)}|_{\mathscr{I}k} = -\sum_k \mathbf{v}_{P_k} \cdot \nabla_{P_k} \mathscr{U}|_{\mathscr{I}}(P_1(t), \ldots, P_N(t)) = -\frac{d\mathscr{U}|_{\mathscr{I}}(P_1(t), \ldots, P_N(t))}{dt} ,$$

che è quanto si voleva provare. □

Osservazioni 4.7. Si osservi che in generale l'energia *meccanica* non è conservata per sistemi fisici reali isolati, a causa delle forze non conservative interne al sistema che in natura sono sempre presenti (gli attriti interni). Estendendo la classe di sistemi fisici studiati, si è visto però che è sempre possibile definire un contributo all'energia totale del sistema (che può non essere totalmente meccanica, per esempio può contenere l'energia del campo elettromagnetico oppure energia interna termodinamica) in modo tale che l'energia totale di un sistema *isolato* sia conservata nel tempo quando valutata in un riferimento inerziale. Il principio generale, che afferma che si possa sempre estendere la nozione di energia in modo tale da avere alla fine una legge di conservazione, prende il nome di **principio di conservazione dell'energia**. La storia della fisica mostra che tale principio è molto più importante del teorema dimostrato sopra in meccanica, ed ha validità sia nelle teorie relativistiche che in quelle quantistiche ed è legato alle proprietà di omogeneità del tempo nei riferimenti inerziali.

Esercizi 4.3.

1. Si consideri un punto materiale P di massa m vincolato ad un'asta orizzontale ℓ liscia passante per il punto O. L'asta è in quiete nel riferimento $\mathscr{I}'$. Definire un sistema di coordinate cartesiane ortonormali destrorse solidali con $\mathscr{I}'$, centrate in O, con asse x diretto lungo ℓ, asse y orizzontale e perpendicolare a ℓ e asse z verticale diretto verso H. P è sottoposto alla forza di gravità $-mg\,\mathbf{e}_z$ ed ad una forza dovuta ad una molla ideale, di lunghezza nulla a riposo e di costante $\kappa > 0$, fissata ad un estremo al punto P ed all'altro estremo al punto fisso H di altezza h sulla verticale di O. Il riferimento $\mathscr{I}'$ ruota, attorno all'asse $\mathbf{e}_z$, con velocità angolare costante $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} = \omega\,\mathbf{e}_z$ rispetto al riferimento inerziale $\mathscr{I}$.
(i) Scrivere le equazioni pure di movimento per il punto P e darne la soluzione generale a seconda del valore del rapporto $m\omega^2/\kappa$.
(ii) Nel caso $m\omega^2/\kappa < 1$, esprimere la reazione vincolare $\boldsymbol{\phi}$ in funzione del tempo, per il moto

con condizioni iniziali $P(0) - O = x_0\,\mathbf{e}_x$, con $x_0 > 0$, e $\mathbf{v}_P|_{\mathscr{I}'}(0) = \mathbf{0}$.
(iii) Dimostrare che la componente parallela all'asse z del momento angolare di P con polo O non è in generale conservata nel riferimento $\mathscr{I}$.
(iv) Individuare tutte le condizioni iniziali (a $t = 0$) che producono moti in cui la componente parallela all'asse z del momento angolare di P con polo O rispetto a $\mathscr{I}$ è conservata nel tempo.

2. Si consideri il sistema fisico dell'esercizio precedente.
(i) Dimostrare che nel riferimento $\mathscr{I}'$ ha senso, dal punto di vista matematico, definire un'energia potenziale di parte delle forze inerziali, in addizione a quella della molla, in modo tale che l'energia meccanica totale sia conservata nel riferimento $\mathscr{I}'$.
(ii) Individuare tutte le condizioni iniziali (a $t = 0$) che producono moti in cui l'energia meccanica è conservata nel riferimento $\mathscr{I}$.

3. Si considerino due punti materiali P e Q di masse M ed m rispettivamente. Tali punti siano vincolati alla superficie cilindrica C di equazione $x^2 + y^2 = R^2$ con $R > 0$, dove x, y, z sono coordinate cartesiane ortonormali solidali con un il riferimento inerziale $\mathscr{I}$. I due punti sono sottoposti alla forza gravitazionale individuata dall'accelerazione $\mathbf{g} := -g\,\mathbf{e}_z$, sono connessi l'un l'altro da una molla di costante elastica $\kappa > 0$ e lunghezza nulla a riposo e ciascuno di essi è connesso all'origine O delle coordinate con una molla di costante elastica $\gamma > 0$ e lunghezza a riposo nulla.
(i) Si scrivano le equazioni pure di movimento in coordinate cilindriche r, φ, z adattate a C con $\varphi = 0$ sull'asse x.
(ii) Si calcolino le reazioni vincolari sui punti all'istante iniziale per le condizioni iniziali $P(0) - O = R\,\mathbf{e}_x$, $Q(0) - O = \mathbf{e}_y$, $\mathbf{v}_P(0) = \mathbf{v}_Q(0) = \mathbf{0}$.
(iii) Si dimostri che la quantità $M\dot{\varphi}_P + m\dot{\varphi}_Q$ (dove il punto indica, al solito, la derivata temporale) è conservata nel tempo su ogni moto del sistema.

4.3 *La necessità della descrizione in termini di continui e di campi in meccanica classica.

La descrizione data dei fondamenti della meccanica classica non è sufficiente ad includere tutti i fenomeni noti nella fisica classica, infatti, vi sono almeno due situazioni molto importanti in cui lo schema dato in termini di punti materiali non è sufficiente a descrivere la realtà fisica.
Il primo caso si ha quando ci si trova a lavorare con corpi fisici che cadono sotto la denominazione di corpi *continui*, in particolare i cosiddetti *fluidi*. Si tratta di sistemi fisici che all'apparenza macroscopica sono *estesi* (usualmente omogenei ed isotropi) e *continui*: nel senso che le configurazioni spaziali di tali corpi, nello spazio assoluto ad ogni istante, sono descritti geometricamente da insiemi omeomorfi ad aperti e connessi di $\mathbb{R}^3$. In tal caso la descrizione data precedentemente deve essere modificata al fine di introdurre opportune *funzioni densità* che generalizzano le varie grandezze introdotte prima. Per esempio si hanno funzioni densità di massa e funzioni (a valori vettoriali) densità di forze che integrate su porzioni di corpo (o superfici di tali porzioni) forniscono rispettivamente la massa e forza totale su tale porzione. È fondamentale in questa ottica (per esempio per poter introdurre la forza totale agente su una porzione finita di continuo) che

lo spazio assoluto ad ogni tempo abbia una struttura affine che permetta di parlare di vettori che non appartengono allo spazio tangente della varietà e di poter sommare vettori che appartengono a spazi tangenti in punti diversi. La generalizzazione della meccanica dei punti materiali al caso dei continui cade sotto il nome di **meccanica dei continui**. Un punto importante che ha avuto influenza fondamentale nello sviluppo delle fisica, è che la meccanica dei continui ammette anche una formulazione *locale* delle leggi della dinamica che invece di riferirsi a porzioni finite di continuo si riferisce a punti di continuo. In tale comoda formulazione, le equazioni integrodifferenziali relative a porzioni finite di continuo vengono tradotte in equazioni puramente differenziali valide punto per punto nel continuo. Tale formulazione è basata sull'esistenza (provata da Cauchy) di un particolare campo tensoriale detto *tensore degli sforzi di Cauchy*. Il punto rilevante è che, una volta enunciata la formulazione locale, la struttura affine non è più necessaria ed il formalismo si presta naturalmente ad essere esteso a situazioni più generali. In particolare la Teoria della Relatività Generale di Einstein sfrutta appieno tale possibilità essendo formulata su uno spaziotempo che non può ammettere struttura affine per motivi profondamente fisici (presenza del campo gravitazionale).
Il secondo caso in cui la formulazione data in termini di punti materiali non è più adeguata a descrivere la realtà fisica è quando si trattano le *forze elettromagnetiche*. In questa situazione le forze agenti sui punti materiali sono *mediate* da *campi di forza*. L'idea base è che ogni punto carico elettricamente (*sorgente*) generi un *campo di forze*, cioè una coppia di campi vettoriali $\mathbf{E} = \mathbf{E}(t, P)$ (**campo elettrico**) e $\mathbf{B} = \mathbf{B}(t, P)$ (**campo magnetico**) definiti in ogni punto dello spazio P e ad ogni istante t, di forma opportuna a seconda delle caratteristiche e dello stato di moto della sorgente. Gli altri punti carichi elettricamente che si trovano nello spazio, risentono di forze esclusivamente attraverso il vettore campo nel punto ove essi si trovano. Più precisamente, per un punto materiale nel vuoto di massa m e carica elettrica q che si trova nel punto $P(t)$ al tempo t nello spazio di quiete del riferimento inerziale $\mathscr{I}$ vale ancora

$$m\mathbf{a}_P|_{\mathscr{I}}(t) = \mathbf{F}_L(t, P(t), \mathbf{v}_P|_{\mathscr{I}}(t))$$

dove la funzione forza a secondo membro è data dalla **Forza di Lorentz**

$$\mathbf{F}_L(t, P(t), \mathbf{v}_P|_{\mathscr{I}}(t)) := q\mathbf{E}(t, P(t)) + \frac{q\mathbf{v}_P|_{\mathscr{I}}(t)}{c} \wedge \mathbf{B}(t, P(t)) \tag{4.14}$$

c è una costante pari alla velocità della luce nel vuoto. Si noti la presenza esplicita di t in $\mathbf{F}_L(t, P(t), \mathbf{v}_P|_{\mathscr{I}}(t))$ che denota la dipendenza dal moto della sorgente dei campi. Il fatto *sperimentale* che rende centrale invece che inutile l'idea di campo, è che se la sorgente modifica il suo stato di moto, il campo nei suoi dintorni si deforma conseguentemente con un certo ritardo: le perturbazioni del campo si muovono ad una velocità finita e le forze agenti sugli altri punti carichi risentono di tale ritardo. È chiaro che in questa situazione, in presenza di cariche in moto, il principio di azione e reazione cessa di valere e con esso *cessa di valere il principio di conservazione dell'impulso, se per impulso si intende la somma degli impulsi dei singoli punti materiali carichi*. Tuttavia si può dimostrare che: il principio di conservazione dell'impulso continua ad essere valido per il sistema complessivo *cariche più campo eletromagnetico*, purchè si definisca un contributo all'impulso totale dovuto al campo stesso. Tale contributo si ottiene

integrando nello spazio una densità di impulso associata al campo elettromagnetico. *In maniera simile si estendono le leggi di conservazione dell'energia e del momento angolare.* Infine si può mostrare che la dinamica del sistema complessivo è formulabile con equazioni locali (nel senso detto sopra) equivalenti alle equazioni integro-differenziali di partenza. Il campo elettromagnetico, in definitiva, viene ad essere trattato, pur con importanti differenze, in maniera simile a quanto appena detto per i continui. (In particolare, viene anche definito un tensore degli sforzi del campo elettromagnetico detto tensore di Maxwell).

Osservazioni 4.8.
(1) È chiaro che nella realtà i continui non esistono perchè i corpi fisici, ad una analisi microscopica, rivelano una natura molecolare e quindi discreta. A tale livello avvengono importantissimi nuovi fenomeni e l'unica descrizione nota è quella basata sulla Meccanica Quantistica.
(2) Anche il campo elettromagnetico ha una natura discreta/quantistica basata sul concetto di *fotone*, che condivide anche una natura relativistica.
(3) Le cariche elettriche elementari, gli elettroni, sembrerebbero essere a tutti gli effetti punti materiali, cioè essere privi di dimensioni (l'unica struttura è quella dovuta al cosiddetto *spin*). In realtà, a livello classico (inclusa la trattazione relativistica non quantistica delle cariche e del campo elettromagnetico) lo schema di carica elettrica puntiforme e campo elettromagnetico non quantistico produce gravi inconsistenze matematiche segnate dall'apparire di quantità infinite (autoforza, autoenergia...)[Jackson]. Tali inconsistenze sono legate al fatto che le cariche elettriche non puntiformi, in base alle stesse equazioni dell'elettromagnetismo risultano autointeragire con il campo elettromagnetico da esse stesse prodotto, tale fenomeno è fisicamente evidente nella *radiazione di frenamento* che si ha quando si cerca di accelerare una carica[1].

[1] Per questo motivo anche la stessa *definizione* di massa di una particella carica è problematica, e si è pensato più volte in passato e in ambito classico o relativistico, di definire la massa di un elettrone attraverso l'autointerazione di frenamento con il campo da essa emesso [Jackson]. Tali teorie, malgrado fossero molto interessanti, non hanno mai avuto successo.

Capitolo 5

Introduzione alla teoria della stabilità .

In questo capitolo introduciamo i primi rudimenti della teoria della stabilità nell'approccio di Liapunov. Lavoreremo con sistema di equazioni differenziali autonomi su $\mathbb{K}^n$, anche se con banali adattamenti, le definizioni ed i risultati che presenteremo si estendono alla teoria dei sistemi di equazioni differenziali su varietà differenziabili.

5.1 Punti singolari e configurazioni di equilibrio

La nozione fondamentale della quale ci occuperemo è quella di *punto singolare*. Dato che consideriamo solo sistemi autonomi, senza perdere generalità , con una traslazione dell'origine del parametro t, possiamo sempre ridurci a lavorare con condizioni iniziali per $t = 0$.

Definizione 5.1. (**Punto singolare**.) Si consideri un sistema di equazioni differenziali del prim'ordine autonomo su $D \subset \mathbb{K}^n$ insieme aperto non vuoto,

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t)) \ , \tag{5.1}$$

con $\mathbf{f} : D \to \mathbb{K}^n$. Il punto $\mathbf{x}_0 \in D$ è detto **singolare** per $\mathbf{f}$ se e solo se ogni soluzione $\mathbf{x} : I \to \mathbb{K}^n$, (con $I \subset \mathbb{R}$ intervallo aperto tale che $0 \in I$) del problema di Cauchy dato dall'equazione (5.1) con condizione iniziale

$$\mathbf{x}(0) = \mathbf{x}_0 \ , \tag{5.2}$$

è la soluzione costante $\mathbf{x}(t) = \mathbf{x}_0$ per ogni t nell'intervallo I. $\diamondsuit$

Osservazioni 5.1.
(1) Dato che la definizione data sopra è completamente *locale*, può essere riformulata immediatamente in riferimento ad un sistema di equazioni differenziali assegnato su una varietà differenziabile

tramite un campo vettoriale nell'ottica della sezione 3.5. lasciamo al lettore questa riformulazione.
(2) La stessa definizione può essere enunciata, con una terminologia leggermente differente considerando un problema del *secondo* ordine, con evidente riferimento al problema fondamentale della dinamica quando $\mathbb{K} = \mathbb{R}$ e $n = 3$ (ma anche in casi più complicati studiando *sistemi* di punti materiali).

Ecco la definizione a cui si riferisce l'osservazione (2) di sopra:

Definizione 5.2. (**Punto di equilibrio**.) Si consideri un sistema di equazioni differenziali del secondo ordine autonomo su $D \times D'$, con $D, D' \subset \mathbb{K}^n$ insiemi aperti non vuoti,

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F}\left(\mathbf{x}(t), \frac{d\mathbf{x}}{dt}\right) , \tag{5.3}$$

con $\mathbf{f} : D \times D' \to \mathbb{K}^n$. Il punto $\mathbf{x}_0 \in D$ è detto **punto di equilibrio** (o equivalentemente **configurazione di equilibrio**) per $\mathbf{F}$, se e solo se ogni soluzione $\mathbf{x} : I \to \mathbb{K}^n$, (con $I \subset \mathbb{R}$ intervallo aperto tale che $0 \in I$) del problema di Cauchy dato dall'equazione (5.3) con condizione iniziale

$$\begin{cases} \mathbf{x}(0) & = \mathbf{x}_0, \\ \dfrac{d\mathbf{x}}{dt} & = \mathbf{0}, \end{cases} \tag{5.4}$$

è la soluzione costante $\mathbf{x}(t) = \mathbf{x}_0$ per ogni t nell'intervallo I. $\diamondsuit$

Osservazioni 5.2.
(1) La definizione data si applica, per esempio, al caso di un punto materiale P descritto dal vettore $\mathbf{x} = P - O$, O in quiete in un riferimento $\mathscr{I}$ e quando $\mathbf{F}$ denota la forza totale agente sul punto, supposta indipendente dal tempo in $\mathscr{I}$. *Si deve quindi tenere conto che la definizione di punto di equilibrio dipende dal riferimento scelto.*
(2) La definizione si può applicare anche al caso di un sistema di N punti, quando

$$D = \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 = \mathbb{R}^{3N} .$$

In questo caso le $3N$ componenti di $\mathbf{x}$ descrivono, a tre a tre, le tre componenti di ciascuno degli N punti, in riferimento ad un fissato sistema di coordinate cartesiane ortonormali, solidali con un riferimento in cui tutte le forze sono indipendenti dal tempo. Le $3N$ componenti della funzione $\mathbf{F}$, descrivono a tre a tre le componenti delle forze totali agenti su ciascuno degli N punti materiali, *ciascuna divisa per la massa del punto corrispondente.*

Abbiamo l'ovvia proposizione seguente, che fornisce condizioni necessarie e sufficienti affinché un punto sia singolare ovvero di equilibrio, nel caso in cui sono soddisfatte le ipotesi per il teorema di esistenza ed unicità globale 3.5.

Proposizione 5.1. *In riferimento alla definizione 5.1, se la funzione* $: \mathbf{D} \to \mathbb{K}^{\mathbf{n}}$ *è localmente lipschitziana, il punto* $\mathbf{x}_0 \in D$ *è singolare se e solo se*

$$\mathbf{f}(\mathbf{x}_0) = \mathbf{0}\,. \tag{5.5}$$

$\diamondsuit$

Dimostrazione. Che la condizione (5.5) sia necessaria è ovvio dalla definizione di punto singolare. La sufficienza segue immediatamente dal teorema 3.5, tenendo conto che, se $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$ allora una soluzione massimale del problema di Cauchy con condizione iniziale $\mathbf{x}(0) = \mathbf{x}_0$ è sicuramente $\mathbf{x}(t) = \mathbf{x}_0$ costantemente per ogni $t \in \mathbb{R}$. Questa deve essere l'unica soluzione massimale per il teorema citato. La dimostrazione della sufficienza si conclude notando che, ogni soluzione non massimale è una restrizione di quella massimale, sempre per il teorema citato. $\square$

Segue l'analoga proposizione per la nozione di configurazione di equilibrio.

Proposizione 5.2. *In riferimento alla definizione 5.2, se la funzione* $\mathbf{F} = \mathbf{F}(\mathbf{x}, \mathbf{v}) \in \mathbb{K}^n$ *con* $(\mathbf{x}, \mathbf{v}) \in D \times D'$ *è localmente lipschitziana nella variabile* $(\mathbf{x}, \mathbf{v})$ *(o più fortemente è di classe* $C^1(D \times D'; \mathbb{K}^n)$*),* $\mathbf{x}_0 \in D$ *è configurazione di equilibrio se e solo se*

$$\mathbf{F}(\mathbf{x}_0, \mathbf{0}) = \mathbf{0}\,. \tag{5.6}$$

$\diamondsuit$

5.1.1 Equilibrio stabile ed instabile.

Dal punto di vista fisico, la nozione di configurazione di equilibrio data sopra non caratterizza in modo appropriato le configurazioni di equilibrio reali dei sistemi fisici (e degli altri sistemi della realtà descrivibili con equazioni differenziali). Una ragione, forse la più importante, è dovuta al fatto che le configurazioni dei sistemi fisici non corrispondono a configurazioni geometriche *ideali*, perché lo stato di ogni sistema fisico (la posizione e la forma stessa degli oggetti) è sottoposto a continue interazioni con l'ambiente circostante (per lo più interazioni di carattere termodinamico), che rendono lo stato fluttuante, sia pur con fluttuazioni molto piccole. Per questo motivo le configurazioni fisiche di equilibrio *osservabili nella realtà fisica* hanno anche una proprietà di *stabilità* sotto piccole perturbazioni dello stato iniziale. Queste proprietà di stabilità vengono definite a seconda del modello matematico adottato. Nel seguito ci riferiremo alle nozioni più elementari possibili, che possono essere formulate all'interno della teoria elementare dei sistemi di equazioni differenziali. Al solito daremo inizialmente le definizioni nel caso di sistemi del prim'ordine anche se poi ci interesserà l'applicazione al caso del secondo ordine.
L'idea centrale della definizione di punto critico stabile è quella che l'orbita rimanga confinata *vicino* al punto singolare anche partendo da condizioni iniziali, che non coincidono con il punto

singolare, ma se ne discostano un poco.

Definizione 5.3. (**Punto singolare stabile**) In riferimento alla definizione 5.2, un punto singolare $\mathbf{y}_0 \in D$ è detto:

(a) stabile nel futuro, se per ogni intorno aperto $U \ni \mathbf{y}_0$ esiste un intorno aperto $V \ni \mathbf{y}_0$ tale che ogni soluzione di (5.1), $\mathbf{x} : I \to D$ con $\mathbf{x}(0) = \mathbf{x}_0 \in V$, soddisfa $\mathbf{x}(t) \in U$ se $I \ni t \geq 0$;

(b) stabile nel passato, se per ogni intorno aperto $U \ni \mathbf{x}_0$ esiste un intorno aperto $V \ni \mathbf{x}_0$ tale che ogni soluzione di (5.1), $\mathbf{x} : I \to D$ con $\mathbf{x}(0) = \mathbf{x}_0 \in V$, soddisfa $\mathbf{x}(t) \in U$ se $I \ni t \leq 0$;

(c) stabile, se è stabile nel futuro e stabile nel passato;

(d) instabile (rispettivamente **nel futuro** o **nel passato**), se non è stabile (rispettivamente nel futuro o nel passato);

(e) asintoticamente stabile nel futuro, se è stabile nel futuro ed esiste un intorno aperto $A \ni \mathbf{y}_0$ tale che ogni soluzione massimale di (5.1) $\mathbf{x} : I \to D$ con $\mathbf{x}(0) = \mathbf{x}_0 \in A$, è completa nel futuro e soddisfa $\mathbf{x}(t) \to \mathbf{y}_0$ per $t \to +\infty$;

(f) asintoticamente stabile nel passato, se è stabile nel passato ed esiste un intorno aperto $A \ni \mathbf{y}_0$ tale che ogni soluzione massimale di (5.1) $\mathbf{x} : I \to D$ con $\mathbf{x}(0) = \mathbf{x}_0 \in A$, è completa nel passato e soddisfa $\mathbf{x}(t) \to \mathbf{y}_0$ per $t \to -\infty$;

Nei casi (e) e (f), l'insieme A è detto essere un **bacino di attrazione** di $\mathbf{x}_0$. $\diamondsuit$.

Osservazioni 5.3.

(1) Nel caso in cui $\mathbf{f}$ sia localmente lipschitziana, se $\mathbf{y}_0$ è stabile nel futuro, in virtù della proposizione 3.5, le soluzioni massimali con condizione iniziale nell'intorno V sono sicuramente complete nel futuro (possiamo prendere U in modo che abbia chiusura compatta). Lo stesso risultato si ottiene per la stabilità nel passato.

(2) Le definizioni date si generalizzano immediatamente al caso di un sistema del secondo ordine (5.3). In tal caso si parla di **punti di equilibrio stabili** (**nel passato** e **nel futuro**) e **punti di equilibrio asintoticamente stabili** (**nel passato** e **nel futuro**).

(3) Bisogna considerare il fatto che, trattando sistemi del secondo ordine, gli intorni U, V e A devono essere presi nello spazio delle coppie $(\mathbf{x}, \mathbf{v})$ e non solo della variabile $\mathbf{x}$. In questo senso, trattando la stabilità di configurazioni di sistemi fisici, *la nozione di stabilità include informazioni e richieste sulle velocità e non solo sulle configurazioni.*

Esempi 5.1.

1. Si consideri l'equazione differenziale in $\mathbb{R}$, dove $k \neq 0$ è fissato:

$$\frac{dx}{dt} = -kx \, .$$

Siamo nelle ipotesi di validità del teorema di esistenza ed unicità globale essendo il secondo membro di classe C^∞. L'unico punto singolare è $x = 0$. Dato che le soluzioni massimali hanno la forma

$$x(t) = x_0 e^{-kt} \, , \quad \text{con } t \in \mathbb{R}.$$

Se $k > 0$, vale la relazione $|x(t)| < |x_0|$ per $t > 0$. Pertanto il punto stabile nel futuro: per ogni scelta dell'intorno aperto $U \ni 0$, l'intorno $V \ni 0$ può essere preso come un qualsiasi intervallo $(-\delta, \delta) \subset U$. Inoltre $x(t) \to 0$ se $t \to +\infty$. Si conclude quindi che, per $k > 0$, il punto singolare $x = 0$ è asintoticamente stabile nel futuro. D'altra parte vale anche, sempre per $k > 0$: $x(t) \to +\infty$ se $t \to -\infty$. Di conseguenza, se $k > 0$, il punto singolare è sicuramente instabile nel passato. La situazione evidentemente si inverte se $k < 0$: il punto singolare $x = 0$ è asintoticamente stabile nel futuro, ma instabile nel passato.

2. Si consideri il sistema di equazioni differenziali su $\mathbb{R}^2$:

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}\,,$$

dove $\mathbf{x} = (x, y)$ e A è la matrice diagonale $diag(-1, 1)$. Siamo nelle ipotesi di validità del teorema di esistenza ed unicità globale essendo il secondo membro di classe C^∞. L'unico punto singolare è l'origine $\mathbf{0}$ di $\mathbb{R}^2$. Si conclude facilmente che tale punto singolare è instabile nel passato e nel futuro. Infatti, il sistema è in realtà costituito da due equazioni disaccoppiate:

$$\frac{dx}{dt} = -x\,, \quad \frac{dy}{dt} = y\,.$$

Dall'esempio precedente sappiamo che ogni condizione iniziale $\mathbf{x}_0 = (x_0, 0)$ produce una soluzione che diverge (nella sola componente x), per $t \to -\infty$. Questo solo fatto è sufficiente a provare che il punto singolare $\mathbf{0}$ è instabile nel passato. Viceversa, sempre dall'esempio precedente, sappiamo che la condizione iniziale $\mathbf{x}_0 = (0, y_0)$ produce una soluzione che diverge (nella sola componente y), per $t \to +\infty$. Questo solo fatto è sufficiente a provare che il punto singolare $\mathbf{0}$ è instabile nel futuro.

3. Consideriamo l'oscillatore armonico, costituito da un punto materiale di massa $m > 0$ attaccato ad una molla di costante $\kappa > 0$, per comodità di lunghezza nulla a riposo, il cui secondo estremo è fisso nel punto O di un riferimento inerziale. L'equazione del sistema è ovviamente

$$\frac{d^2\mathbf{x}}{dt^2} = -k^2\mathbf{x}\,,$$

dove $k := \sqrt{\kappa/m}$ e $\mathbf{x} = P - O$. Siamo nelle ipotesi di validità del teorema di esistenza ed unicità globale essendo il secondo membro di classe C^∞. Passando al prim'ordine, si ha il sistema equivalente in $\mathbb{R}^6$:

$$\frac{d\mathbf{v}}{dt} = -k\mathbf{x}\,, \quad \frac{d\mathbf{x}}{dt} = \mathbf{v}\,.$$

In coordinate cartesiane ortonormali solidali con il riferimento inerziale il sistema si disaccoppia in 3 sistemi equazioni della forma (ne scriviamo una sola):

$$\frac{dv}{dt} = -kx\,, \quad \frac{dx}{dt} = v\,.$$

Si osservi che l'unica *configurazione di equilibrio* è dunque $x = 0$, ovvero $\mathbf{x} = \mathbf{0}$ tenendo conto delle 3 distinte componenti. La soluzione generale del sistema si determina subito con i metodi

del capitolo 3. Essa è

$$x(t) = A\sin(kt) + B\cos(kt)\ , \quad v(t) = kA\cos(kt) - kB\sin(kt)\ ,$$

dove $A, B \in \mathbb{R}$ sono costanti arbitrarie. Ricostruendo le 3 componenti e fissando le costanti in funzione delle condizioni iniziali, le soluzioni generali sono scrivibili come

$$\mathbf{x}(t) = k^{-1}\sin(kt)\mathbf{v}_0 + \cos(kt)\mathbf{x}_0\ , \quad \mathbf{v}(t) = \cos(kt)\mathbf{v}_0 - k\sin(kt)\mathbf{x}_0\ . \tag{5.7}$$

Da queste si ricava che, tenendo conto che le funzioni seno e coseno oscillano in $[-1, 1]$:

$$k||\mathbf{x}(t)|| \leq ||\mathbf{v}_0|| + k||\mathbf{x}_0||\ , \quad ||\mathbf{v}(t)|| \leq ||\mathbf{v}_0|| + k||\mathbf{x}_0||\ , \quad \text{per ogni } t \in \mathbb{R}\ . \tag{5.8}$$

Dato che su $\mathbb{K}^n$ tutte le norme sono equivalenti (producono la stessa topologia), possiamo usare la norma su $\mathbb{R}^6$:

$$||(\mathbf{x}, \mathbf{v})||' := \max(k||\mathbf{x}||, ||\mathbf{v}||)\ .$$

Le disuguaglianza (5.8) di sopra implica immediatamente che

$$||(\mathbf{x}(t), \mathbf{v}(t))||' \leq 2||(\mathbf{x}(0), \mathbf{v}(0))||'\ , \quad \text{per ogni } t \in \mathbb{R}\ .$$

Questa disuguaglianza, a sua volta, ci assicura che la configurazione $\mathbf{x}_0 = 0$ sia di equilibrio stabile. Infatti, fissato un intorno aperto U di $(\mathbf{x}_0, \mathbf{v}_0) = (\mathbf{0}, \mathbf{0})$, se prendiamo una palla metrica B_δ (rispetto alla norma $||\cdot||'$) aperta centrata in $(\mathbf{0}, \mathbf{0})$ e, con raggio $\delta > 0$ tale che valga l'inclusione $B_\delta \subset U$, scegliendo condizioni iniziali in $B_{\delta/2}$, siamo sicuri che l'orbita della soluzione massimale e completa non uscirà mai da B_δ e quindi, a maggior ragione, dall'aperto U. Ovviamente la configurazione $\mathbf{x}_0 = \mathbf{0}$ non è di equilibrio asintoticamente stabile nel passato e neppure nel futuro, perché non esistono i limiti a $t \to \pm\infty$ dei secondi membri di (5.7), se si escludono quelli valutati sulla soluzione banale.

5.1.2 Introduzione ai metodi di Liapunov per lo studio della stabilità .

Negli esempi visti nella sezione precedente, lo studio della stabilità dei punti singolari è stato sviluppato avendo preventivamente calcolato la forma generale delle soluzioni delle equazioni differenziali in considerazione. Dal punto di vista pratico questa è una procedura con scarsa validità , in quanto la forma esplicita delle soluzioni è nota in pochissimi casi, rarissime volte di interesse applicativo. Per questo motivo sono stati sviluppati teoremi che forniscono condizioni sufficienti per la stabilità del punto critico considerato, che si possano controllare senza dover risolvere l'equazione differenziale.
Introduciamo ora due teoremi, in questo contesto, dovuti a Liapunov che rappresentano solo il caso più elementare della tecnologia matematica sviluppata da Liapunov e altri matematici.

Consideriamo un sistema di equazioni differenziali del prim'ordine autonomo su $D \subset \mathbb{K}^n$ insieme aperto non vuoto,

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t))\ , \tag{5.9}$$

con $\mathbf{f} : D \to \mathbb{K}^n$. Consideriamo poi una funzione $W : U \to \mathbb{K}$ differenziabile con $U \subset D$ sottoinsieme aperto (eventualmente coincidente con D stesso). Se conosciamo una soluzione $\mathbf{x} = \mathbf{x}(t)$ di (5.9), la derivata della funzione composta $t \mapsto W(\mathbf{x}(t))$, ammesso che abbia senso, si scrive:

$$\frac{dW(\mathbf{x}(t))}{dt} = \frac{d\mathbf{x}}{dt} \cdot \nabla W(\mathbf{x})|_{\mathbf{x}(t)} = [\mathbf{f}(\mathbf{x}) \cdot \nabla W(\mathbf{x})]|_{\mathbf{x}=\mathbf{x}(t)} \ .$$

Se definiamo la funzione:

$$\mathring{W}(\mathbf{x}) := \mathbf{f}(\mathbf{x}) \cdot \nabla W(\mathbf{x}) \ , \quad \text{per ogni } \mathbf{x} \in U \ , \tag{5.10}$$

abbiamo allora che, per ogni soluzione $\mathbf{x} = \mathbf{x}(t)$ del sistema (5.9), se il secondo membro dell'equazione che segue è definito, vale:

$$\frac{dW(\mathbf{x}(t))}{dt} = \mathring{W}(\mathbf{x}(t)) \ . \tag{5.11}$$

Osservazioni 5.4.
(1) La funzione $\mathring{W}$ è definita senza dover risolvere esplicitamente l'equazione differenziale, tuttavia contiene informazioni sulle soluzioni.
(2) Data $\mathbf{f} : D \to \mathbb{K}^n$, se $W : U \to \mathbb{K}$ (con $U \subset D$ aperto) è differenziabile, si dice che W è un **integrale primo** del sistema (5.9) su U, se $I \ni t \mapsto W(\mathbf{x}(t))$ è costante per ogni soluzione $I \ni t \mapsto \mathbf{x}(t) \subset U$ di (5.9), al variare di $t \in I$.
Vale il seguente risultato:

Proposizione 5.3. *Se $\mathbf{f} : D \to \mathbb{K}^n$ ($D \subset \mathbb{K}^n$ aperto) è localmente lipschitziana, allora $W : U \to \mathbb{K}$ (con $U \subset D$ aperto) soddisfa, rispettivamente:*
$\mathring{W}(\mathbf{x}) < 0$, $\forall \mathbf{x} \in U$,
$\mathring{W}(\mathbf{x}) \leq 0$, $\forall \mathbf{x} \in U$,
$\mathring{W}(\mathbf{x}) = 0$, $\forall \mathbf{x} \in U$,
se e solo se per ogni soluzione $\mathbf{x} : I \to U$ del sistema (5.9) vale rispettivamente:
$dW(\mathbf{x}(t))/dt < 0$, $\forall t \in I$,
$dW(\mathbf{x}(t))/dt \leq 0$, $\forall t \in I$,
$dW(\mathbf{x}(t))/dt = 0$, $\forall t \in I$.
Vale un analoga proposizione scambiando ovunque $<$ con $>$.

Dimostrazione. Se $\mathring{W}(\mathbf{x}) < 0$ su U allora, da (5.11), segue che $dW(\mathbf{x}(t))/dt < 0$ per ogni soluzione $\mathbf{x} : I \to U$ del sistema (5.9). Viceversa, se vale $dW(\mathbf{x}(t))/dt < 0$ per ogni soluzione $\mathbf{x} : I \to U$ del sistema (5.9), deve valere, per (5.11), $\mathring{W}(\mathbf{x}_0) < 0$ se $\mathbf{x}_0 \in U$ è un punto attraversato dall'orbita di una soluzione del sistema (5.9). La locale lipschitzianità di $\mathbf{f}$ assicura che per ogni punto $\mathbf{x}_0 \in U$ c'è una soluzione di (5.9) che ha come condizione iniziale $\mathbf{x}(0) = \mathbf{x}_0$. L'orbita di tale soluzione attraversa, per definizione, $\mathbf{x}_0$. Gli altri casi si provano nello stesso modo. □

Ora che abbiamo definito la funzione $\mathring{W}$ per ogni assegnata funzione $W : U \to \mathbb{K}$, possiamo enunciare il primo teorema di Liapunov-Barbasin.

D'ora in poi, se $W : A \to \mathbb{R}$ è una funzione con $A \subset \mathbb{K}^n$, diremo che essa ha un **minimo stretto** in $\mathbf{x}_0 \in A$ se $F(\mathbf{x}) > F(\mathbf{x}_0)$ per $\mathbf{x} \in A$ con $\mathbf{x} \neq \mathbf{x}_0$.

Teorema 5.1. (**di Liapunov-Barbasin**) *Se $D \subset \mathbb{K}^n$ è un aperto, sia $\mathbf{x}_0 \in D$ un punto singolare per la funzione localmente lipschitziana $\mathbf{f} : D \to \mathbb{K}^n$.*
(a) *$\mathbf{x}_0$ è un punto singolare stabile nel futuro se in un'intorno aperto di $\mathbf{x}_0$, $G_0 \subset D$ è definita una funzione differenziabile $W : G_0 \to \mathbb{R}$* [1] *che soddisfa:*
(i) W ha un minimo stretto in $\mathbf{x}_0$,
(ii) $\mathring{W}(\mathbf{x}) \leq 0$ per $\mathbf{x} \in G_0$.
(b) *$\mathbf{x}_0$ è un punto singolare stabile nel passato se in un'intorno aperto di $\mathbf{x}_0$, $G_0 \subset D$ è definita una funzione differenziabile $W : G_0 \to \mathbb{R}$ che soddisfa:*
(i) W ha un minimo stretto in $\mathbf{x}_0$,
(ii) $\mathring{W}(\mathbf{x}) \geq 0$ per $\mathbf{x} \in G_0$.
(c) *$\mathbf{x}_0$ è un punto singolare stabile se in un'intorno aperto di $\mathbf{x}_0$, $G_0 \subset D$ è definita una funzione differenziabile $W : G_0 \to \mathbb{R}$ che soddisfa:*
(i) W ha un minimo stretto in $\mathbf{x}_0$,
(ii) W è un integrale primo su G_0 per il sistema associato a $\mathbf{f}$. $\diamondsuit$

Dimostrazione La dimostrazione di (b) è del tutto analoga a quella di (a). (c) segue immediatamente da (a), da (b) e dalla proposizione 5.3.
Dimostriamo (a). Per comodità lavoreremo con un sistema di coordinate cartesiane su $\mathbb{K}^n$ centrate nel punto $\mathbf{x}_0$ che di conseguenza diventa il punto $\mathbf{0}$. Consideriamo un intorno aperto $U \ni \mathbf{x}_0$ e supponiamo $U \subset G_0$. Dobbiamo esibire il secondo intorno aperto $V \ni \mathbf{0}$, tale che prendendo soluzioni con condizioni iniziali in V, la loro orbita non esce da U. Definiamo V come segue. Sia $B \subset U$ una palla aperta centrata in $\mathbf{0}$ con $\overline{B} \subset U$. Dato che W è continua sul compatto ∂B, poniamo:

$$\ell := \min \{W(\mathbf{x}) \ |\mathbf{x} \in \partial B \ \} \ .$$

Definiamo V come una palla aperta centrata in $\mathbf{0}$ e con $V \subset B \cap \{\mathbf{x} \in G_0 \,|\, W(\mathbf{x}) < \ell\}$. Un tale V esiste perché l'insieme $\{\mathbf{x} \in B \ | \ W(\mathbf{x}) < \ell\}$ è sicuramente aperto (controimmagine di un aperto secondo una funzione continua) e non vuoto, dato che contiene almeno $\mathbf{0}$, avendo W un minimo stretto in $\mathbf{0} \in B$. Mostriamo che V soddisfa la proprietà che ogni soluzione con condizione iniziale in V ha orbita che non esce da B e quindi da U. Supponiamo che la soluzione $\mathbf{x} = \mathbf{x}(t)$, con $t \in I$ intervallo aperto (che include 0) e con con $\mathbf{x}(0) \in V$, abbia un punto $\mathbf{x}(t_1) \notin B$ per $t_1 > 0$. Se $R > 0$ è il raggio di B, la funzione continua $I \ni t \mapsto R - ||\mathbf{x}(t)||$ è positiva per $t = 0$ e non positiva per $t = t_1$, di conseguenza si annulla per qualche $t_2 \in (0, t_1]$. Quindi $\mathbf{x}(t_2) \in \partial B$ e dunque, per definizione di ℓ:

$$W(\mathbf{x}(t_2)) \geq \ell \ . \tag{5.12}$$

D'altra parte, per ipotesi $[0, t_2] \ni t \mapsto W(\mathbf{x}(t))$ è non crescente essendo, nelle nostre ipotesi:

$$\frac{dW(\mathbf{x}(t))}{dt} = \mathring{W}(\mathbf{x}(t)) \leq 0 \ ,$$

[1] La funzione W assume valori in $\mathbb{R}$ anche se lavoriamo su $\mathbb{K}^n = \mathbb{C}^n$.

di conseguenza deve anche essere $W(\mathbf{x}(t_2)) \leq W(\mathbf{x}(0)) < \ell$ e quindi

$$W(\mathbf{x}(t_2)) < \ell\,,$$

che contraddice (5.12). L'assurdo trovato prova che $\mathbf{x} = \mathbf{x}(t)$ rimane confinata in $B \subset U$ per ogni $t \in I$. $\square$

La funzione W usata nel teorema viene chiamata **funzione di Liapunov**. Tutta la tecnica di Liapunov per studiare la stabilità è bastata su simili funzioni. Possiamo completare il teorema precedente con il seguente risultato che si occupa della stabilità asintotica.

Teorema 5.2. (**Liapunov-Barbasin 2**) *Nelle ipotesi del teorema 5.1, se in aggiunta alle ipotesi in (a) oppure (b) vale più fortemente, rispettivamente,* $\mathring{W}(\mathbf{x}) < 0$ *per* $\mathbf{x} \neq \mathbf{x}_0$ *oppure* $\mathring{W}(\mathbf{x}) > 0$ *per* $\mathbf{x} \neq \mathbf{x}_0$*, allora* $\mathbf{x}_0$ *è asintoticamente stabile rispettivamente nel futuro oppure nel passato.* $\diamondsuit$

Dimostrazione. Dimostriamo la stabilità asintotica nel futuro, quella nel passato si prova in modo del tutto analogo. Ridefinendo W per una costante additiva, possiamo ridurci sempre ad assumere $W(\mathbf{x}_0) = 0$. Sia B una palla aperta centrata in $\mathbf{x}_0$ con $\overline{B} \subset G_0$. Se $\epsilon > 0$, definiamo l'insieme aperto e non vuoto (contiene x_0):

$$A_\epsilon := \{\mathbf{x} \in B \ \ |W(x) < \epsilon\}\,.$$

Consideriamo una soluzione massimale $\mathbf{x} = \mathbf{x}(t)$ con condizione iniziale in A_1. Prima di tutto notiamo che l'orbita non può raggiungere $\mathbf{x}_0$ in un tempo finito T, perché il problema di Cauchy con condizione iniziale $\mathbf{x}(T) = \mathbf{0}$ ha come unica soluzione massimale (quindi anche nel passato di T), quella costante $\mathbf{x}(t) = \mathbf{0}$ per ogni $t \in \mathbb{R}$. Quindi l'orbita rimarrà in un insieme in cui $\mathring{W} < 0$. Dato che $\mathring{W} < 0$, allora $dW(\mathbf{x}(t))/dt < 0$, per cui la funzione $t \mapsto W(\mathbf{x}(t))$ è strettamente decrescente. In particolare deve esistere il suo limite ℓ quando $t \to +\infty$. Se fosse $\ell > 0$, fissato $\epsilon > 0$ sufficientemente piccolo, a partire da un certo $T > 0$, l'orbita sarebbe sempre confinata nel compatto dato da $\overline{B} \setminus A_\epsilon$ per ϵ sufficientemente piccolo. Ma allora, se $-M = \max\{\mathring{W}(\mathbf{x}) \,|\, \mathbf{x} \in \overline{B} \setminus A_\epsilon\} < 0$, avremmo che $dW(\mathbf{x}(t))/dt < -M$ per ogni $t > T$ e quindi applicando il teorema di Lagrange (dove $\xi \in (T, t)$):

$$W(\mathbf{x}(t)) - W(\mathbf{x}(T)) = (t - T)\frac{dW(\mathbf{x}(t))}{dt}|_\xi \leq -M(t - T) \to -\infty\,, \quad \text{se } t \to +\infty\,.$$

Questo è impossibile dato che W è continua sul compatto $\overline{B} \setminus A_n$ e quindi limitata. Di conseguenza deve essere $W(\mathbf{x}(t)) \to \ell = 0$ per $t \to +\infty$. In altre parole:

$$\text{per ogni } \epsilon > 0 \text{ esiste } T_\epsilon > 0 \text{ tale che, se } t > T_\epsilon, \text{ vale } \mathbf{x}(t) \in A_\epsilon. \tag{5.13}$$

Se B_r denota la palla aperta di raggio r e centrata in x_0, per ogni $n = 1, 2, \ldots$, deve esistere $\epsilon > 0$

sufficientemente piccolo, per cui $A_\epsilon \subset B_{1/n}$[2]. Di conseguenza e tenendo conto della proposizione (5.13) concludiamo che: per ogni $n > 0$ esiste T_n per cui $\mathbf{x}(t) \in B_{1/n}$ se $t > T_n$. Questo significa che: $\mathbf{x}(t) \to \mathbf{x}_0$ se $t \to +\infty$. $\square$

Esempi 5.2.
1. Consideriamo ancora il punto singolare $x = 0$ dell'equazione differenziale in $\mathbb{R}$, dove $k \neq 0$ è fissato:

$$\frac{dx}{dt} = -kx\,.$$

La funzione $W(x) := x^2$, definita su $\mathbb{R}$, soddisfa: $W(x) > 0$ se $x \neq 0$, per cui ha un minimo stretto in $x = 0$, inoltre $\mathring{W} = -2kx^2 < 0$ se $x \neq 0$. Possiamo applicare il primo teorema di Liapunov concludendo che $x = 0$ è punto singolare stabile nel futuro. Dal secondo teorema di Liapunov concludiamo che il punto singolare è anche asintoticamente stabile nel futuro.
2. Consideriamo l'equazione differenziale in $\mathbb{R}^2$:

$$\frac{d\mathbf{x}}{dt} = -\sin(||\mathbf{x}||^2)A\mathbf{x}\,.$$

dove:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\,.$$

Dato che $\det A = 3 > 0$ e $\mathrm{tr}A = 4 > 0$, gli autovalori della matrice simmetrica A sono strettamente positivi e pertanto $A\mathbf{x} = \mathbf{0}$ se e solo se $\mathbf{x} = \mathbf{0}$. $\mathbf{0}$ è un punto singolare del sistema. Possiamo studiarne la stabilità tenendo conto che ora le soluzioni esplicite non si riescono a scrivere. La funzione $W(\mathbf{x}) = ||\mathbf{x}||^2$ definita su $\mathbb{R}^2$ ha un minimo stretto nell'origine, inoltre

$$\mathring{W}(\mathbf{x}) = -2\sin(||\mathbf{x}||^2)\mathbf{x}\cdot A\mathbf{x}$$

è strettamente negativa in un intorno dell'origine escludendo l'origine stessa. Infatti $\sin(||\mathbf{x}||^2) > 0$ in un intorno di $\mathbf{0}$ se $\mathbf{x} \neq 0$, inoltre la forma quadratica $\mathbf{x}\cdot A\mathbf{x}$ è strettamente positiva avendo, la matrice simmetrica A, autovalori positivi.
Concludiamo che l'origine è un punto singolare stabile nel futuro ed asintoticamente stabile nel futuro.
3. Consideriamo ancora l'equazione differenziale, questa volta in $\mathbb{C}^2$:

$$\frac{d\mathbf{x}}{dt} = -\sin(||\mathbf{x}||^2)A\mathbf{x}\,.$$

dove ora:

$$A = \begin{bmatrix} 2 & 1+i \\ 1-i & 2 \end{bmatrix}\,.$$

[2]Se ciò non fosse, significherebbe che fuori da una certa palla $B_{1/n*}$ esisterebbe una successione di punti $\mathbf{x}_1, \mathbf{x}_2, \ldots$ con $W(\mathbf{x}_k) \to 0$ per $k \to +\infty$. Dato che gli $\mathbf{x}_k$ apparterrebbero al compatto $K := \overline{B} \setminus B_{1/n*}$, esisterebbe una sottosuccessione $\{\mathbf{x}_{k_m}\}_m$ con $\mathbf{x}_{k_m} \to \mathbf{x}* \in K$ per $m \to +\infty$. Per continuità sarebbe anche $W(\mathbf{x}^*) = 0$. Dato che $\mathbf{x}^* \neq \mathbf{x}_0$, W non potrebbe quindi avere il minimo stretto in x_0.

Vale ancora $\det A = 2 > 0$ e $\mathrm{tr} A = 4 > 0$, gli autovalori della matrice hermitiana A sono pertanto strettamente positivi e $A\mathbf{x} = \mathbf{0}$ se e solo se $\mathbf{x} = \mathbf{0}$. $\mathbf{x} = 0$ è ancora un punto singolare. Se $\mathbf{x}^*$ indica il vettore le cui componenti sono le complesse coniugate delle componenti del vettore $\mathbf{x}$ e il puntino $\cdot$ indica l'ordinario prodotto "righe per colonne", la funzione a valori reali $W(\mathbf{x}) = \mathbf{x}^* \cdot \mathbf{x} = ||\mathbf{x}||^2$, definita su $\mathbb{C}^2$, ha un minimo stretto nell'origine. Inoltre

$$\mathring{W}(\mathbf{x}) = \sin(||\mathbf{x}||^2)\,(\mathbf{x}^* \cdot A\mathbf{x} + \mathbf{x} \cdot A\mathbf{x}^*)\ ,$$

è strettamente negativa in un intorno dell'origine escludendo l'origine stessa. Infatti vale, in un intorno di $\mathbf{0}$, $\sin(||\mathbf{x}||^2) > 0$ se $\mathbf{x} \neq 0$, inoltre la forma quadratiche hermitiana $\mathbf{x}^* \cdot A\mathbf{x}$ è strettamente positiva avendo, la matrice hermitiana A, autovalori positivi, infine, se Rez denota la parte reale del numero complesso $z \in \mathbb{C}$:

$$(\mathbf{x}^* \cdot A\mathbf{x} + \mathbf{x} \cdot A\mathbf{x}^*) = 2Re(\mathbf{x}^* \cdot A\mathbf{x})\ .$$

Concludiamo che l'origine è un punto singolare stabile nel futuro ed asintoticamente stabile nel futuro.

5.1.3 *Ancora sulla stabilità asintotica.

Dimostriamo un secondo teorema che generalizza l'enunciato del precedente teorema dando una condizione sufficiente più debole per la stabilità asintotica. Introduciamo alcune nozioni tecniche di grande utilità .

Definizione 5.4. (**Insiemi** $\Lambda_\pm(\mathbf{x}_0)$) Si consideri un sistema di equazioni differenziali del prim'ordine autonomo su $D \subset \mathbb{K}^n$ insieme aperto non vuoto,

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t))\ , \tag{5.14}$$

con $\mathbf{f} : D \to \mathbb{K}^n$ localmente lipschitziana. Indichiamo con $I \ni t \mapsto \mathbf{x}(t|\mathbf{x}_0)$, la soluzione massimale di (5.14) con condizione iniziale $\mathbf{x}_0 \in D$.
(a) Se $I \ni t \mapsto \mathbf{x}(t|\mathbf{x}_0)$ è completa nel futuro, si definisce

$$\Lambda_+(\mathbf{x}_0) := \{\mathbf{y} \in D \mid \mathbf{x}(t_k|\mathbf{x}_0) \to \mathbf{y}\ ,\ \text{se } k \to +\infty \text{ per qualche } \{t_k\}_{k\in\mathbb{N}} \subset I \text{ con } t_k \to +\infty \text{ se } k \to +\infty\}\ .$$

(b) Se $I \ni t \mapsto \mathbf{x}(t|\mathbf{x}_0)$ è completa nel passato, si definisce

$$\Lambda_-(\mathbf{x}_0) := \{\mathbf{y} \in D \mid \mathbf{x}(t_k|\mathbf{x}_0) \to \mathbf{y}\ ,\ \text{se } k \to +\infty \text{ per qualche } \{t_k\}_{k\in\mathbb{N}} \subset I \text{ con } t_k \to -\infty \text{ se } k \to +\infty\}\ .$$

$\diamondsuit$

Si osservi che $\Lambda_\pm(\mathbf{x}_0)$ può essere l'insieme vuoto.

Proposizione 5.4. *In riferimento alla definizione 5.4, se $\Lambda_+(\mathbf{x}_0)$ ($\Lambda_-(\mathbf{x}_0)$) è non vuoto allora valgono i seguenti fatti.*
(a) *Ogni soluzione massimale con condizione iniziale in $\Lambda_+(\mathbf{x}_0)$ (risp. $\Lambda_-(\mathbf{x}_0)$) è completa.*
(b) *L'orbita di ogni soluzione massimale con condizione iniziale in $\Lambda_+(\mathbf{x}_0)$ (risp. $\Lambda_-(\mathbf{x}_0)$) è un sottoinsieme di $\Lambda_+(\mathbf{x}_0)$ (risp. $\Lambda_-(\mathbf{x}_0)$).* $\diamondsuit$

Dimostrazione. Facciamo la dimostrazione per $\Lambda_+(\mathbf{x}_0)$, l'altro caso è analogo. Consideriamo $\{t_k\}_{k\in\mathbb{N}} \subset I$ con $\mathbf{x}(t_k|\mathbf{x}_0) \to \mathbf{y}_0$ per $k \to +\infty$. Fissato $t \in \mathbb{R}$, esisterà $k \in \mathbb{N}$ sufficientemente grande tale che $t + t_k > 0$. In tal caso ha senso scrivere (si noti che vale il teorema di unicità): $\mathbf{x}(t + t_k|\mathbf{x}_0) = \mathbf{x}(t|\mathbf{x}(t_k|\mathbf{x}_0))$. Prendendo il limite per $k \to +\infty$, abbiamo che il secondo membro tende a $\mathbf{x}(t|\mathbf{y}_0)$ per (b) in teorema 3.8. Questo prova che esiste il limite per $k \to +\infty$ di $\mathbf{x}(t + t_k|\mathbf{x}_0)$ ed è :

$$\lim_{k\to+\infty} \mathbf{x}(t + t_k|\mathbf{x}_0) = \mathbf{x}(t|y_0) \, .$$

In particolare, dato che t è arbitrario in $\mathbb{R}$, la soluzione $t \mapsto \mathbf{x}(t|\mathbf{y}_0)$ è completa ed (a) è in tal modo dimostrato. Per provare (b) si osservi che, definiti i numeri $t'_k := t + t_k$, vale:

$$\mathbf{x}(t|\mathbf{y}_0) = \lim_{k\to+\infty} \mathbf{x}(t'_k|\mathbf{x}_0) \, ,$$

ma allora, per definizione di $\Lambda_+(\mathbf{x}_0)$, si ha che $\mathbf{x}(t|y_0) \in \Lambda_+(\mathbf{x}_0)$ per ogni $t \in \mathbb{R}$. $\square$

Possiamo allora provare un terzo teorema di Liapunov.

Teorema 5.3. (**Liapunov-Barbasin 3**) *Nelle ipotesi del teorema 5.1, se in aggiunta alle ipotesi in (a) oppure (b) oppure (c), per qualche $\epsilon > 0$, l'insieme*

$$A_\epsilon := \{\mathbf{x} \in G_0 \ | \ 0 \neq ||\mathbf{x} - \mathbf{x}_0|| \leq \epsilon \, , \ \mathring{W}(\mathbf{x}) = 0\}$$

non include completamente soluzioni massimali complete, allora $\mathbf{x}_0$ è asintoticamente stabile, rispettivamente, nel futuro, nel passato o in entrambi. $\diamondsuit$

***Dimostrazione**. La proposizione da provare equivale all'affermazione che: se $\mathbf{x}_0$ è stabile, ma non asintoticamente stabile (nei tre casi considerati), allora ogni insieme A_ϵ, con $\epsilon > 0$, contiene almeno un'orbita di una soluzione massimale completa. Dimostriamo questa affermazione. Supporremo, per fissare le idee che $\mathbf{x}_0$ non sia asintoticamente stabile *nel futuro*, gli altri casi sono analoghi.
Se non c'è stabilità asintotica nel futuro, fissata una palla B_{ϵ_0} con centro $\mathbf{x}_0$ e di raggio $\epsilon_0 > 0$ piccolo a piacere, esiste una condizione iniziale $\mathbf{y}_0 \in B_{\epsilon_0}$ per cui la soluzione massimale associata $t \mapsto \mathbf{x}(t|\mathbf{y}_0)$ non tende a $\mathbf{x}_0$ per $t \to +\infty$. Si osservi che $t \mapsto \mathbf{x}(t|\mathbf{y}_0)$ è comunque confinabile in una palla a chiusura compatta B_ϵ, con $\epsilon > \epsilon_0 > 0$, per la stabilità nel futuro. Questo implica la completezza nel futuro di $t \mapsto \mathbf{x}(t|\mathbf{y}_0)$ per (a) in proposizione 3.5. Il fatto che $\mathbf{x}(t|\mathbf{y}_0) \not\to \mathbf{x}_0$ per $t \to +\infty$ implica che, comunque fissiamo t_k grande a piacere, c'è un tempo successivo, t_{k+1}, per cui $\mathbf{x}(t_{k+1}|\mathbf{y}_0) \notin B_{\epsilon_0}$. In definitva ci sarà una successione $t_1 < t_2 < \cdots < t_k \to +\infty$ per cui

ogni $\mathbf{x}(t_k|\mathbf{y}_0)$ è confinato nel compatto $K := \overline{B_\epsilon \setminus B_{\epsilon_0}}$. Si potrà allora estrarre una sottosuccessione $\{\mathbf{x}(t_{k_r}|\mathbf{y}_0)\}_{r\in\mathbb{N}}$ convergente a $\mathbf{x}' \in K$. Per definizione $\mathbf{x}' \in \Lambda_+(\mathbf{y}_0)$. Dato che $W(\mathbf{x}(t|\mathbf{y}_0))$ è continua e non decrescente, concludiamo allora che deve essere

$$\lim_{t\to+\infty} W(\mathbf{x}(t|\mathbf{y}_0)) = \lim_{r\to+\infty} W(\mathbf{x}(t_{t_r}|\mathbf{y}_0)) = W(\mathbf{x}')\,,$$

per cui:

$$W(\Lambda_+(\mathbf{y}_0)) = W(\mathbf{x}') \quad \text{costantemente}\,. \tag{5.15}$$

Quest'ultimo fatto esclude che $\mathbf{x}_0 \in \Lambda_+(\mathbf{y}_0)$ dato che $W(\mathbf{x}') > W(\mathbf{x}_0)$, in quanto W ha un minimo stretto in $\mathbf{x}_0$.
Consideriamo infine $\mathbf{z} \in \mathbf{\Lambda}_+(\mathbf{y_0})$. Sappiamo che la soluzione massimale $t \mapsto \mathbf{x}(t|\,\mathbf{z})$ è completa e che $\mathbf{x}(t|\,\mathbf{z}) \in \mathbf{\Lambda}_+(\mathbf{y_0})$ per ogni $t \in \mathbb{R}$, per la proposizione 5.4. D'altra parte, per (5.15),

$$\frac{dW(\mathbf{x}(t|\,\mathbf{z}))}{dt} = 0\,,$$

ossia:

$$\mathring{W}(\mathbf{x}(t|\,\mathbf{z})) = \mathbf{0}\,, \text{per ogni } t \in \mathbb{R}.$$

In definitiva, in

$$A_\epsilon := \{\mathbf{x} \in G_0 \;|\; 0 \neq ||\mathbf{x}-\mathbf{x}_0|| \leq \epsilon\,,\;\; \mathring{W}(\mathbf{x}) = 0\}$$

è contenuta completamente l'orbita di una soluzione massimale completa: $\mathbf{x} = \mathbf{x}(t|\,\mathbf{z})$ con $t \in \mathbb{R}$.
□

Osservazioni 5.5. Se la funzione di Liapunov del teorema 5.1 soddisfa più fortemente $\mathring{W}(\mathbf{x}) < 0$ per $\mathbf{x} \neq \mathbf{x}_0$ oppure $\mathring{W}(\mathbf{x}) > 0$ per $\mathbf{x} \neq \mathbf{x}_0$ tutti gli insiemi A_ϵ sono vuoti, per cui ricadiamo banalmente nelle ipotesi del teorema 5.3. Questo prova che il teorema 5.2 è in realtà un corollario del teorema 5.3.

5.1.4 Un criterio per l'instabilità basato sulla procedura di linearizzazione.

Per concludere diamo un semplice criterio di sufficienza, non basato sulle funzioni di Liapunov, per stabilire se un punto singolare è instabile. L'idea è quella di "linearizzare" il sistema intorno al punto singolare, studiare il sistema linearizzato ed infine estendere il risultato trovato al sistema completo.
Consideriamo il solito sistema del prim'ordine autonomo, definito su D sottoinsieme aperto non vuoto di $\mathbb{K}^n$:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x})\,. \tag{5.16}$$

Supponiamo che $\mathbf{f}$ sia di classe $C^1(D)$ e che $\mathbf{x}_0$ sia un punto singolare. Usando lo sviluppo di Taylor centrato in $\mathbf{x}_0$ possiamo scrivere che:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} + \nabla\mathbf{f}|_{\mathbf{x}_0} \cdot (\mathbf{x}-\mathbf{x}_0) + ||\mathbf{x}-\mathbf{x}_0||\mathbf{O}(\mathbf{x}-\mathbf{x}_0)\,,$$

dove la funzione $\mathbf{O}$ soddisfa $\mathbf{O}(\mathbf{x}) \to \mathbf{0}$ se $\mathbf{x} \to \mathbf{0}$. Senza perdere generalità , possiamo sempre supporre di cambiare variabili passando a $\mathbf{y} := \mathbf{x} - \mathbf{x}_0$, in modo che lo sviluppo di sopra si scriva:

$$\mathbf{f}(\mathbf{x}) = A\mathbf{y} + ||\mathbf{y}||\mathbf{O}(\mathbf{y}) \,, \tag{5.17}$$

dove trasformazione lineare $A : \mathbb{K}^n \to \mathbb{K}^n$ è definita da $A\mathbf{y} := \nabla \mathbf{f}|_{\mathbf{x}_0} \cdot \mathbf{y}$. Il sistema di equazioni differenziali:

$$\frac{d\mathbf{y}}{dt} = A\mathbf{y} \,, \tag{5.18}$$

dove $A\mathbf{y} := \mathbf{y} \cdot \nabla \mathbf{f}|_{\mathbf{x}_0}$ per ogni $\mathbf{y} \in \mathbb{K}^n$ è detto **sistema linearizzato attorno al punto singolare** $\mathbf{x}_0$ associato al sistema (5.16) .

Abbiamo ora un risultato semplice ma importante.

Proposizione 5.5. *Si consideri il sistema di equazioni differenziali su* $\mathbb{K}^n$*:*

$$\frac{d\mathbf{y}}{dt} = A\mathbf{y} \,,$$

dove A *è una qualsiasi matrice* $n \times n$ *a coefficienti in* $\mathbb{K}$*.*
(a) *Se* A *ha un autovalore con parte reale positiva, il punto singolare* $\mathbf{y}_0 = 0$ *è instabile nel futuro.*
(b) *Se* A *ha un autovalore con parte reale negativa, il punto singolare* $\mathbf{y}_0 = 0$ *è instabile nel passato.* $\diamondsuit$

Dimostrazione. Proviamo (a), dato che la prova di (b) è del tutto analoga. Esisterà $\mathbf{u} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ tale che $Au = \lambda u$ e $Re\lambda > 0$.
Supponiamo dapprima che $\mathbb{K} = \mathbb{C}$. Si consideri, per $\delta > 0$ piccolo a piacere, la funzione:

$$\mathbf{x}(t) = \delta e^{\lambda t}\mathbf{u} \,, \quad \text{con } t \in \mathbb{R} \,. \tag{5.19}$$

Essa soddisfa $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$, con condizione iniziale $\mathbf{y}(0) = \delta\mathbf{u}$ come è immediato verificare (ed è anche l'unica soluzione massimale con queste condizioni iniziali, dato che la funzione $\mathbf{y} \mapsto A\mathbf{y}$ è C^∞). D'altra parte, essendo $Re\lambda > 0$, vale anche:

$$||\mathbf{x}(t)||^2 = \delta^2(\mathbf{u}^* \cdot \mathbf{u})e^{2(Re\lambda)t} \to +\infty \quad \text{se } t \to +\infty.$$

Fissato un intorno aperto $U \ni \mathbf{0}$, comunque scegliamo piccolo un secondo intorno aperto $V \ni \mathbf{0}$, esso conterrà sempre un punto $\delta\mathbf{u}$, per cui la condizione iniziale $\mathbf{y}(0) = \delta\mathbf{u}$ produce una soluzione che diverge per $t \to +\infty$ e quindi, a maggior ragione, la cui orbita esce da U, per t sufficientemente grande. Quindi il punto singolare $\mathbf{0}$ non può essere stabile nel futuro.
Consideriamo ora il caso $\mathbb{K} = \mathbb{R}$. La soluzione (5.19) non si può più considerare in quanto assume valori in $\mathbb{C}^n$ mentre noi ora vogliamo lavorare in $\mathbb{R}^n$. Dato che A è reale, valendo $A\mathbf{u} = \lambda\mathbf{u}$ (con

$\mathbf{u} \neq \mathbf{0}$) varrà anche $A\mathbf{u}^* = \lambda^*\mathbf{u}^*$. Moltiplichiamo $\mathbf{u}$ per un numero complesso, se necessario, in modo che $\mathbf{u} + \mathbf{u}^* \neq \mathbf{0}$ e consideriamo infine la funzione a valori in $\mathbb{R}^n$:

$$\mathbf{x}(t) = \delta e^{\lambda t}\mathbf{u} + \delta e^{\lambda^* t}\mathbf{u}^* \,, \quad \text{con } t \in \mathbb{R} \,.$$

Si verifica immediatamente che è una soluzione della nostra equazione differenziale con condizione iniziale $\mathbf{x}(0) = \delta(\mathbf{u} + \mathbf{u}^*)$. D'altra parte

$$||\mathbf{x}(t)||^2 = \delta^2 \left(e^{\lambda t}\mathbf{u} + e^{\lambda^* t}\mathbf{u}^*\right) \cdot \left(e^{\lambda t}\mathbf{u} + e^{\lambda^* t}\mathbf{u}^*\right) = \delta^2 \left[2e^{2(Re\lambda)t}||\mathbf{u}||^2 + 2Re\left(\mathbf{u}\cdot\mathbf{u}e^{2\lambda t}\right)\right]$$

$$= \delta^2 e^{2(Re\lambda)t}\left[2||\mathbf{u}||^2 + 2Re\left(\mathbf{u}\cdot\mathbf{u}e^{2iIm(\lambda)t}\right)\right]$$

$$= \delta^2 e^{2(Re\lambda)t}\left[2||\mathbf{u}||^2 + 2\left(Re(\mathbf{u}\cdot\mathbf{u})\cos(2Im(\lambda)t) - Im(\mathbf{u}\cdot\mathbf{u})\sin(2Im(\lambda)t)\right)\right] \,.$$

Quindi:

$$||\mathbf{x}(t)||^2 = 2\delta^2 e^{2(Re\lambda)t}\left[||\mathbf{u}||^2 + |\mathbf{u}\cdot\mathbf{u}|\cos\left(2Im(\lambda)t + \phi\right)\right] \,,$$

dove l'angolo ϕ è definito da: $|\mathbf{u}\cdot\mathbf{u}|\cos\phi = Re(\mathbf{u}\cdot\mathbf{u})$ e $|\mathbf{u}\cdot\mathbf{u}|\sin\phi = Im(\mathbf{u}\cdot\mathbf{u})$.
Nel caso in cui $Im\lambda = 0$, le formule trovate[3] implicano che $||\mathbf{x}(t)||^2 \to +\infty$ per $t \to +\infty$. Se viceversa $Im\lambda \neq 0$, vale che $||\mathbf{x}(t_k)||^2 \to +\infty$ per $k \to +\infty$, se abbiamo definito $t_k := \frac{2\pi k - \phi}{2Im\lambda}$ con $k \in \mathbb{N}$.
In entrambi i casi, fissato un intorno aperto $U \ni \mathbf{0}$, comunque scegliamo piccolo un secondo intorno aperto $V \ni \mathbf{0}$, esso conterrà sempre un punto $\delta(\mathbf{u} + \mathbf{u}^*)$, per cui la condizione iniziale $\mathbf{y}(0) = \delta(\mathbf{u} + \mathbf{u}^*)$ produce una soluzione $\mathbf{x} = \mathbf{x}(t)$ con $\mathbf{x}(t_k) \notin U$ per una serie infinita di valori $t_k \to +\infty$. Quindi $\mathbf{0}$ non può essere stabile nel futuro. $\square$

La proposizione si può estendere, in un teorema, valido per i sistemi non lineari. Di fatto l'instabilità di un punto singolare per un sistema linearizzato implica l'instabilità dello stesso punto per il sistema completo a cui è associato il sistema linearizzato. Vale infatti il seguente teorema [Malkin]:

Teorema 5.4. *Si consideri il sistema di equazioni differenziali del prim'ordine, autonomo, con $\mathbf{f} : D \to \mathbb{K}^n$ localmente lipschitziana, per $D \subset \mathbb{K}$ aperto non vuoto:*

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t)) \,.$$

Se $\mathbf{x}_0 \in D$ è un punto singolare e valgono le seguenti condizioni:

(i) $\mathbf{f}(\mathbf{x}) = \nabla\mathbf{f}|_{\mathbf{x}_0} \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{R}(\mathbf{x})$ con $||\mathbf{R}(\mathbf{x})|| \leq K||\mathbf{x} - \mathbf{x}_0||^{1+\alpha}$ in un intorno di $\mathbf{x}_0$ per qualche $K \geq 0$ e $\alpha > 0$,

(ii) La matrice $A = \nabla\mathbf{f}|_{\mathbf{x}_0}$ ha almeno un autovalore con parte reale positiva o, rispettivamente, negativa;

[3] Non può essere $||\mathbf{u}||^2 + |\mathbf{u}\cdot\mathbf{u}|\cos\phi = 0$ cioé $||\mathbf{u}||^2 + Re(\mathbf{u}\cdot\mathbf{u}) = 0$ perché il calcolo diretto prova che in tal caso sarebbe $\mathbf{u}^* = -\mathbf{u}$ cosa che abbiamo escluso a priori.

allora $\mathbf{x}_0$ *è instabile, rispettivamente, nel futuro o nel passato.* $\diamondsuit$

Osservazioni 5.6.
(1) La condizione (i) è automaticamente soddisfatta se $\mathbf{f} \in C^2(D; \mathbb{K}^n)$. Ciò segue banalmente dallo sviluppo di Taylor fino al secondo ordine. In tal caso, il teorema di Taylor mostra che, per ogni componente f^i di $\mathbf{f}$:

$$f^i(\mathbf{x}) = f^i(\mathbf{x}_0) + \nabla f^i|_{\mathbf{x}_0} \cdot (\mathbf{x} - \mathbf{x}_0) + ||\mathbf{x} - \mathbf{x}_0||^2 \mathbf{n} \cdot \mathbf{R}^i_{\mathbf{x}} \mathbf{n} \,,$$

dove $\mathbf{n}$ denota il versore di $\mathbf{x} - \mathbf{x}_0$ e la matrice $\mathbf{R}^i_{\mathbf{x}}$ definisce una funzione limitata al variare di $\mathbf{x}$ in un intorno di $\mathbf{x}_0$. Si noti che la sola differenziabilità di $\mathbf{f}$ in D non è sufficiente ad assicurare la validità di (i).
(2) Esiste un analogo criterio per studiare la stabilità piuttosto che l'instabilit1à : si dimostra abbastanza facilmente, per il sistema lineare (5.18), che se la matrice A ha *tutti* gli autovalori con parte reale negativa, allora il punto critico $\mathbf{y}_0 = \mathbf{0}$ è asintoticamente stabile nel futuro. Tale risultato si estende ad un sistema non linearizzato esattamente come nel caso del teorema 5.4. Si ha stabilità nel futuro per il punto critico $\mathbf{x}_0$ se vale la condizione (i) e del teorema 5.4, e vale anche che, (ii), tutti gli autovalori di $A = \nabla \mathbf{f}|_{\mathbf{x}_0}$ hanno parte reale negativa [Malkin]. Un analogo teorema vale per la stabilità nel passato richiedendo che tutti gli autovalori di $A = \nabla \mathbf{f}|_{\mathbf{x}_0}$ abbiano parte reale negativa.

5.2 Applicazioni a sistemi di punti materiali con forze conservative.

Applichiamo la teoria sviluppata nella sezione precedente a sistemi fisici particolari. Ci occupiamo di sistemi di punti materiali sottoposti a sistemi di forze conservative, rispetto ad un fissato riferimento $\mathscr{I}$, con l'aggiunta eventuale di forze girostatiche/o dissipative. Più precisamente considereremo un sistema di N punti materiali P_k con masse, rispettivamente, m_k e $k = 1, \ldots, N$. Il sistema di punti è sottoposto ad un sistema di forze conservativo rispetto al sistema di riferimento $\mathscr{I}$ con energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(P_1, \ldots, P_N)$. Ulteriormente, ciascun punto è sottoposto alla forza, che nel riferimento $\mathscr{I}$ ha forma funzionale indipendente dal tempo $\mathbf{F}_k = \mathbf{F}_k(P_1, \ldots, P_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}})$ in modo tale che le due condizioni seguenti sono verificate:

(i) $\mathbf{F}_k = \mathbf{F}_k(P_1, \ldots, P_N, \mathbf{0}, \ldots, \mathbf{0}) = \mathbf{0}$ per ogni $k = 1, \ldots, N$;

(ii) $\displaystyle\sum_{k=1}^{N} \mathbf{F}_k(P_1, \ldots, P_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} \leq 0$ per ogni scelta dei vettori velocità rispetto a $\mathscr{I}$: $\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}} \in V_{\mathscr{I}}$.

Esempi 5.3.
1. L'esempio più semplice per la situazione in esame, si ha quando le forze $\mathbf{F}_k$ sono identicamente nulle ed il sistema è sotto sposto a sole forze conservative. Un esempio concreto è un sistema di

cariche elettriche, trascurando l'irraggiamento, oppure un sistema di pianeti che interagiscono con la forza di gravità .
2. Un esempio meno drastico in cui le forze $\mathbf{F}_k$ non sono nulle, ma soddisfano (i) e (ii) è il sistema di cariche suddetto con l'aggiunta di forze magnetiche dovute ad un campo esterno. Tali forze hanno potenza nulla: $\sum_{k=1}^{N} \mathbf{F}_k(P_1, \ldots, P_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} = 0$, per ogni scelta di $(\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}})$.
3. Un altro esempio è quello di particelle che interagiscono con forze conservative, per esempio dovute a forze elastiche, ed il sistema di punti è immerso in un liquido in quiete nel riferimento $\mathscr{I}$ stesso. In tal caso la potenza delle forze conservative è non positiva e soddisfa più fortemente

$$\sum_{k=1}^{N} \mathbf{F}_k(P_1, \ldots, P_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} < 0 \,,$$

per $(\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \neq (\mathbf{0}, \ldots, \mathbf{0})$.

5.2.1 Il teorema di Lagrange-Dirichlet.

Fissiamo un sistema di coordinate cartesiane ortogonali e solidali con $\mathscr{I}$ con origine O. In questo modo, ogni punto P_k è individuato dal vettore di $\mathbb{R}^3$, $\mathbf{x}_k$ le cui componenti sono quelle del vettore $P_k - O$ nelle coordinate dette. Le equazioni che determinano il moto sono allora, in base al secondo principio della dinamica e riportando al prim'ordine il sistema, dove omettiamo ovunque $|_{\mathscr{I}}$ per semplicità notazionale:

$$\frac{d\mathbf{v}_k}{dt} = -\frac{1}{m_k}\nabla_{\mathbf{x}_k}\mathscr{U}(\mathbf{x}_1, \ldots, \mathbf{x}_n) + \frac{1}{m_k}\mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N)\,, \quad k = 1, \ldots, N\,, \tag{5.20}$$

$$\frac{d\mathbf{x}_k}{dt} = \mathbf{v}_k\,, \quad k = 1, \ldots, N\,. \tag{5.21}$$

Assumeremo che ogni $\mathbf{F}_k \in C^1(\mathbb{R}^{6N}; \mathbb{R}^{3N})$, mentre $\mathscr{U} \in C^2(\mathbb{R}^{3N})$.
Le configurazioni di equilibrio del sistema sono, come detto nella sezione 5.1, le configurazioni $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}) \in \mathbb{R}^{3N}$, tali che, il secondo membro del sistema di sopra si annulla quando è valutato sul vettore di $\mathbb{R}^{3N} \times \mathbb{R}^{3N}$: $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}, \mathbf{0}, \ldots, \mathbf{0})$, dove gli N vettori nulli $\mathbf{0}$ si riferiscono alle velocità $\mathbf{v}_k$. A causa dell'assunzione (i) fatta sopra (e ciò vale in particolare quando le forze non conservative $\mathbf{F}_k$ sono assenti), concludiamo che vale la seguente proposizione.

Proposizione 5.6. *Si consideri di N punti materiali P_k con masse, rispettivamente, $m_k > 0$ e $k = 1, \ldots, N$, le cui equazioni della dinamica nel riferimento $\mathscr{I}$ hanno forma (se $\mathbf{x}_i = P_i - O$ con O in quiete in $\mathscr{I}$):*

$$m_k \frac{d^2\mathbf{x}_k}{dt^2} = -\nabla_{\mathbf{x}_k}\mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) + \mathbf{F}_k\left(\mathbf{x}_1, \ldots, \mathbf{x}_N, \frac{d\mathbf{x}_1}{dt}|_{\mathscr{I}}, \ldots, \frac{d\mathbf{x}_N}{dt}|_{\mathscr{I}}\right)\,, \tag{5.22}$$

dove l'energia potenziale energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ *corrisponde ad un sistema di forze conservative nel riferimento* $\mathscr{I}$ *e le forze non conservative* $\mathbf{F}_k$ *in aggiunta alle precedenti (eventualmente assenti), soddisfano:*

$$\mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{0}, \ldots, \mathbf{0}) = \mathbf{0}\,, \quad \textit{per ogni } k = 1, \ldots, N \textit{ ed ogni } (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in V^N_{\mathscr{I}}\,.$$

Le funzioni $\mathbf{F}_k$ *sono assunte di classe* C^1 *e la funzione* $\mathscr{U}|_{\mathscr{I}}$ *di classe* C^2 *in un sottoinsieme aperto non vuoto rispettivamente di* $V^N_{\mathscr{I}} \times V^N_{\mathscr{I}}$ *e* $V^N_{\mathscr{I}}$.
Condizione necessaria e sufficiente affinché, $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}) \in V^N_{\mathscr{I}}$ *sia una configurazione di equilibrio è che sia un punto di stazionarietà dell'energia potenziale, cioé valga:*

$$\nabla_{\mathbf{x}_k}\mathscr{U}|_{\mathscr{I}}(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}) = \mathbf{0}\,, \quad \text{per ogni } k = 1, \ldots, N. \tag{5.23}$$

$\diamondsuit$

La questione che ci interessa è ora quella di stabilire quando una configurazione di equilibrio sia stabile. Il teorema di Lagrange-Dirichlet che ora, enunceremo e proveremo, costituisce un utilissimo criterio.

Teorema 5.5. (**Lagrange-Dirichlet**) *Si consideri di* N *punti materiali* P_k *con masse, rispettivamente,* $m_k > 0$ *e* $k = 1, \ldots, N$, *le cui equazioni della dinamica nel riferimento* $\mathscr{I}$ *hanno forma (se* $\mathbf{x}_i = P_i - O$ *con* O *in quiete in* $\mathscr{I}$*):*

$$m_k \frac{d^2\mathbf{x}_k}{dt^2} = -\nabla_{\mathbf{x}_k}\mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) + \mathbf{F}_k\left(\mathbf{x}_1, \ldots, \mathbf{x}_N, \frac{d\mathbf{x}_1}{dt}|_{\mathscr{I}}, \ldots, \frac{d\mathbf{x}_N}{dt}|_{\mathscr{I}}\right)\,, \tag{5.24}$$

dove l'energia potenziale energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ *corrisponde ad un sistema di forze conservative nel riferimento* $\mathscr{I}$ *e le forze non conservative* $\mathbf{F}_k$ *in aggiunta alle precedenti (eventualmente assenti), soddisfano:*

$$\mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{0}, \ldots, \mathbf{0}) = \mathbf{0}\,, \quad \textit{per ogni } k = 1, \ldots, N \textit{ ed ogni } (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in V^N_{\mathscr{I}}\,.$$

Le funzioni $\mathbf{F}_k$ *sono assunte di classe* C^1 *e la funzione* $\mathscr{U}|_{\mathscr{I}}$ *di classe* C^2 *in un sottoinsieme aperto non vuoto rispettivamente di* $V^N_{\mathscr{I}} \times V^N_{\mathscr{I}}$ *e* $V^N_{\mathscr{I}}$.
Una configurazione di equilibrio $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N})$ *è :*

(a) *stabile nel futuro se la restrizione di* $\mathscr{U}|_{\mathscr{I}}$ *ad un intorno aperto di di* $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N})$ *ha un minimo stretto in tale punto in aggiunta alla condizione:*
$\sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} \leq 0$ *per ogni scelta dei vettori velocità* $\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}$*;*

(b) *stabile nel passato e nel futuro se la restrizione di* $\mathscr{U}|_{\mathscr{I}}$ *ad un intorno aperto di di* $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N})$ *ha un minimo stretto in* $((\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N})$ *in aggiunta alla condizione:*
$\sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} = 0$ *per ogni scelta dei vettori velocità* $\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}$.

$\diamondsuit$

Dimostrazione. (a) Fissate coordinate cartesiane ortonormali solidali con $\mathscr{I}$ e con origine O, individuiamo i punto P_k tramite i vettori $\mathbf{x}_k \in \mathbb{R}^3$ le cui componenti sono quelle dei corrispondenti vettori $P_k - O$. Le velocità $\mathbf{v}_{P_k}|_{\mathscr{I}}$ saranno biunivocamente individuate da corrispondenti vettori $\mathbf{v}_k \in \mathbb{R}^3$. Consideriamo la funzione che definisce l'energia meccanica in $\mathscr{I}$:

$$\mathscr{E}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N) := \sum_{k=1}^{N} \frac{1}{2} m_k \mathbf{v}_k^2 + \mathscr{U}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,,$$

dove abbiamo omesso ogni specificazione $|_{\mathscr{I}}$ per semplicità notazionale. Sia U un intorno aperto di $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}) \in \mathbb{R}^{3N}$ in cui la restrizione di $\mathscr{U}$ ha un minimo stretto. Se $G_0 := U \times \mathbb{R}^{3N}$, la restrizione di $\mathscr{E}$: $W := \mathscr{E} \restriction_{G_0}$: $G_0 \to \mathbb{R}$ ha evidentemente un minimo stretto nel punto $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}, \mathbf{0}, \ldots, \mathbf{0})$. Inoltre, per il teorema 4.5, su ogni soluzione delle equazioni del moto (5.20)-(5.21), vale:

$$\frac{d}{dt} W(\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t), \mathbf{v}_1(t), \ldots, \mathbf{v}_N(t)) = \sum_{k=1}^{n} \mathbf{F}_K(\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t), \mathbf{v}_1(t), \ldots, \mathbf{v}_N(t)) \cdot \mathbf{v}_k \leq 0 \,.$$

In altre parole, per la proposizione 5.3 su G_0 vale:

$$\mathring{W}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N) \leq 0 \,.$$

Siamo allora nelle ipotesi del teorema 5.1 rispetto alla funzione di Liapunov W, per poter asserire che $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}, \mathbf{0}, \ldots, \mathbf{0})$ è un punto singolare stabile nel futuro e quindi $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N})$ è una configurazione di equilibrio stabile nel futuro.
(b) Procedendo esattamente come nella dimostrazione del caso (a), troviamo che ora vale più fortemente

$$\mathring{W}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N) = 0 \,.$$

Allora, per lo stesso teorema 5.1, in questo caso, la stabilità si ha anche nel passato. $\square$

Osservazioni 5.7. Sottolineamo che la richiesta di avere un punto di minimo stretto dell'energia potenziale in corrispondenza di una configurazione di equilibrio stabile *non* è una condizione necessaria, anche in assenza delle forze non conservative. Un esempio elementare è quello del sistema costituito da un unico punto materiale sulla retta reale sottoposto alla forza conservativa con energia potenziale di classe $C^2(\mathbb{R})$ data da: $\mathscr{U}(x) := x^4 \sin(1/x)$ se $x \neq 0$ e $\mathscr{U}(0) := 0$. La configurazione $x = 0$ è di equilibrio stabile, ma non corrisponde ad un minimo stretto.

Il seguente risultato completa l'enunciato del teorema di Lagrange-Dirichlet, dando condizioni sufficienti per la stabilità asintotica.

Proposizione 5.7. *Nelle ipotesi del teorema 5.5, allora vale:*

(c) *la configurazione di equilibrio* $(\mathbf{x}_{0,1},\ldots,\mathbf{x}_{0,N})$ *è stabile nel futuro e asintoticamente stabile nel futuro se il punto di minimo stretto* $(\mathbf{x}_{0,1},\ldots,\mathbf{x}_{0,N})$ *è anche un punto stazionario* isolato[4] *dell'energia potenziale* $\mathscr{U}_{\mathscr{I}}|$ *e vale la condizione sulle forze non conservative:*
$\sum_{k=1}^{N}\mathbf{F}_k(\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{v}_{P_1}|_{\mathscr{I}},\ldots,\mathbf{v}_{P_N}|_{\mathscr{I}})\cdot\mathbf{v}_{P_k}|_{\mathscr{I}}<0$, *per* $(\mathbf{v}_{P_1}|_{\mathscr{I}},\ldots,\mathbf{v}_{P_N}|_{\mathscr{I}})\neq(\mathbf{0},\ldots,\mathbf{0})$. $\diamondsuit$

Dimostrazione. Useremo le convenzioni e le notazioni usate nella dimostrazione del teorema precedente. Restringiamo l'intorno U (e quindi $G_0=U\times\mathbb{R}^{3N}$) del punto di stazionarietà in modo tale che, in U, non ci siano altri punti stazionari per l'energia potenziale. Vogliamo applicare il teorema 5.3 e pertanto vogliamo inizialmente determinare gli insiemi A_ϵ che appaiono nella tesi del teorema 5.3. Nelle ipotesi fatte, procedendo come nella dimostrazione del teorema precedente, si vede subito che l'insieme dei punti per cui:

$$\mathring{W}(\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{v}_1,\ldots,\mathbf{v}_N)=0$$

coincide con quello per cui vale:

$$\sum_{k=1}^{N}\mathbf{F}_k(\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{v}_1,\ldots,\mathbf{v}_N)\cdot\mathbf{v}_k=0\,.$$

Valendo $\sum_{k=1}^{N}\mathbf{F}_k(P_1,\ldots,P_N,\mathbf{v}_{P_1}|_{\mathscr{I}},\ldots,\mathbf{v}_{P_N}|_{\mathscr{I}})\cdot\mathbf{v}_{P_k}|_{\mathscr{I}}<0$, per $(\mathbf{v}_{P_1}|_{\mathscr{I}},\ldots,\mathbf{v}_{P_N}|_{\mathscr{I}})\neq(\mathbf{0},\ldots,\mathbf{0})$, questo è possibile solo se le velocità dei punti sono tutte nulle. In definitva l'insieme A_ϵ contiene i punti di G_0 che soddisfano: $(\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{0},\ldots,\mathbf{0})\neq(\mathbf{x}_{0,1},\ldots,\mathbf{x}_{0,N},\mathbf{0},\ldots,\mathbf{0})$ e

$$||(\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{0},\ldots,\mathbf{0})-(\mathbf{x}_{0,1},\ldots,\mathbf{x}_{0,N},\mathbf{0},\ldots,\mathbf{0})||\leq\epsilon\,.$$

Se $\mathbb{R}\ni t\mapsto(\mathbf{x}_1(t),\ldots,\mathbf{x}_N(t),\mathbf{v}_1(t),\ldots,\mathbf{v}_N(t))$ è una soluzione massimale completa che appartiene a qualche A_ϵ, deve quindi valere $\mathbf{v}_k(t)=\mathbf{0}$ per ogni $t\in\mathbb{R}$ e, di conseguenza da (5.20), deve anche essere:

$$\nabla_{\mathbf{x}_k}\mathscr{U}(\mathbf{x}_1(t),\ldots,\mathbf{x}_n(t))+\mathbf{F}_k(\mathbf{x}_1(t),\ldots,\mathbf{x}_N(t),\mathbf{0},\ldots,\mathbf{0})=0\,.$$

In virtù della condizione $\mathbf{F}_k=\mathbf{F}_k(P_1,\ldots,P_N,\mathbf{0},\ldots,\mathbf{0})=\mathbf{0}$ per ogni $k=1,\ldots,N$, sulla soluzione massimale e completa suddetta deve essere soddisfatta l'identità :

$$\nabla_{\mathbf{x}_k}\mathscr{U}(\mathbf{x}_1(t),\ldots,\mathbf{x}_n(t))=\mathbf{0}\,.$$

Quindi tutte le configurazioni $(\mathbf{x}_1(t),\ldots,\mathbf{x}_n(t))$ (che appartengono a U per costruzione) sono punti di stazionarietà dell'energia potenziale. Ma nell'aperto U questo è possibile solo se

$$(\mathbf{x}_1(t),\ldots,\mathbf{x}_n(t))=(\mathbf{x}_{0,1},\ldots\mathbf{x}_{0,N})\,,$$

[4] Il fatto che $(\mathbf{x}_{0,1},\ldots,\mathbf{x}_{0,N})$ individui un minimo stretto dell'energia potenziale non è sufficiente ad assicurare che sia anche un punto di stazionarietà isolato dell'energia potenziale: si consideri, per $N=1$, la funzione $\mathscr{U}(x):=x^4(2+\sin(1/x))$ se $x\neq0$ e $\mathscr{U}(0):=0$. $\mathscr{U}\in C^2(\mathbb{R})$ ed ha un minimo stretto per $x=0$ che non corrisponde ad un punto di stazionarietà isolato.

in quanto $(\mathbf{x}_{0,1}, \dots \mathbf{x}_{0,N})$ è un punto stazionario isolato di $\mathscr{U}$. Dato che $(\mathbf{x}_{0,1}, \dots \mathbf{x}_{0,N}) \notin A_\epsilon$, concludiamo che A_ϵ non contiene orbite di soluzioni massimali complete e quindi vale la stabilità asintotica nel futuro. □

Osservazioni 5.8. La dimostrazione del della proposizione 5.6, del teorema di Lagrange-Dirichlet e della proposizione 5.7 si basa unicamente sulla forma delle equazioni (5.22) e sul fatto che le masse m_k siano positive. Non è importante la dimensione dello spazio vettoriale in cui varia ogni vettore $\mathbf{x}_i$ che può essere diversa da 3 (ma comunque finita). Tale osservazione è importante perché permette di estendere i teoremi al caso in cui si lavora con sistemi vincolati e la forma (5.22) delle equazioni vale per il sistema di equazioni pure di movimento, ed in cui $\mathbf{x}_i$ viene sostituita dalla coordinata k-esima di un sistema di coordinate con il quale si descrive il sistema e la massa m_i è una costante strettamente positiva che non ha necessariamente il significato di una massa.

5.2.2 Un criterio per l'instabilità

Per concludere proviamo un risultato, riguardate sistemi di punti materiali sottoposti a sistemi di forze puramente conservative, che stabilisce una condizione sufficiente per l'instabilità basata sulle proprietà della matrice hessiana dell'energia potenziale.

Proposizione 5.8. *Si consideri di N punti materiali P_k con masse, rispettivamente, $m_k > 0$ e $k = 1, \dots, N$, le cui equazioni della dinamica nel riferimento $\mathscr{I}$ hanno forma (se $\mathbf{x}_i = P_i - O$ con O in quiete in $\mathscr{I}$):*

$$m_k \frac{d^2\mathbf{x}_k}{dt^2} = -\nabla_{\mathbf{x}_k} \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \dots, \mathbf{x}_n)\,, \tag{5.25}$$

dove l'energia potenziale energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \dots, \mathbf{x}_N)$ corrisponde ad un sistema di forze conservative nel riferimento $\mathscr{I}$
Se in una configurazione di equilibrio $(\mathbf{x}_{0,1}, \dots, \mathbf{x}_{0,N})$ la matrice hessiana dell'energia potenziale ha un autovalore negativo (rispettivamente positivo), allora $(\mathbf{x}_{0,1}, \dots, \mathbf{x}_{0,N})$ è instabile nel futuro (rispettivamente nel passato). ♢

Dimostrazione. Dimostreremo il caso relativo all'instabilità nel futuro. L'altro caso si prova analogamente. Consideriamo coordinate cartesiane solidali con $\mathscr{I}$ con origine O coincidente con il punto stazionario $(\mathbf{x}_{0,1}, \dots, \mathbf{x}_{0,N}) = (\mathbf{0}, \dots \mathbf{0})$. In tali coordinate le equazioni del moto del sistema, riportate al prim'ordine sono:

$$\frac{d\mathbf{v}'_k}{dt} = -\nabla_{\mathbf{x}'_k} \mathscr{U}(\mathbf{x}'_1/\sqrt{m_1}, \dots, \mathbf{x}'_n/\sqrt{m_N})\,, \quad k = 1, \dots, N\,, \tag{5.26}$$

$$\frac{d\mathbf{x}'_k}{dt} = \mathbf{v}'_k\,, \quad k = 1, \dots, N\,. \tag{5.27}$$

dove abbiamo introdotto le nuove variabili $\mathbf{x}_i := \sqrt{m_i}\mathbf{x}_i$ e $\mathbf{v}'_i := \sqrt{m_i}\mathbf{v}_i$. La richiesta di $\mathscr{U}$ di classe C^3 è sufficiente ad assicurare la validità delle ipotesi del teorema 5.4 tenendo conto di

(1) in osservazioni 5.6. Se linearizziamo tali equazioni in corrispondenza del punto singolare $(\mathbf{0},\dots,\mathbf{0},\mathbf{0},\dots,\mathbf{0})$, otteniamo il sistema del primo ordine lineare:

$$\frac{du'_j}{dt} = -\sum_{i=1}^{N} H'_{ji}\, y'_i\,, \quad j=1,\dots,3N\,, \tag{5.28}$$

$$\frac{dy'_j}{dt} = u'_j\,, \quad j=1,\dots,3N\,. \tag{5.29}$$

dove y'_1, y'_2, y'_3 sono le componenti di $\mathbf{x}'_1$. Nello stesso modo, y'_4, y'_5, y'_6 sono le componenti di $\mathbf{x}'_2$, e via di seguito.

$$H'_{ji} := M_j^{-1/2} H_{ji} M_i^{-1/2}\,,$$

inoltre $M_1 = M_2 = M_3 = m_1$ e $M_4 = M_5 = M_6 = m_2$ e via di seguito, ed infine la matrice simmetrica H di coefficienti H_{ij} è la matrice hessiana di $\mathscr{U}(\mathbf{x}_1,\dots,\mathbf{x}_N)$ valutata nella configurazione di stazionarietà $(\mathbf{x}_{01},\dots,\mathbf{x}_{0N})$. Nelle nostre ipotesi, la matrice reale simmetrica H ha un autovettore:

$$\mathbf{y}_\lambda = (y_{\lambda 1},\dots,y_{\lambda 3N})^t\,,$$

con autovalore $\lambda < 0$: vale cioé $H\mathbf{y}_\lambda = \lambda \mathbf{y}_\lambda$. Ma allora la matrice H', i cui coefficienti sono i numeri H'_{ji}, ha anch'essa un autovalore, dato dal vettore non nullo:

$$\mathbf{y}'_\lambda := \left(\sqrt{M_1}\, y_{\lambda 1},\dots,\sqrt{M_{3N}}\, y_{\lambda 3N}\right)^t\,,$$

con autovalore $\lambda < 0$:

$$H'\mathbf{y}'_\lambda = \lambda \mathbf{y}'_\lambda\,. \tag{5.30}$$

Se introduciamo il vettore colonna su $\mathbb{R}^{6N}$, $\mathbf{Y} = (u'_1,\dots,u'_{3N}, y'_1,\dots y'_{3N})^t$, il sistema lineare (5.28)-(5.29) si può scrivere:

$$\frac{d\mathbf{Y}}{dt} = A\mathbf{Y}\,, \tag{5.31}$$

dove la matrice A è definita da:

$$A := \left[\begin{array}{c|c} 0 & -H' \\ \hline I & 0 \end{array}\right]\,. \tag{5.32}$$

Consideriamo infine il vettore non nullo (notare che $-\lambda > 0$ per ipotesi!):

$$\mathbf{U} := \left(y'_{\lambda,1},\ \dots,\ y'_{\lambda,3N},\ (-\lambda)^{-1/2} y'_{\lambda,1},\ \dots,\ (-\lambda)^{-1/2} y'_\lambda\right)^t\,.$$

Da (5.32) e (5.30), segue subito che:

$$A\mathbf{U} = \sqrt{-\lambda}\,\mathbf{U}\,.$$

Abbiamo provato che il sistema lineare (5.31), cioé il sistema (5.28)-(5.29), ha matrice a secondo membro con un autovalore reale positivo ($\sqrt{-\lambda}$). In base alla proposizione 5.5, il punto

singolare di questo sistema dato dall'origine di $\mathbb{R}^{6N}$ è instabile nel futuro. Applicando il teorema 5.4, concludiamo che il punto singolare $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N}, \mathbf{0}, \ldots, \mathbf{0})$ del sistema (5.26)-(5.27) è instabile nel futuro. Questo risultato implica immediatamente che la configurazione di equilibrio $(\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,N})$ per il sistema di punti materiali sia instabile nel futuro. □

Osservazioni 5.9. La dimostrazione della proposizione 5.8 si basa unicamente sulla forma delle equazioni (5.25) e sul fatto che le masse m_k siano positive. Non è importante la dimensione dello spazio vettoriale in cui varia ogni vettore $\mathbf{x}_i$ che può essere diversa da 3 (ma comunque finita). Tale osservazione è importante perché permette di estendere i teoremi al caso in cui si lavora con sistemi vincolati e la forma (5.22) delle equazioni vale per il sistema di equazioni pure di movimento, ed in cui $\mathbf{x}_i$ viene sostituita dalla coordinata k-esima di un sistema di coordinate con il quale si descrive il sistema e la massa m_i è una costante strettamente positiva che non ha necessariamente il significato di una massa.

Esempi 5.4.
1. L'esempio più elementare possibile è quello di un punto materiale P sottoposto ad una forza, conservativa rispetto al riferimento $\mathscr{I}$, con energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(P)$ (di classe C^2). Se $P_0 \in E_{\mathscr{I}}$ è un punto di minimo stretto in un intorno di P_0 stesso, questa è una configurazione di equilibrio stabile, nel passato e nel futuro per il teorema di Lagrange-Dirichlet. L'introduzione di una forza viscosa $-\gamma \mathbf{v}_P|_{\mathscr{I}}$ rende tale punto di equilibrio asintoticamente stabile nel futuro a causa della proposizione 5.7.
2. Consideriamo un punto materiale sottoposto ad una forza conservativa conservativa rispetto al riferimento $\mathscr{I}$, con energia potenziale $\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(P)$ (di classe C^2). (Quanto stiamo per dire si estende banalmente al caso di un sistema di punti materiali sottoposti unicamente ad un sistema di forze conservative.) Se non sono presenti altre forze, lo studio della stabilità di un punto di stazionarietà , P_0, di $\mathscr{U}|_{\mathscr{I}}$ può essere fatto analizzando la matrice hessiana di $\mathscr{U}|_{\mathscr{I}}$ nel punto P_0. Più precisamente:
(a) se la matrice hessiana nel punto P_0 ha solo autovalori strettamente positivi, allora si ha un minimo stretto nell'intorno di P_0 e, di conseguenza, il punto di stazionarietà è una configurazione di equilibrio stabile (nel passato e nel futuro).
(b) se la matrice hessiana nel punto P_0 ha almeno un autovalore strettamente negativo, allora il punto di stazionarietà è una configurazione di equilibrio instabile (nel futuro).
Si osservi che la presenza di un minimo stretto non richiede necessariamente autovalori strettamente positivi: qualcuno potrebbe essere nullo. Si consideri ad esempio la funzione

$$\mathscr{U}|_{\mathscr{I}}(x, y, x) := x^2 + y^2 + z^4 \,.$$

Essa ha un minimo stretto nell'origine ma l'autovalore della matrice hessiana corrispondente al versore $\mathbf{e}_z$ è nullo (i rimanenti sono positivi). Viceversa l'instabilità non è sempre evidenziata da un autovalore negativo della matrice hessiana. Per esempio un punto materiale sottoposto alla sola forza conservativa (rispetto al riferimento $\mathscr{I}$) data da:

$$\mathscr{U}|_{\mathscr{I}}(x, y, x) := x^2 + y^2 - z^4 \,,$$

ha una configurazione di equilibrio instabile nel futuro nell'origine. D'altra parte gli autovalori della matrice hessiana nell'origine sono non negativi. L'instabilità è dovuta alla direzione z corrispondente ad un autovalore nullo.

Esercizi 5.1.

1. Considerare un punto materiale di massa m sottoposto alla forza conservativa

$$\mathscr{U}|_{\mathscr{I}}(x, y, x) := \kappa(x^2 + y^2 - z^4) \,,$$

rispetto a coordinate cartesiane ortonormali x, y, z solidali con il riferimento $\mathscr{I}$ e dove $\kappa > 0$ è una costante. Dimostrare che l'origine delle coordinate determina una configurazione di equilibrio instabile.

2. Si consideri un punto materiale P di massa $m > 0$ sottoposto alla forza di gravità $-mg\,\mathbf{e}_z$ e connesso al punto O tramite una molla, di lunghezza a riposo nulla, di costante elastica $\kappa > 0$ e immerso in un liquido. Il punto O ed il liquido sono in quiete nel riferimento non inerziale $\mathscr{I}'$ che ruota, attorno all'asse z uscente da O, rispetto al riferimento inerziale $\mathscr{I}$ con velocità angolare $\omega_{\mathscr{I}'}|_{\mathscr{I}} = \omega\,\mathbf{e}_z$ dove $\omega > 0$ è costante. A causa della presenza del liquido, P è sottoposto alla forza viscosa $\mathbf{F}(\mathbf{v}_P|_{\mathscr{I}'}) = -\gamma \mathbf{v}_P|_{\mathscr{I}'}$ con $\gamma \geq 0$ costante.
(i) Si scrivano le equazioni del moto del punto P nel riferimento $\mathscr{I}'$.
(ii) Si dimostri che, nel riferimento $\mathscr{I}'$, se non sono presenti le forze viscose ($\gamma = 0$), allora il punto è sottoposto ad una forza girostatica ed a una rimanente forza conservativa. Si scriva l'energia potenziale di quest'ultima.
(iii) Si determinino le configurazioni di equilibrio di P *nel riferimento* $\mathscr{I}'$ e se ne discuta la stabilità , per $k > m\omega^2$.

Capitolo 6

Fondamenti di Meccanica Lagrangiana.

In questo capitolo ci occuperemo di introdurre la formulazione lagrangiana della meccanica classica. Ricordiamo che la presenza di reazioni vincolari di cui non è nota la forma funzionale rende, in generale, *non deterministiche* le equazioni di Newton, proprio perché le reazioni vincolari appaiono come incognite aggiuntive. Per rendere il sistema di equazioni deterministico è necessario aggiungere informazioni di vario genere: geometrico/cinematico ("forma dei vincoli") e fisico (caratterizzazione costitutiva delle reazioni vincolari). La *meccanica lagrangiana* o *fomulazione lagrangiana della meccanica classica* nasce come una procedura universale per estrarre equazioni pure di movimento (le cosiddette *equazioni di Eulero-Lagrange*) per sistemi di punti materiali vincolati, quando i vincoli soddisfano il requisito di essere *vincoli ideali*. Tale requisito generalizza, come vedremo, la caratterizzazione di vincolo liscio includendo situazioni estremamente generali.
La meccanica lagrangiana è applicabile, con piccole modificazioni, anche a sistemi materiali rigidi continui. La formulazione lagrangiana si applica anche, ovviamente, al caso di sistemi di punti materiali non sottoposti a vincoli. In questo caso le equazioni sono completamente equivalenti a quelle deterministiche di Newton, ma hanno la particolarità di essere scrivibili, nella stessa forma, in un qualsiasi sistema di coordinate. Infine, in tutti i casi, nella formulazione di Lagrange della meccanica risulta molto chiaro il legame, strettissimo, tra simmetria di un sistema fisico sotto un gruppo di trasformazioni ed esistenza di particolari quantità fisiche conservate nel tempo sul moto del sistema.
La meccanica lagrangiana è un approccio potentissimo la cui applicazione è andata ben oltre la meccanica newtoniana classica con un'enorme influenza, sicuramente decisiva, nello sviluppo della fisica matematica e teorica moderna.

N.B. In tutto questo capitolo *differenziabile* significa di classe almeno C^2 se non è specificato altrimenti. Inoltre *sottovarietà* significherà *sottovarietà embedded* (vedi appendice A).

6.1 Sistemi vincolati ed equazioni di Eulero-Lagrange.

Consideriamo un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$ con masse $m_1, \ldots, m_N$ rispettivamente. Per varie ragioni e in vari rami della fisica classica (in particolare in *Meccanica Statistica*) è spesso comodo studiare la cinematica e la dinamica di tale sistema non nello spaziotempo $\mathbb{V}^4$, ma in una varietà $\mathbb{V}^{3N+1}$ di dimensione $3N+1$ detto **spaziotempo delle configurazioni**[1]. La sua struttura è identica a quella di $\mathbb{V}^4$ con la differenza che le foglie a tempo assoluto costante, $T : \mathbb{V}^{3N+1} \to \mathbb{R}$ pensato come una funzione suriettiva e regolare, Σ_t sono ora sostituite dalle sottovarietà di dimensione $3N+1$, date prodotto cartesiano di N copie dello spazio euclideo Σ_t, cioé $\Sigma_t^N = \Sigma_t \times \cdots \times \Sigma_t$. La k-esiama copia di Σ_t si deve pensare come lo spazio assoluto di P_k, e questo per ogni $k = 1, \ldots, N$.
Un punto $p \in \mathbb{V}^{3N+1}$ è quindi individuato da un numero $t := T(p)$ e da una N-pla di punti $\mathcal{C}_t := (Q_1, \ldots, Q_N) \in \Sigma_t$ detta **configurazione al tempo** t. Ovviamente assegnare la configurazione $\mathcal{C}_t$ al tempo t equivale ad assegnare il valore di $3N$ coordinate: una terna di numeri per ogni punto P_k. Tenuto conto del valore t assunto dal tempo assoluto nel momento in cui si assegna $\mathcal{C}_t$, per determinare un punto in $\mathbb{V}^{3N+1}$ bisogna fissare $3N+1$ coordinate. Di conseguenza lo spaziotempo delle configurazioni è una varietà di dimensione $3N+1$ e questo spiega la notazione $\mathbb{V}^{3N+1}$.
L'evoluzione del sistema $\mathfrak{S}$ è assegnata da una **linea di universo**, cioé una curva differenziabile $\Gamma = \Gamma(t) \in \mathbb{V}^{3N+1}$ con $t \in I$ intervallo aperto di $\mathbb{R}$ che non possa intersecare più di una volta ogni Σ_t^N per evitare i soliti paradossi temporali. Dal punto di vista matematico questa restrizione viene imposta richiedendo, come nel caso di una sola particella, che $T(\Gamma(t)) = t$ (per una scelta opportuna della costante arbitraria nella definizione del tempo assoluto) per ogni valore di $t \in I$. Identificando il tempo assoluto di $\mathbb{V}^{3N+1}$ con il tempo assoluto dello spaziotempo $\mathbb{V}^4$, per ogni $k = 1, \ldots, N$, l'applicazione differenziabile surgettiva:

$$\Pi_k : \mathbb{V}^{3N+1} \ni (t, (Q_1, \ldots, Q_N)) \mapsto (t, Q_k) \in \mathbb{V}^4$$

estrae la configurazione del singolo punto P_k dalla configurazione totale del sistema, ad ogni istante t del tempo assoluto. $\gamma_k(t) := \Pi_k(\Gamma(t))$, al variare di $t \in I$, definisce la linea di universo del k-esimo punto, quando è data la linea di universo Γ del sistema $\mathfrak{S}$. Tramite le funzioni Π_k ha senso, per esempio, definire la distanza tra due punti $P_k(t)$ e $P_h(t)$ della configurazione $\mathcal{C}_t = \Gamma(t)$ ad un fissato tempo t tramite la distanza d_t di cui è dotato ogni spazio assoluto $\Sigma_t \subset \mathbb{V}^4$. A tal fine è sufficiente pensare i punti $P_k(t)$ e $P_h(t)$ come in $\mathbb{V}^4$ tramite l'azione delle due funzioni Π_k e Π_h.

6.1.1 Spaziotempo delle configurazioni in presenza di vincoli olonomi.

Supponiamo ora che gli N punti del sistema Σ siano sottoposti a vincoli posizionali espressi da $C < 3N$ ($C \geq 0$) funzioni differenziabili $f_j : \mathbb{V}^{3N+1} \to \mathbb{R}$, tramite le C richieste

$$f_j(t, P_1(t), \ldots, P_N(t)) = 0\,, \quad j = 1, \ldots, C \tag{6.1}$$

[1] In meccanica statistica si studiano sistemi fisici il cui numero di elementi N, dati da atomi o molecole, è estremamente grande, dell'ordine del *numero di Avogadro*: 6.02×10^{23}.

che devono essere soddisfatte su ogni moto del sistema $\Gamma = \Gamma(t)$, $t \in I$. In questa situazione, molto frequente nella pratica, le configurazioni del sistema fisico ad un tempo t non sono tutte quelle rappresentate dai punti di Σ_t^N. I vincoli, ad ogni istante t, determineranno un sottoinsieme $\mathbb{Q}_t \subset \Sigma_t^N$ delle configurazioni realmente accessibili al sistema fisico in quell'istante. Vogliamo, per esigenze di semplicità fisico-matematica, che questi sottoinsiemi $\mathbb{Q}_t$ abbiano una struttura matematica che permettano l'uso delle solite tecniche analitiche basate sulle equazioni differenziali descriventi le leggi di evoluzione del sistema. La situazione ottimale (ma non l'unica possibile) è quella in cui i sottoinsiemi $\mathbb{Q}_t$ abbiano a loro volta una struttura di varietà differenziabile di dimensione $n \leq 3N$. È naturale assumere che questa struttura sia quella ereditata dalla struttura di ogni Σ_t^N. In altre parole assumeremo che i sottoinsiemi $\mathbb{Q}_t$ abbiamo la struttura di *sottovarietà* (vedi Appendice A) delle varietà Σ_t^N.
La dimensione n sarà fissata dal numero delle funzioni di vincolo f_j: ci si aspetta che se imponiamo C equazioni di vincolo *indipendenti* ad ogni tempo t, queste siano in grado di lasciare libera la scelta di $n := 3N - C$ coordinate degli N punti materiali che costituiscono il sistema fisico.
Dal punto di vista matematico questa situazione si ottiene sempre quando le equazioni di vincolo (6.1) soddisfano le seguenti ipotesi tecniche.

(H1) Per ogni $t \in \mathbb{R}$ le condizioni

$$f_j(t, P_1, \ldots, P_N) = 0\,, \quad j = 1, \ldots, C \tag{6.2}$$

siano soddisfatte su un insieme non vuoto.

(H2) Ad ogni istante $t \in \mathbb{R}$, le funzioni f_j risultino essere **funzionalmente indipendenti**.

N.B. *Funzionalmente indipendenti* significa che, nell'intorno di ogni punto $p \in \mathbb{V}^{3N+1}$, esprimendo le funzioni f_j in termini di coordinate locali $t, x^1, \ldots, x^{3N}$ su $\mathbb{V}^{3N+1}$, con t tempo assoluto e $x^1, \ldots, x^{3N}$ coordinate locali su Σ_t^N[2], la matrice jacobiana di elementi

$$\frac{\partial f_j}{\partial x^i}\,, \quad i = 1, \ldots, 3N\,, j = 1, \ldots, C\,,$$

ha rango C, per ogni fissato istante t, sull'insieme dei punti di $\mathbb{V}^{3N+1}$ per per i quali valgano

$$f_j(t, x^1, \ldots, x^{3N}) = 0\,.$$

Come chiarito in appendice A, tale caratterizzazione è indipendente dalla scelta delle coordinate locali.

Diremo **sistema di vincoli olonomi** un sistema di vincoli, imposti su un sistema di N punti materiali, che rispetta i requisiti (H1) e (H2) sopra imposti.

[2]Tali coordinate esistono sempre (essendo le superfici a $T = t$ costante delle sottovarietà) e si possono costruire nello stesso modo in cui abbiamo definito, per $N = 1$, le coordinate cartesiane ortonormali solidali con un riferimento.

Vale la seguente proposizione.

Proposizione 6.1. *Per un sistema $\mathfrak{S}$ di N punti materiali descritti in $\mathbb{V}^{3N+1}$ e sottoposto ai vincoli (6.2) descritti da $C \leq 3N$ funzioni differenziabili f_j, si consideri l'insieme $\mathbb{V}^{n+1} \subset \mathbb{V}^{3N+1}$ individuato dalle condizioni (6.2).*
Nell'ipotesi che i vincoli siano olonomi valgono i fatti seguenti.
(1) *$\mathbb{V}^{n+1}$ è una sottovarietà di $\mathbb{V}^{3N+1}$ di dimensione $n+1$.*
(2) *Per ogni fissato tempo t, l'insieme $\mathbb{Q}_t := \mathbb{V}^{n+1} \cap \Sigma_t^{3N}$ è una sottovarietà di Σ_t^{3N} di dimensione n.*
(3) *Nell'intorno di ogni punto $p \in \mathbb{V}^{n+1}$ possiamo sempre scegliere un sistema di coordinate $(t, q^1, \ldots, q^n)$ in cui: t misura il tempo assoluto e, per ogni fissato valore del tempo assoluto t, $q^1, \ldots, q^n$ siano coordinate locali ammissibili su $\mathbb{Q}_t$. (In particolare $q^1, \ldots, q^n$ possono essere scelte come n tra le coordinate cartesiane degli N punti rispetto ad un sistema di riferimento arbitrariamente fissato.)* $\diamondsuit$

Dimostrazione. Mettiamo su $\mathbb{V}^{3N+1}$ un sistema di coordinate cartesiane $t, x^1, \ldots, x^{3N}$ solidali con un fissato riferimento $\mathscr{I}$. Quindi, per $i = 1, \ldots, N$, le coordinate $x^{i+1}, x^{i+2}, x^{i+3}$ sono le coordinate cartesiane in $\mathscr{I}$ del punto $P_i \in \mathfrak{S}$. Si consideri ora l'insieme $\mathbb{V}^{n+1} \subset \mathbb{V}^{3N+1}$ individuato dalle condizioni (6.2), nell'ipotesi che i vincoli siano olonomi. Per il teorema (dei valori regolari) A.2, $\mathbb{V}^{n+1}$ è una sottovarietà di $\mathbb{V}^{3N+1}$ di dimensione $n+1$. Questo segue immediatamente dal fatto che, su $\mathbb{V}^{n+1}$, la matrice jacobiana di elementi – dove le coordinate locali sono definite da: $y^0 := t$ e $x^i := y^i$ per $i = 1, \ldots, 3N$ –

$$\frac{\partial f_j}{\partial y^i}\,, \quad i = 0, \ldots, 3N\,, j = 1, \ldots, C\,,$$

ha rango C, avendo C colonne linearmente indipendenti per il requisito (2). Il requisito (2) implica per la stessa ragione che, per ogni fissato tempo t, $\mathbb{Q}_t := \mathbb{V}^{n+1} \cap \Sigma_t^{3N}$ sia una sottovarietà di Σ_t^{3N} di dimensione n. Si osservi che $\mathbb{Q}_t \neq \emptyset$ per il requisito (1).
La seconda parte del teorema A.2 ha un'altra conseguenza. Dato che la matrice jacobiana

$$\frac{\partial f_j}{\partial y^i}\,, \quad i = 0, \ldots, 3N\,, j = 1, \ldots, C\,,$$

ha rango C, e che questo fatto è dovuto alla sottomatrice

$$\frac{\partial f_j}{\partial y^i}\,, \quad i = 1, \ldots, 3N\,, j = 1, \ldots, C\,,$$

nell'intorno di ogni punto $p \in \mathbb{V}^{n+1}$, possiamo usare $n+1$ coordinate per descrivere $\mathbb{V}^{n+1}$ scegliendo le coordinate tra le $y^0, \ldots, y^{3N}$ come segue. La prima coordinata è sempre y^0 (il tempo assoluto) e le rimanenti $n = 3N - C$ sono scelte tra le $y^1, \ldots, y^{3N}$, in modo tale che la sotto matrice jacobiana $C \times C$, corrispondente alle coordinate y^k non scelte, sia non singolare in p. $\square$

Definizione 6.1. (**Spaziotempo delle configurazioni con vincoli olonomi**.) Si consideri un sistema $\mathfrak{S}$ di N punti materiali sottoposto ad un sistema di $C < 3N$ vincoli olonomi $f_j(t, P_1, \ldots, P_N) = 0$ con $j = 1, \ldots, C$, dove le funzioni $f_j : \mathbb{V}^{3N+1} \to \mathbb{R}$ sono differenziabili.
La sottovarieà $\mathbb{V}^{n+1}$ di dimensione $n := 3N - C$, individuata da (6.2), è detta **spaziotempo delle configurazioni** per il sistema vincolato $\mathfrak{S}$.
Il numero $n := 3N - C$ è detto **numero dei gradi di libertà** del sistema $\mathfrak{S}$.
Per ogni istante del tempo assoluto $t \in \mathbb{R}$, $\mathbb{Q}_t := \mathbb{V}^{n+1} \cap \Sigma_t^{3N}$ è detto **spazio delle configurazioni** al tempo t, per il sistema vincolato $\mathfrak{S}$. Una **configurazione** $\mathcal{C}_t$ del sistema vincolato $\mathfrak{S}$ al tempo t è dunque un punto in $\mathbb{Q}_t$.
Ogni sistema di coordinate locali $\phi : U \ni p \mapsto (t, q^1, \ldots q^n) \in V$ – dove $U \subset \mathbb{V}^{n+1}$ e $V \subset \mathbb{R}^{n+1}$ sono insiemi aperti, $t = T(p)$, e le coordinate $(q^1, \ldots q^n)$ individuano coordinate locali su $\mathbb{Q}_t$ per ogni t – è detto **sistema di coordinate locali naturali** su $\mathbb{V}^{n+1}$. Le coordinate $q^1, \ldots, q^n$ su ogni $\mathbb{Q}_t$ sono dette **coordinate libere**.
Un sistema di coordinate naturali $(t, q^1, \ldots, q^n)$ su $\mathbb{V}^{n+1}$ si dice **solidale con il riferimento** $\mathscr{I}$ se, detto $\mathbf{x}_i = P_i - O$ il vettore posizione del punto P rispetto all'origine O di $\mathscr{I}$ (le componenti di $\mathbf{x}_i$ sono dunque le coordinate di P in un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$), la relazione parametrica tra $\mathbf{x}_i$ e le coordinate dette non contiene esplicitamente il tempo: $\mathbf{x}_i = \mathbf{x}_i(q^1, \ldots, q^n)$, per ogni $i = 1, \ldots, N$. $\diamondsuit$

Osservazioni 6.1.
(1) Chiaramente i sistemi di coordinate locali naturali formano un atlante di $\mathbb{V}^{n+1}$.
Se $t, q^1, \ldots, q^n$ e $t', q'^1, \ldots q'^n$ sono due distinti sistemi di coordinate locali naturali definite su due domini $U, U' \subset \mathbb{V}^{n+1}$ con $U \cap U' \neq \emptyset$, nell'intersezione dei due domini, la relazione tra i due sistemi di coordinate è del tipo:

$$q'^k = q'^k(t, q^1, \ldots, q^n) \,, \tag{6.3}$$

$$t' = t + c \,, \tag{6.4}$$

dove c è una costante relativa alla nota ambiguità nella definizione del tempo assoluto.
(2) Uno spazio delle configurazioni $\mathbb{Q}_t$ al tempo t si dice anche **fibra** di $\mathbb{V}^{n+1}$ al tempo t. I sistemi di coordinate locali naturali di $\mathbb{V}^{n+1}$ si chiamano anche equivalentemente, sistemi di coordinate locali naturali **adattate alle fibre** di $\mathbb{V}^{n+1}$. Le coordinate libere sono anche chiamate **coordinate lagrangiane**.
(3) Le configurazioni $\mathcal{C}_t$ si possono pensare contemporaneamente come punti di $\mathbb{V}^{n+1}$, ma anche di $\mathbb{V}^{3N+1}$, dato che ogni $\mathbb{Q}_t$ è sottovarietà di Σ_t^N
(4) Se i punti $P_i \in \mathbb{V}^4$ del sistema vincolato sono individuati da vettori posizione $P_i - O = \mathbf{x}_i$ rispetto ad un'origine O di evoluzione assegnata, e se $t, q^1, \ldots, q^n$ sono coordinate naturali definite nell'intorno di $\mathcal{C}_t$, si dovrà poter scrivere come già utilizzato nella definizione 6.1:

$$\mathbf{x}_i = \mathbf{x}_i(t, q^1, \ldots, q^n) \,,$$

dove, in generale il tempo t appare esplicitamente. Il moto del sistema sarà dunque individuato, almeno localmente, da funzioni differenziabili $q^k = q^k(t)$ attraverso le quali si descrivono, in

funzione del tempo, i vettori posizione dei punti del sistema fisico e quindi la loro evoluzione:

$$\mathbf{x}_i = \mathbf{x}_i(t, q^1(t), \ldots, q^n(t)) \,,$$

Esempi 6.1.
(1) Consideriamo il caso elementare di $N = 1$, quindi un punto materiale P unicamente, sottoposto alla coppia di vincoli:

$$x^2 + y^2 = 1 + ct^2 \quad \text{insieme a} \quad z = x \,, \quad \text{per ogni } t \in \mathbb{R} \,.$$

Le coordinate x, y, z sono coordinate cartesiane ortonormali solidali con un riferimento $\mathscr{I}$ e t è il solito tempo assoluto e $c > 0$. Abbiamo allora due funzioni di vincolo che si annullano insieme su un insieme non vuoto ad ogni $t \in \mathbb{R}$. Le due funzioni sono: $f_1(t, x, y, z) = x^2 + y^2 - (1 + ct^2)$ e $f_2(t, x, y, z) = z - x$. La matrice jacobiana è costituita dai due vettori riga

$$(2x, 2y, 0) \,, \quad (1, 0, -1) \,.$$

Si osservi che quando $x^2 + y^2 = 1 + ct^2$ il primo vettore non può mai annullarsi. Inoltre esso non può mai essere linearmente dipendente dal secondo (cioé proporzionale ad esso). Concludiamo che le ipotesi (H1) e (H2) sono soddisfatte. Per ogni tempo t i due vincoli individuano una curva $\mathbb{Q}_t = \Gamma_t$ data dall'intersezione del cilindro di raggio $\sqrt{1 + ct^2}$ ed asse dato dall'asse z, e dal piano $z = x$. Tale curva è una circonferenza di raggio variabile nel tempo. Il punto materiale ha dunque un solo grado di libertà : $1 = 3 - 2$.
Vediamo di determinare delle coordinate libere. Un sistema di coordinate libere è sicuramente dato dalla quota del punto $q^1 := z$ del punto P. È chiaro che tale coordinata è solo locale in quanto, descrivendo la proiezione dei Γ_t sul piano x, y in termini dell'angolo polare $\varphi \in (-\pi, +\pi)$ (riferito a coordinate polari piane nel piano x, y), i punti su Γ_t con angolo polare $\pm\varphi$ hanno la stessa z. Tuttavia, in ciascuno dei rami $(-\pi, 0)$ e $(0, \pi)$, la coordinata z può essere presa come coordinata libera. Una scelta per molti versi più appropriata è data da $q^1 := \varphi$, dove φ è la coordinata polare descritta sopra. Le coordinate cartesiane di P nel riferimento $\mathscr{I}$ sono allora date, in funzione di q^1 e per ogni istante di tempo t, come:

$$x(t, q^1) = \sqrt{1 + ct^2} \cos q^1 \,, \tag{6.5}$$

$$y(t, q^1) = \sqrt{1 + ct^2} \sin q^1 \,, \tag{6.6}$$

$$z(t, q^1) = \sqrt{1 + ct^2} \cos q^1 \,. \tag{6.7}$$

(2) Consideriamo il caso meno elementare di $N = 2$ punti materiali P, Q sottoposti al vincolo di rimanere a distanza $d > 0$ fissata. Lo spaziotempo delle configurazione ha inizialmente $6 + 1 = 7$ dimensioni. Usando un riferimento $\mathscr{I}$, possiamo riferirci a coordinate globali su $\mathbb{V}^{6+1}$, t, x, y, z, X, Y, Z, dove t è il solito tempo assoluto e (x, y, z) e (X, Y, Z) sono le coordinate

cartesiane di P e Q rispettivamente in un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ di origine O. Il vincolo detto, se definiamo:

$$f(t,x,y,z,X,Y,Z) := (x-X)^2 + (y-Y)^2 + (z-Z)^2 - d^2 \,,$$

si esprime come:

$$f(x,y,z,X,Y,Z) = 0 \,.$$

Per verificare l'indipendenza funzionale, dobbiamo calcolare il gradiente di f e controllare che non si annulli mai sui punti in cui l'equazione del vincolo è soddisfatta. Tale gradiente è il vettore riga:

$$(2(x-X), 2(y-Y), 2(z-Z), 2(X-x), 2(Y-y), 2(Z-z)) \,.$$

È evidentemente impossibile che tale vettore sia nullo se $(x-X)^2+(y-Y)^2+(z-Z)^2 = d^2$ con $d > 0$. Per ogni tempo $t \in \mathbb{R}$ possiamo risolvere l'equazione di vincolo esprimendo per esempio X in funzione delle rimanenti 5 variabili:

$$X = x \pm \sqrt{d^2 - (y-Y)^2 - (z-Z)^2}$$

La presenza di $\pm$ denota il fatto che le coordinate $(t, q^1, q^2, q^3, q^4, q^5) := (t, x, y, z, Y, Z)$ sono coordinate locali ed sono necessari più , sistemi di coordinate locali per ricoprire, ad ogni istante, lo spaziotempo delle configurazioni $\mathbb{V}^{5+1}$. Un sistema di coordinate locali più utile nelle applicazioni e sicuramente più visualizzabile è dato come segue. q^1, q^2, q^3 sono le tre coordinate cartesiane, rispetto a $\mathscr{I}$, del punto medio M tra P e Q, q^3, q^4 sono gli angoli polari di $P - M$ rispetto ad un sistema di assi cartesiani ortonormali di origine M che si muove restando parallelo al sistema di assi cartesiani ortonormali solidali con $\mathscr{I}$ in O. Si può provare per esercizio che questo sistema di coordinate è ammissibile su $\mathbb{Q}_t$ e che $(t, q^1, \ldots, q^5)$ individuate come detto definiscano un sistema di coordinate locali naturali di $\mathbb{V}^{5+1}$.

6.1.2 Spostamenti virtuali e vincoli ideali.

Consideriamo il solito sistema di N punti materiali $\mathfrak{S}$ sottoposto a C vincoli olonomi e descritto sullo spaziotempo delle configurazioni $\mathbb{V}^{n+1}$ (visto come sottovarietà di $\mathbb{V}^{3N+1}$). Supponiamo che, fissato un riferimento $\mathscr{I}$, sul generico i-esimo punto agisca la forza:

$$\mathbf{F}_{\mathscr{I}i}(t, P_1, \ldots, P_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}})$$

che, per distinguerla dalla forza reattiva dovuta ai vincoli, sarà denominata **forza attiva**. Abbiamo scritto esplicitamente l'indice $_{\mathscr{I}}$ in quanto assumeremo che $\mathbf{F}_{\mathscr{I}}$ possa includere forze inerziali quando $\mathscr{I}$ è non inerziale. Le equazioni di Newton si scrivono allora:

$$m_{P_i}\mathbf{a}_{P_i}|_{\mathscr{I}} = \mathbf{F}_{\mathscr{I}i} + \boldsymbol{\phi}_i \,, \quad \text{per ogni } i = 1, \ldots, N. \tag{6.8}$$

Come accennato sopra i vincoli, dal punto di vista fisico, sono imposti tramite forze, più precisamente **reazioni vincolari** o **forze reattive** o **forze passive** agenti su ciascun punto P_i. Indicheremo con $\boldsymbol{\phi}_i$, $i = 1, \ldots, N$ la reazione vincolare agente sull'i-esimo punto materiale. Vogliamo ora

dare la nozione di *vincolo ideale* che generalizza la nozione di vincolo liscio vista precedentemente.

Consideriamo, per il nostro sistema $\mathfrak{S}$ sottoposto a C vincoli olonomi e descritto sullo spaziotempo delle configurazioni $\mathbb{V}^{n+1}$ visto come sottovarietà di $\mathbb{V}^{3N+1}$, due configurazioni *entrambe allo stesso tempo* t: $\mathcal{C}_t = (P_1, \ldots, P_N) \subset \Sigma_t \times \cdots \times \Sigma_t$ e $\mathcal{C}'_t = (P'_1, \ldots, P'_N) \subset \Sigma_t \times \cdots \times \Sigma_t$. Possiamo definire, nello spazio $\mathbb{V}^{3N+1}$ il vettore riga di vettori $\Delta P = (\Delta P_1, \ldots, \Delta P_N)$ con

$$\Delta P = (\Delta P_1, \ldots, \Delta P_N) \quad \text{dove } \Delta P_i = P'_i - P_i \text{ per } i = 1, \ldots, N\,.$$

Teniamo ora conto che $\mathbb{Q}_t \subset \Sigma_t^N$. Se i punti P_i e P'_i sono individuati da vettori posizione $P_i - O = \mathbf{x}_i$ e, rispettivamente $P'_i - O = \mathbf{x}'_i$, in riferimento ad un'origine O di evoluzione assegnata, e se $t, q^1, \ldots, q^n$ sono coordinate naturali definite nell'intorno di $\mathcal{C}_t$, avremo che

$$\mathbf{x}_i = \mathbf{x}_i(t, q^1, \ldots, q^n) \quad \text{e} \quad \mathbf{x}'_i = \mathbf{x}_i(t, q'^1, \ldots, q'^n)\,.$$

Di conseguenza, il punto i-esimo avrà uno spostamento, dato da:

$$\Delta P_i = \mathbf{x}_i(t, q'^1, \ldots, q'^n) - \mathbf{x}_i(t, q^1, \ldots, q^n) = \sum_{k=1}^{n} \frac{\partial \mathbf{x}_i}{\partial q^k}|_{\mathcal{C}_t} \delta q^k + O_i((\delta q)^2)\,.$$

Dove $\delta q^k := q'^k - q^k$ e O_i è una funzione infinitesima del secondo ordine per $\sqrt{\sum_{k=1}^{n} |\delta q^k|} \to 0$. In prima approssimazione possiamo trascurare i termini infinitesimi di ordine superiore, se le configurazioni $\mathcal{C}_t$ e $\mathcal{C}'_t$ sono abbastanza vicine. In dipendentemente dalla sensatezza di tale approssimazione, si definisce uno **spostamento virtuale (rispetto alla configurazione $\mathcal{C}_t$ al tempo t)** il vettore riga di vettori:

$$\delta P = \left(\sum_{k=1}^{n} \frac{\partial \mathbf{x}_1}{\partial q^k}|_{\mathcal{C}_t} \delta q^k, \cdots, \sum_{k=1}^{n} \frac{\partial \mathbf{x}_N}{\partial q^k}|_{\mathcal{C}_t} \delta q^k \right)\,, \tag{6.9}$$

dove i $\delta q^k \in \mathbb{R}$ sono numeri arbitrari (anche grandi!).
Il **lavoro virtuale** dovuto alle reazioni vincolari $(\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N)$ rispetto allo spostamento virtuale $\delta P = (\delta P_1, \ldots, \delta P_N)$ riferito alla configurazione $\mathcal{C}_t$ è definito come:

$$\delta L := \sum_{i=1}^{N} \boldsymbol{\phi}_i \cdot \delta P_i\,, \tag{6.10}$$

che equivale a:

$$\delta L = \sum_{k=1}^{n} \left(\sum_{i=1}^{N} \boldsymbol{\phi}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q^k}|_{\mathcal{C}_t} \right) \delta q^k\,. \tag{6.11}$$

Osservazioni 6.2.
(1) Ogni singolo vettore δP_i, *non* dipende dalla scelta del punto O di riferimento. Pertanto, in particolare non dipende da alcuna scelta del sistema di riferimento nello spaziotempo $\mathbb{V}^4$.

(2) Il lavoro virtuale *non* dipende dalla scelta di alcun sistema di riferimento dato che le forze reattive $\boldsymbol{\phi}_i$ hanno valori invarianti in quanto forze vere, e i vettori δP_i non dipendono dal riferimento come chiarito sopra.
(3) * Consideriamo un riferimento $\mathscr{I}$ ed un sistema di coordinate cartesiane ortonormali con esso solidale. Consideriamo le coordinate indotte in ciascuno spazio Σ_t relativo ad ogni punto P_i del sistema. Infine denotiamo con $x^1, x^2, \ldots, x^{3N}$ le coordinate indotte su Σ_t^N: x^1, x^2, x^3 saranno le coordinate cartesiane del punto P_1, x^4, x^5, x^6 saranno le coordinate cartesiane del punto P_2 e via di seguito. La N-pla di vettori definente δP può essere equivalentemente vista come una $3N$-pla di numeri:

$$\delta P = \left(\sum_{k=1}^{n} \frac{\partial x^1}{\partial q^k}|_{\mathcal{C}_t} \delta q^k, \sum_{k=1}^{n} \frac{\partial x^2}{\partial q^k}|_{\mathcal{C}_t} \delta q^k, \cdots, \sum_{k=1}^{n} \frac{\partial x^{3N}}{\partial q^k}|_{\mathcal{C}_t} \delta q^k \right) .$$

Questa a sua volta porta le stesse informazioni del vettore tangente a Σ_t^N in $\mathcal{C}_t$:

$$\delta P := \sum_{j=1}^{3N} \sum_{k=1}^{n} \delta q^k \frac{\partial x^j}{\partial q^k}|_{\mathcal{C}_t} \frac{\partial}{\partial x^j}|_{\mathcal{C}_t} .$$

Dato infine che:

$$\frac{\partial}{\partial q^k} = \sum_{j=1}^{3N} \frac{\partial x^j}{\partial q^k} \frac{\partial}{\partial x^j} ,$$

possiamo equivalentemente pensare lo spostamento virtuale δP come un vettore tangente alla varietà $\mathbb{Q}_t$ applicato a $\mathcal{C}_t \in \mathbb{Q}_t$:

$$\delta P = \sum_{k=1}^{n} \delta q^k \frac{\partial}{\partial q^k}|_{\mathcal{C}_t} . \tag{6.12}$$

Definizione 6.2. (**Vincoli ideali**.) Si consideri un sistema $\mathfrak{S}$ di N punti materiali sottoposto ad un sistema di $C < 3N$ vincoli olonomi. Le reazioni vincolari che si esercitano su una linea di universo $\Gamma = \Gamma(t)$ del sistema sono dette **ideali** se per ogni configurazione $\mathcal{C}_t$ del sistema attraversata dalla linea di universo e per ogni spostamento virtuale rispetto a tale configurazione del sistema, il lavoro virtuale risulta essere nullo[3], cioé :

$$\sum_{k=1}^{n} \left(\sum_{i=1}^{N} \boldsymbol{\phi}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q^k}|_{\mathcal{C}_t} \right) \delta q^k = 0 , \quad \text{per ogni scelta dei } \delta q^k \in \mathbb{R} \text{ con } k = 1, \ldots, n. \tag{6.13}$$

I vincoli sono detti **ideali** se le reazioni vincolari sono ideali su ogni linea di universo che soddisfa le equazioni del moto del sistema (6.8). $\diamondsuit$

[3] I *vincoli ideali* in letteratura sono anche detti "vincoli che soddisfano il principio del lavori virtuali", oppure "vincoli che soddisfano il pricipio di D'Alembert".

Esempi 6.2.
1. L'esempio più semplice di sistema di punti materiali soggetto a vincoli olonomi ideali è quello di un sistema di punti materiali in interazione tra di essi e soggetti a forze esterne, ciascuno dei quali è vincolato a muoversi su una superficie o su una curva liscia in quiete in qualche riferimento. Non è richiesto che il riferimento sia lo stesso per tutti i punti materiali. In questo caso, ad ogni istante, gli spostamenti virtuali sono tangenti alle superfici o curve in questione, mentre le reazioni vincolari sono normali ad essi. Ciascuno dei termini $\boldsymbol{\phi}_i \cdot \delta P_i$ è nullo separatamente e pertanto il lavoro virtuale delle reazioni vincolari è nullo.
2. Un altro esempio molto importante è quello del *vincolo di rigidità* (assumendo anche che il principio di azione e reazione valga in forma forte): un sistema di N punti materiali collegati tra di essi con supporti ideali, in modo tale da rimanere a distanze reciproche costanti per ogni moto del sistema. Le diverse configurazioni ammissibili al sistema (che soddisfano cioé il vincolo di rigidità) ad un tempo fissato, saranno connesse le une alle altre attraverso un'isometria generale di $\mathbb{E}^3$:

$$\mathbf{x}'_i = \mathbf{x}_i + \mathbf{t} + R(\mathbf{x}_i)\,, \quad \text{per ogni } i = 1, \ldots, N,$$

dove $\mathbf{t}$ e R sono una generica traslazione (vettore dello spazio delle traslazioni di Σ_t) ed una rotazione di $O(3)$, entrambe le trasformazioni agiscono attivamente. $\mathbf{x}_i := P_i - O$ è il vettore posizione del punto i-esimo rispetto ad un'origine $O \in \Sigma_t$ fissata una volta per tutte. Fissiamo più in generale un sistema di coordinate cartesiane ortonormali centrate in O con assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Per descrivere la classe delle trasformazioni rigide suddette sono necessari e sufficienti 6 parametri reali $q^1, \ldots q^6$. Le coordinate q^1, q^2, q^3 individuano le componenti sulla terna $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ della traslazione $\mathbf{t}$ mentre i tre angoli q^4, q^5, q^6 individuano la rotazione R. Si può dimostrare, ma non è per nulla banale farlo, che gli angoli q^4, q^5, q^6 possono essere scelti come angoli di rotazione attorno agli assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ rispettivamente, almeno per valori sufficientemente piccoli dei tre angoli. Gli spostamenti virtuali hanno allora la struttura:

$$\delta P_i = \sum_{k=1}^{3} \delta q^k\, \mathbf{e}_k + \sum_{k=1}^{3} \delta q^{k+3}\, \mathbf{e}_k \wedge \mathbf{x}_i\,, \quad \text{per ogni } i = 1, \ldots, N\,.$$

Gli spostamenti virtuali del tipo:

$$\delta P_i = \delta q^k\, \mathbf{e}_k\,, \quad \text{per ogni } i = 1, \ldots, N\,,$$

non sono altro che traslazioni rigide dell'intero sistema l'ungo l'asse $\mathbf{e}_k$, per una distanza δq^k. Gli spostamenti virtuali del tipo:

$$\delta P_i = \delta q^{k+3}\, \mathbf{e}_k \wedge \mathbf{x}_i\,, \quad \text{per ogni } i = 1, \ldots, N\,,$$

non sono altro che "rotazioni infinitesime" rigide dell'intero sistema di un angolo δq^{k+3} attorno all'asse $\mathbf{e}_k$.
Se $\boldsymbol{\phi}_{ij}$ è la reazione vincolare sul punti i-esimo dovuta al punto j-esimo ($j \neq i$), il lavoro virtuale si scrive:

$$\delta L = \sum_{i,j\ i\neq j} \sum_{k=1}^{3} \delta q^k\, \mathbf{e}_k \cdot \boldsymbol{\phi}_{ij} + \sum_{i,j\ i\neq j} \sum_{k=1}^{3} \delta q^{k+3}\, \mathbf{e}_k \wedge \mathbf{x}_i \cdot \boldsymbol{\phi}_{ij}\,.$$

Ossia, con banali passaggi (ricordando che $\mathbf{a} \wedge \mathbf{b} \cdot \mathbf{c} = \mathbf{b} \wedge \mathbf{c} \cdot \mathbf{a} = \mathbf{a} \cdot \mathbf{b} \wedge \mathbf{c}$):

$$\delta L = \sum_{k=1}^{3} \delta q^k \, \mathbf{e}_k \cdot \sum_{i,j \; i \neq j}^{3} \boldsymbol{\phi}_{ij} + \sum_{k=1}^{3} \delta q^{k+3} \, \mathbf{e}_k \cdot \sum_{i,j \; i \neq j} \mathbf{x}_i \wedge \boldsymbol{\phi}_{ij} \, .$$

Ma si prova facilmente che

$$\sum_{i,j \; i \neq j} \boldsymbol{\phi}_{ij} = \mathbf{0} \quad \text{e} \quad \sum_{i,j \; i \neq j} \mathbf{x}_i \wedge \boldsymbol{\phi}_{ij} = \mathbf{0} \, ,$$

essendo $\boldsymbol{\phi}_{ij} = -\boldsymbol{\phi}_{ji}$ ed essendo $\boldsymbol{\phi}_{ij}$ diretta lungo $\mathbf{x}_i - \mathbf{x}_j$ (la prova procede come nella dimostrazione del teorema 4.1 in riferimento alle forze interne). Quindi il vincolo di rigidità soddisfa il requisito di vincolo ideale.
(3) Sono ideali vincoli su un sistema di punti materiali costituiti dalla sovrapposizione dei due precedenti casi. Per esempio un sistema di tre punti materiali P_1, P_2, P_3, con P_1, P_2 connessi al punto P_3 tramite aste rigide prive di massa, assumendo che gli angoli tra le aste in P_3 siano variabili a piacimento (si può anche richiedere che le aste stiano in un piano fissato oppure no) a causa di cerniere ideali senza attrito (cioé senza momenti delle forze d'attrito). Infine il punto P_3 è vincolato a scorrere su una curva assegnata, liscia ed in quiete in un riferimento.
Come caso similare si può considerare un sistema di punti materiali $\mathfrak{S}$ soggetto al vincolo di rigidità e connesso, tramite uno dei punti materiali di $\mathfrak{S}$ e per mezzo di un asta rigida di massa nulla, ad un punto geometrico in quiete con un riferimento. Le connessioni devono avvenire tramite cerniere ideali senza momenti delle forze d'attrito. Questo pendolo soddisfa ancora la richiesta di vincoli ideali. In particolare, la reazione vincolare sul sistema dovuta all'asta rigida fornisce un contributo nullo al computo del lavoro virtuale, dato che ogni spostamento virtuale è normale all'asse detto mentre la reazione vincolare ne è parallela (potrebbe non esserlo se l'asta fosse dotata di massa).
(4) Si può infine considerare il caso di punti materiali vincolati a curve o superfici lisce, la cui forma si modifica nel tempo. Il calcolo del lavoro virtuale, dato che viene effettuato a tempo fissato, non è soggetto a modifiche rispetto a caso di curve o superfici lisce di forma fissata nel tempo: il lavoro virtuale risulta sempre essere nullo.
(5) Citiamo ancora il caso di *vincoli di rotolamento integrabili* che avremo occasione di incontrare pù avanti.

6.1.3 Grandezze cinematiche ed energia cinetica.

Consideriamo il solito sistema di N punti materiali $P_1, \ldots, P_N$ sottoposto a $C < 3N$ vincoli olonomi ideali e descritto nello spaziotempo delle configurazioni $\mathbb{V}^{n+1}$. usando coordinate naturali $t, q^1, \ldots, q^n$, possiamo esprimere le posizioni dei punti come $P_i = P_i(t, q^1, \ldots, q^n)$. Una linea di universo $\Gamma = \Gamma(t) \in \mathbb{V}^{3N+1}$ che rispetta i vincoli, può essere descritta localmente in termini delle coordinate naturali tramite una curva differenziabile $q^k = q^k(t)$. Fissato un riferimento $\mathscr{I}$,

se $P_i(t) = \mathbf{x}_i(t) + O$ dove O è in quiete nel riferimento, possiamo scrivere:

$$\mathbf{v}_{P_i}|_{\mathscr{I}}(t) = \frac{\partial \mathbf{x}_i}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathbf{x}_i}{\partial q^k} \frac{dq^k}{dt} \,. \tag{6.14}$$

In questo modo, possiamo esprimere l'energia cinetica del sistema rispetto al riferimento $\mathscr{I}$, usando le coordinate naturali. Tenendo conto che $\mathscr{T}|_{\mathscr{I}} = \sum_i \frac{1}{2} m_{P_i} \mathbf{v}_{P_i}|^2_{\mathscr{I}}$, si ha immediatamente:

$$\mathscr{T}|_{\mathscr{I}} = \mathscr{T}_2|_{\mathscr{I}} + \mathscr{T}_1|_{\mathscr{I}} + \mathscr{T}_0|_{\mathscr{I}} \tag{6.15}$$

dove:

$$\mathscr{T}_2|_{\mathscr{I}} := \sum_{h,k=1}^{n} a_{hk}(t, q^1, \ldots, q^n) \frac{dq^h}{dt} \frac{dq^k}{dt}, \quad \mathscr{T}_1|_{\mathscr{I}} := \sum_{k=1}^{n} b_k(t, q^1, \ldots, q^n) \frac{dq^k}{dt}, \quad \mathscr{T}_0|_{\mathscr{I}} := c(t, q^1, \ldots, q^n),$$

con le definizioni:

$$a_{hk}(t, q^1, \ldots, q^n) := \frac{1}{2} \sum_{i=1}^{N} m_i \frac{\partial \mathbf{x}_i}{\partial q^h} \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} \,, \tag{6.16}$$

$$b_k(t, q^1, \ldots, q^n) := \sum_{i=1}^{N} m_i \frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \frac{\partial \mathbf{x}_i}{\partial t} \,, \tag{6.17}$$

$$c(t, q^1, \ldots, q^n) := \frac{1}{2} \sum_{i=1}^{N} m_i \frac{\partial \mathbf{x}_i}{\partial t} \cdot \frac{\partial \mathbf{x}_i}{\partial t} \,. \tag{6.18}$$

6.2 Equazioni di Eulero-Lagrange e loro proprietà elementari.

Consideriamo il solito sistema di N punti materiali $P_1, \ldots, P_N$, con masse $m_1, \ldots, m_N$ rispettivamente, sottoposto a $C < 3N$ vincoli olonomi ideali e descritto nello spaziotempo delle configurazioni $\mathbb{V}^{n+1}$. Ad ogni linea di universo $\Gamma = (\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t)) \in \mathbb{V}^{3N+1}$ che soddisfi i vincoli imposti a $\mathfrak{S}$ corrisponde biunivocamente una curva $\gamma = \gamma(t) \in \mathbb{V}^{n+1}$ espressa usando, localmente, coordinate libere $q^k = q^k(t)$, $k = 1, \ldots, n$. Vale cioé $\mathbf{x}_i(t) = \mathbf{x}_i(q^1(t), \ldots, q^n(t))$ per ogni $i = 1, \ldots, N$. Tenendo conto di ciò , vale il seguente teorema.

Teorema 6.1. *Consideriamo un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$, con masse $m_1, \ldots, m_N$ rispettivamente, sottoposto a $C < 3N$ vincoli olonomi descritto nello spaziotempo delle configurazioni $\mathbb{V}^{n+1}$. Supponiamo che, fissato un riferimento $\mathscr{I}$, sul generico i-esimo punto agisca la forza attiva $\mathbf{F}_{\mathscr{I}i}(t, P_1, \ldots, P_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}})$.*
Una linea di universo $\Gamma = \Gamma(t) \in \mathbb{V}^{3N+1}$ del sistema $\mathfrak{S}$ soddisfa le equazioni del moto di Newton (6.8) con reazioni vincolari ideali se e solo se, la curva $\gamma = \gamma(t) \in \mathbb{V}^{n+1}$ che corrisponde a

Γ *ed è descritta in coordinate locali naturali da* $q^k = q^k(t)$*, soddisfa le* **equazioni di Eulero-Lagrange***:*

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^k} - \dfrac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^k} & = \mathscr{Q}_k|_{\mathscr{I}}\ , \\ \dfrac{dq^k}{dt} & = \dot{q}^k(t) \end{cases} \quad \textit{per } k = 1, \ldots, n, \tag{6.19}$$

essendo $\mathscr{T}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ *l'energia cinetica del sistema* $\mathfrak{S}$ *rispetto al riferimento* $\mathscr{I}$ *(6.15) e* $\mathscr{Q}_k|_{\mathscr{I}}$ *le* **componenti lagrangiane delle forze attive***:*

$$\mathscr{Q}_k|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \sum_{i=1}^{N} \mathbf{F}_{\mathscr{I}i}(t, q^1, \ldots q^n, \dot{q}^1, \ldots, \dot{q}^n) \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} \tag{6.20}$$

dove abbiamo usato le relazioni (6.14) per esprimere le velocità dei punti di $\mathfrak{S}$ *in termini delle* q^k *e delle loro derivate temporali* $\dot{q}^k$*, pensate come variabili indipedenti dalle* q^k *in (6.19).* $\diamondsuit$

Dimostrazione. Abbiamo bisogno di due risultati preliminari. Ecco il primo. Usando le equazioni di Newton (6.8), la caratterizzazione di vincoli ideali (valide sulla configurazione $\mathcal{C}_t$ raggiunta dalla linea di universo del sistema $\mathfrak{S}$ che omettiamo di scrivere per semplicità notazionale):

$$\sum_{i=1}^{N} \boldsymbol{\phi}_i \cdot \delta P_i = 0\ , \quad \text{per ogni spostamento virtuale } \delta P, \tag{6.21}$$

si può equivalentemente riscrivere nella seguente forma[4]:

$$\sum_{i=1}^{N} (m_{P_i} \mathbf{a}_{P_i}|_{\mathscr{I}} - \mathbf{F}_{\mathscr{I}i}) \cdot \delta P_i = 0\ , \quad \text{per ogni spostamento virtuale } \delta P\ .$$

Tenendo conto della (6.9) e ponendo $\mathbf{x}_i := P_i - O$ con O in quiete nel riferimento $\mathscr{I}$, si ha l'ulteriore equazione equivalente alla (6.21):

$$\sum_{k=1}^{n} \left(\sum_{i=1}^{N} (m_{P_i} \mathbf{a}_{P_i}|_{\mathscr{I}} - \mathbf{F}_{\mathscr{I}i}) \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} \right) \delta q^k = 0\ , \quad \text{per ogni scelta di } \delta q^1, \ldots, \delta q^k \in \mathbb{R}\ . \tag{6.22}$$

Infine (6.22) è equivalente alle n condizioni (valide sulla configurazione $\mathcal{C}_t$ raggiunta dalla linea di universo del sistema $\mathfrak{S}$):

$$\sum_{i=1}^{N} (m_{P_i} \mathbf{a}_{P_i}|_{\mathscr{I}} - \mathbf{F}_{\mathscr{I}i}) \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} = 0\ , \quad \text{per ogni } k = 1, \ldots, n\ . \tag{6.23}$$

[4]Questa equazione porta il nome un pò pomposo di **equazione simbolica della dinamica** o anche **principio di D'Alembert** o anche **principio dei lavori virtuali**.

Se infatti vale (6.23) la (6.22) è evidentemente verificata. Se viceversa vale (6.22), scegliendo $\delta q^k = \delta_{kr}$, si ottiene che (6.23) deve essere verificata per $k = r$. Dato che $r = 1, \ldots, n$ è generico, questo mostra che (6.23) è verificata.
Come secondo risultato mostriamo che:

$$\frac{d}{dt}\frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^k} - \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^k} = \sum_{i=1}^{N} m_i \mathbf{a}_{P_i}|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} \,, \tag{6.24}$$

dove abbiamo imposto il vincolo $\dot{q}^k = \frac{dq^k}{dt}$ dopo aver calcolato il primo membro. Per provare l'identità suddetta, notiamo che essa si esplicita in

$$\sum_i \frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}\left(\frac{1}{2} m_i \mathbf{v}_{P_i}|^2_{\mathscr{I}}\right) - \sum_i \frac{\partial}{\partial q^k}\left(\frac{1}{2} m_i \mathbf{v}_{P_i}|^2_{\mathscr{I}}\right) = \sum_i m_i \mathbf{a}_{P_i}|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} \,. \tag{6.25}$$

Proviamo la validità di questa identità . Dato che

$$\mathbf{v}_{P_i}|^2_{\mathscr{I}} = \left(\sum_r \frac{\partial \mathbf{x}_i}{\partial q^r}\dot{q}^r + \frac{\partial \mathbf{x}_i}{\partial t}\right) \cdot \left(\sum_h \frac{\partial \mathbf{x}_i}{\partial q^h}\dot{q}^h + \frac{\partial \mathbf{x}_i}{\partial t}\right) \,, \tag{6.26}$$

si ha:

$$\frac{\partial}{\partial \dot{q}^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} = 2\frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \left(\sum_h \frac{\partial \mathbf{x}_i}{\partial q^h}\dot{q}^h + \frac{\partial \mathbf{x}_i}{\partial t}\right) = 2\frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \mathbf{v}_{P_i}|_{\mathscr{I}} \,,$$

e quindi

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} = 2\left(\frac{d}{dt}\frac{\partial \mathbf{x}_i}{\partial q^k}\right) \cdot \mathbf{v}_{P_i}|_{\mathscr{I}} + 2\frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \mathbf{a}_{P_i}|_{\mathscr{I}} \,.$$

D'altra parte vale anche:

$$2\left(\frac{d}{dt}\frac{\partial \mathbf{x}_i}{\partial q^k}\right) \cdot \mathbf{v}_{P_i}|_{\mathscr{I}} = 2\sum_h \frac{\partial^2 \mathbf{x}_i}{\partial q^k \partial q^h}\dot{q}^h \cdot \mathbf{v}_{P_i}|_{\mathscr{I}} + 2\frac{\partial^2 \mathbf{x}_i}{\partial q^k \partial t} \cdot \mathbf{v}_{P_i}|_{\mathscr{I}} = \frac{\partial}{\partial q^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} \,,$$

dove l'ultima identità segue immediatamente da (6.26). In definitiva:

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} = \frac{\partial}{\partial q^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} + 2\frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \mathbf{a}_{P_i}|_{\mathscr{I}} \,,$$

che equivale a

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} - \frac{\partial}{\partial q^k}\mathbf{v}_{P_i}|^2_{\mathscr{I}} = 2\frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \mathbf{a}_{P_i}|_{\mathscr{I}} \,.$$

Moltiplicando per $m_i/2$ si trova infine:

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}\left(\frac{1}{2} m_i \mathbf{v}_{P_i}|^2_{\mathscr{I}}\right) - \frac{\partial}{\partial q^k}\left(\frac{1}{2} m_i \mathbf{v}_{P_i}|^2_{\mathscr{I}}\right) = m_i \mathbf{a}_{P_i}|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} \,.$$

Sommando su i si ottiene (6.25), cioé (6.24), immediatamente.
Quindi le equazioni di Eulero Lagrange si possono riscrivere, tenendo conto della definizione (6.20) delle componenti lagrangiane delle forze attive $\mathscr{Q}_k$:

$$\begin{cases} \displaystyle\sum_{i=1}^{N} m_i \mathbf{a}_{P_i}|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} & = \displaystyle\sum_{i=1}^{N} \mathbf{F}_{\mathscr{I}i}(t, q^1, \dots q^n, \dot{q}^1, \dots, \dot{q}^n) \cdot \frac{\partial \mathbf{x}_i}{\partial q^k}\,, \\ \displaystyle\frac{dq^k}{dt} & = \dot{q}^k(t) \end{cases} \quad \text{per } k = 1, \dots, n,$$

ovvero

$$\begin{cases} \displaystyle\sum_{i=1}^{N} \big(m_i \mathbf{a}_{P_i}|_{\mathscr{I}} - \mathbf{F}_{\mathscr{I}i}(t, q^1, \dots q^n, \dot{q}^1, \dots, \dot{q}^n)\big) \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} & = 0\,, \\ \displaystyle\frac{dq^k}{dt} & = \dot{q}^k(t) \end{cases} \quad \text{per } k = 1, \dots, n, \tag{6.27}$$

Supponiamo ora che $q^k = q^k(t)$ definisca una linea di universo $\Gamma = \Gamma(t) \in \mathbb{V}^{3N+1}$ per il sistema $\mathfrak{S}$ che soddisfi le equazioni di Newton (6.8) con vincoli ideali. In tal caso $m_i \mathbf{a}_{P_i}|_{\mathscr{I}} - \mathbf{F}_{\mathscr{I}i} = \boldsymbol{\phi}_i$ è la reazione vincolare sull'i-esimo punto ed il sistema (6.27) di sopra è verificato in virtù della caratterizzazione (6.23) della richiesta di vincoli ideali.
Supponiamo viceversa che $q^k = q^k(t)$ definisca una linea di universo $\Gamma = \Gamma(t) \in \mathbb{V}^{3N+1}$ per il sistema $\mathfrak{S}$ che soddisfi le equazioni di Eulero-Lagrange (6.27). In tal caso, banalmente, vale la (6.8), dove le ϕ_i soddisfano la (6.23) cioé la caratterizzazione di vincolo ideale (6.21). □

Osservazioni 6.3.
(1) Si deve sottolineare che le componenti lagrangiane delle forze attive sono *indipendenti* dal riferimento inerziale $\mathscr{I}$ usato per definire i vettori posizione $\mathbf{x}_i$ quando le forze attive considerate sono *forze vere*. Cambiando riferimento ed usando vettori posizione $\mathbf{x}'_i$ riferiti al nuovo riferimento inerziale $\mathscr{I}'$, si ha infatti $\mathbf{x}'_i(t, q^1, \dots, q^n) = \mathbf{x}_i(t, q^1, \dots, q^n) + \mathbf{X}(t)$, dove $\mathbf{X}(t) = O'(t) - O(t)$ è la funzione, *assegnata* che descrive il moto relativo delle due origini delle coordinate solidali con $\mathscr{I}$ e $\mathscr{I}'$. Abbiamo allora che

$$\mathscr{Q}_k|_{\mathscr{I}'}(t, q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n) := \sum_{i=1}^{N} \mathbf{F}_{\mathscr{I}'i}(t, q^1, \dots q^n, \dot{q}^1, \dots, \dot{q}^n) \cdot \frac{\partial \mathbf{x}'_i}{\partial q^k}$$

$$= \sum_{i=1}^{N} \mathbf{F}_{\mathscr{I}i}(t, q^1, \dots q^n, \dot{q}^1, \dots, \dot{q}^n) \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} = \mathscr{Q}_k|_{\mathscr{I}}(t, q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n)\,,$$

dove si è tenuto conto del fatto che le forze vere hanno valori che non dipendono dal riferimento. Nel caso si lavori in un riferimento non inerziale, una parte delle forze attive è di natura inerziale e pertanto *dipendente* dal riferimento. Tale dipendenza si riperquote sulle componenti lagrangiane delle forze attive.
(2) È molto importante sottolineare che le variabili $\dot{q}^k$ e le variabili q^k sono pensate come

indipendenti e diventano dipendenti solo al momento di imporre il sistema delle equazioni di Eulero-Lagrange, che è di fatto un sistema del primo ordine. Senza assumere l'indipendenza delle $\dot{q}^k$ dalle q^k la prima riga in (6.19) sarebbe anche di difficile interpretazione.
(3) La superiorità delle equazioni di Eulero-Lagrange rispetto a quelle di Newton si basa principalmente sul fatto che le prime sono automaticamente *equazioni pure di movimento* non contenendo le reazioni vincolari ed essendo, come vedremo tra poco, riscrivibili in forma normale e quindi soddisfacendo l'ipotesi principale del teorema di esistenza ed unicità . Dal punto di vista pratico le equazioni di Eulero-Lagrange sono scrivibili in riferimento ad un qualsiasi sistema di coordinate locali adattate alla foliazione. Questo solo fatto semplifica enormemente la matematica del problema.
(4) Come provato nella dimostrazione, la condizione di idealità dei vincoli (sulle configurazioni raggiunte dalle una linea di universo del sistema fisico) equivale alla (6.23). Quest'ultima può ancora essere trascritta come:

$$\sum_{i=1}^{N} \boldsymbol{\phi}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q^k} = 0 \,,$$

che a sua volta si può illustrare dicendo che: *le componenti lagrangiane delle reazioni vincolari ideali sono nulle (sulle configurazioni raggiunte dalle una linea di universo del sistema fisico).*

6.2.1 Normalità delle equazioni di Eulero-Lagrange.

Mostriamo ora che le equazioni di Eulero-Lagrange sono sempre scrivibili in forma normale. Questo risultato, come sappiamo, assicura che le equazioni ammettano *una ed una sola soluzione per fissate condizioni iniziali*, sotto debite ipotesi di regolarità delle funzioni note che appaiono nelle equazioni di Eulero-Lagrange stesse.
Tenendo conto della forma (6.15) dell'energia cinetica, le equazioni di Eulero-Lagrange hanno forma esplicita (notare che $a_{hk} = a_{kh}$):

$$\sum_{h=1}^{n} 2a_{kh}(t, q^1(t), \ldots, q^n(t)) \frac{d^2 q^h}{dt^2} = G_k \left(t, q^1(t), \ldots, q^n(t), \frac{dq^1}{dt}, \ldots, \frac{dq^n}{dt} \right) \,, \; k = 1, \ldots, n \,, \quad (6.28)$$

dove le funzioni G_k si ottengono sommando l'espressione esplicita delle $\mathscr{Q}_k$ a $-2\sum_h \frac{da_{kh}}{dt}\frac{dq^h}{dt}$ ed alle derivate dei termini $\mathscr{T}_1|_{\mathscr{I}}$, $\mathscr{T}_0|_{\mathscr{I}}$. Se i vincoli sono funzioni almeno C^3 e le forze attive sono funzioni almeno C^1, i coefficienti $a_{kh} = a_{kh}(t, q^1, \ldots, q^n)$ ed il secondo membro di (6.28) sono almeno di classe C^1. Se proviamo che la matrice quadrata $a(t, q^1, \ldots, q^n)$ dei coefficienti $a_{kh}(t, q^1, \ldots, q^n)$ è invertibile (cioé $deta \neq 0$) per ogni scelta di $(t, q^1, \ldots, q^n)$, abbiamo automaticamente provato che le equazioni di Eulero-Lagrange si possono scrivere in forma normale con secondo membro di classe C^1 – la funzione $(a^{-1})_{kh} = (a^{-1})_{kh}(t, q^1, \ldots, q^n)$ risulta essere anch'essa di classe C^1 se esiste, come si prova facilmente dall'espressione esplicita della matrice inversa – e questo assicura la validità del teorema di esistenza ed unicità del problema di Cauchy connesso. Consideriamo dunque il *nucleo* della matrice quadrata a. Vogliamo mostrare che

contiene solo il vettore nullo, cioé che a è biettiva. $(c^1, \dots, c^n) \in \mathbb{R}^m$ è nel nucleo di a se e solo se $\sum_{h=1}^{n} a_{kh} c^h = 0$ per ogni $k = 1, \dots, n$. In tal caso vale anche: $\sum_{h,k=1}^{n} c^k c^h a_{kh} = 0$. Dall'espressione (6.16) per la matrice a, questo è equivalente a dire che:

$$\sum_{i=1}^{N} m_i \left(\sum_h c^h \frac{\partial \mathbf{x}_i}{\partial q^h} \right) \cdot \left(\sum_k c^k \frac{\partial \mathbf{x}_i}{\partial q^k} \right) = 0 \,, \quad \text{cioé} \quad \sum_{i=1}^{N} m_i \left\| \sum_h c^h \frac{\partial \mathbf{x}_i}{\partial q^h} \right\|^2 = 0 \,.$$

Dato che le masse m_i sono strettamente positive, giungiamo alla conclusione che deve essere

$$\sum_{h=1}^{n} c^h \frac{\partial x_i^j}{\partial q^h} = \mathbf{0} \,, \quad \text{per ogni } i = 1, \dots N \text{ e } j = 1, 2, 3. \tag{6.29}$$

Mostriamo che (6.29) implica: $c^1 = c^2 = \cdots = c^n = 0$, cioé il nucleo di a contiene solo il vettore nullo. Lasciamo al lettore la prova del fatto che, se si passa a nuove coordinate libere $q'^h = q'^h(t, q^1, \dots, q^n)$ (unitamente a $t' = t+$ costante) in modo da individuare una trasformazione locale di coordinate differenziabile, biettiva e con inversa differenziabile, la validità di (6.29) implica la validità delle analoghe equazioni che si ottengono esprimendo i vettori $\mathbf{x}_i$ in funzione di $t', q'^1, \dots, q'^n$. Dalla Proposizione 6.1 risulta che le coordinate $t, q^1, \dots, q^n$ possono sempre scegliersi localmente come: il tempo assoluto t e n tra le coordinate x_i^j. Senza perdere generalità (basta cambiare il nome alle coordinate x_i^j) possiamo sempre assumere che le coordinate libere siano $q^1 := x_1^1$, $q^2 := x_2^2$ e via di seguito fino a $q^n := x_n^n$. Scegliendo $x_i^j = x_1^1$, la (6.29) si riduce a $c_1 = 0$, scegliendo $x_i^j = x_2^2$, la (6.29) si riduce a $c_2 = 0$ e via di seguito fino a $c_n = 0$. Quindi la matrice quadrata a è non singolare. Si osservi che valendo:

$$\sum_{h,k=1}^{n} c^k c^h a_{kh} = \sum_{i=1}^{N} m_i \left\| \sum_h c^h \frac{\partial \mathbf{x}_i}{\partial q^h} \right\|^2 \geq 0 \,,$$

abbiamo anche provato che la matrice dei coefficienti a_{kh} è *strettamente definita positiva.*

Osservazioni 6.4.* Un modo più rapido di arrivare a $c^1 = c^2 = \cdots c^n = 0$ è il seguente, che però usa la nozione di vettore tangente ad una varietà . L'indipendenza funzionale dei vincoli assicura che $q^1, \dots, q^n$ siano coordinate sulla sottovarietà $\mathbb{Q}_t$. usando le coordinate cartesiane su Σ_t^N, $x_1^1, x_1^2, x_1^3, \dots, x_N^1, x_N^2, x_N^3$, l'identità (6.29) si riscrive: $\sum_{h=1}^{n} c^h \sum_{j,i} \frac{\partial x_i^j}{\partial q^h} \frac{\partial}{\partial x_i^j} = \mathbf{0}$, che possiamo infine riscrivere usando la (A.17): $\sum_{h=1}^{n} c^h \frac{\partial}{\partial q^h} = \mathbf{0}$. Questa identità è possibile solo se tutti i coefficienti c^h sono nulli, dato che i vettori $\frac{\partial}{\partial q^h}$ formano una base dello spazio tangente a $T_{\mathfrak{C}_t}\mathbb{Q}_t$ e sono pertanto linearmente indipendenti.

Abbiamo provato la seguente proposizione che tiene conto del teorema 3.3 e della proposizione 3.2.

Proposizione 6.2. *Nelle ipotesi del teorema 6.1, in ogni sistema di coordinate locali naturali su $\mathbb{V}^{n+1}$ adattate alle fibre, $(t, q^1, \ldots, q^n)$, la matrice quadrata $n \times n$, definita in (6.16), dei coefficienti $a_{kh}(t, q^1, \ldots, q^n)$ è strettamente definita positiva per ogni scelta di $(t, q^1, \ldots, q^n)$. Di conseguenza le equazioni di Eulero-Lagrange (6.19) sono sempre scrivibili in forma normale:*

$$\frac{d^2 q^k}{dt^2} = z^k \left(t, q^1(t), \ldots, q^n(t), \frac{dq^1}{dt}(t), \ldots, \frac{dq^n}{dt}(t) \right) \quad \textit{per ogni } k = 1, \ldots, n. \tag{6.30}$$

Il secondo membro dell'identità di sopra è una funzione di classe C^1 congiuntamente in tutte le sue variabili quando i vincoli sono funzioni almeno C^3 su $\mathbb{V}^{N+1}$ e le componenti lagrangiane $\mathscr{Q}_k$ delle forze attive sono funzioni almeno C^1 delle coordinate $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots \dot{q}^n)$ congiuntamente (oppure, equivalentemente, le forze attive $\mathbf{F}_{\mathscr{I}i}$ sono funzioni almeno C^1 del tempo, delle posizioni e delle velocità congiuntamente dei punti del sistema). In tal caso vale il teorema di esistenza ed unicità locale per il sistema di equazioni di Eulero-Lagrange (6.19) se sono assegnati dati di Cauchy $q^k(t_0) = q_0^k$, $\dot{q}^k(t_0) = \dot{q}_0^k$ per $k = 1, \ldots, n$.

6.2.2 Spaziotempo degli atti di moto ed invarianza delle equazioni di Eulero-Lagrange.

La forma delle equazioni di Eulero-Lagrange (6.19) deve essere la stessa indipendentemente dalle coordinate naturali usate localmente su $\mathbb{V}^{n+1}$. Questo segue dal fatto che noi non abbiamo scelto un particolare sistema, ma siamo rimasti con la scelta generica. Vogliamo esaminare più a fondo la questione. Per procedere consideriamo dunque due insiemi aperti $U, V \subset \mathbb{V}^{n+1}$, dotati di coordinate naturali $t, q^1, \ldots, q^n$ e $t', q'^1, \ldots, q'^n$ rispettivamente e con $U \cap V \neq \emptyset$. Le equazioni di Eulero-Lagrange varranno sia su U che su V e quindi, in particolare su $U \cap V$. In tale insieme, i due membri di (6.19) riferiti a coordinate non primate possono essere confrontati con i corrispondenti riferiti alle coordinate primate. Per eseguire il confronto dobbiamo conoscere la relazione che intercorre tra le coordinate usate. Riguardo alle coordinate non puntate valgano le leggi di trasformazione (6.3)-(6.4). Le coordinate puntate hanno uno status più complesso. Le coordinate puntate sono infatti considerate *indipendenti* da quelle puntate nelle equazioni di Eulero-Lagrange: esse diventano dipendenti solo sulle soluzioni delle equazioni. Se $q^k = q^k(t)$ soddisfa le equazioni di Eulero-Lagrange avremo che:

$$\dot{q}'^k(t) = \frac{dq'^k}{dt'} = \sum_{h=1}^{n} \frac{\partial q'^k}{\partial q^h} \frac{dq^h}{dt} + \frac{\partial q'^k}{\partial t} = \sum_{h=1}^{n} \frac{\partial q'^k}{\partial q^h} \dot{q}^h(t) + \frac{\partial q'^k}{\partial t} \, .$$

Possiamo assumere che queste relazioni valgano *anche se non è stato fissato alcun moto del sistema*, visto che su ogni moto si riducono alle corrette relazioni. Assumeremo quindi che, nella

situazione detta:

$$
\begin{aligned}
t' &= t + c\,, && (6.31)\\
q'^k &= q'^k(t, q^1, \ldots, q^n)\,, && (6.32)\\
\dot{q}'^k &= \sum_{h=1}^{n} \frac{\partial q'^k}{\partial q^h}\dot{q}^h + \frac{\partial q'^k}{\partial t}\,. && (6.33)
\end{aligned}
$$

Si osservi che ne segue l'identità utile nella pratica:

$$
\frac{\partial \dot{q}'^k}{\partial \dot{q}^h} = \frac{\partial q'^k}{\partial q^h}\,, \tag{6.34}
$$

e anche, su ogni curva $q'^k = q'^k(t)$ ed identificando le variabili puntate con le derivate temporali di quelle non puntate:

$$
\left(\frac{d}{dt}\frac{\partial q^h}{\partial q'^k}\right) = \frac{\partial \dot{q}^h}{\partial q'^k}\,. \tag{6.35}
$$

Assumendo valide le trasformazioni (6.31),(6.32) e (6.33) seguono subito le ulteriori utili identità :

$$
\begin{aligned}
\frac{\partial}{\partial t'} &= \frac{\partial}{\partial t} + \sum_{h=1}^{n} \frac{\partial q^h}{\partial t'}\frac{\partial}{\partial q^h} + \sum_{h=1}^{n} \frac{\partial \dot{q}^h}{\partial t'}\frac{\partial}{\partial \dot{q}^h}\,, && (6.36)\\
\frac{\partial}{\partial \dot{q}'^k} &= \sum_{h=1}^{n} \frac{\partial \dot{q}^h}{\partial \dot{q}'^k}\frac{\partial}{\partial \dot{q}^h} = \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k}\frac{\partial}{\partial \dot{q}^h}\,, && (6.37)\\
\frac{\partial}{\partial q'^k} &= \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k}\frac{\partial}{\partial q^h} + \sum_{h=1}^{n} \frac{\partial \dot{q}^h}{\partial q'^k}\frac{\partial}{\partial \dot{q}^h}\,. && (6.38)
\end{aligned}
$$

Diamo la definizione di *spaziotempo degli atti di moto*. La costruzione ed il significato e la rigoroso di tale spazio nell'ambito della teoria dei fibrati sono presentati nella sezione 6.3.

Definizione 6.3. **(Spaziotempo degli atti di moto).** Lo **spaziotempo degli atti di moto** $j^1(\mathbb{V}^{n+1})$ è una varietà differenziabile di dimensione $2n+1$ che ammette un atlante dato da tutte le carte locali con coordinate della forma $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ che si trasformano con trasformazioni di coordinate differenziabili con inversa differenziabile date in (6.31), (6.32), (6.33) e dove:

(i) $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \in V$ sottoinsieme aperto di $\mathbb{R}^{n+1}$ sono coordinate locali naturali su $\mathbb{V}^{n+1}$,

(ii) $(\dot{q}^1, \ldots, \dot{q}^n) \in \mathbb{R}^n$,

I sistemi di coordinate locali suddetti si dicono **sistemi di coordinate locali naturali di** $j^1(\mathbb{V}^{n+1})$. La coordinata t si identifica nel dominio di ogni sistema coordinate locali naturali, a meno di costanti additive, con il tempo assoluto visto come funzione suriettiva differenziabile Lo **spazio degli atti di moto al tempo** t è ogni sottovarietà (embedded) $2n$-dimensionale $\mathbb{A}_t$

di $j^1(\mathbb{V}^{n+1})$ che si ottiene fissando un valore di tempo assoluto t. Un **atto di moto al tempo** t, $\mathcal{A}_t$, è quindi ogni punto di $\mathbb{A}_t$. $\diamondsuit$

Osservazioni 6.5.
(1) Si osservi che le n-ple $(q^1, \ldots, q^n)$ variano in insiemi aperti di $\mathbb{R}^n$ la cui forma può essere diversa a seconda del caso, viceversa le n-ple $(\dot{q}^1, \ldots, \dot{q}^n)$ variano sempre su *tutto* $\mathbb{R}^n$.
(2) Ogni spazio degli atti di moto al tempo t, $\mathbb{A}_t$, si dice anche **fibra di** $j^1(\mathbb{V}^{n+1})$ **al tempo** t pertanto i sistemi di coordinate locali naturali su $j^1(\mathbb{V}^{n+1})$ sono anche detti sistemi di coordinate locali naturali **adattati alle fibre** di $j^1(\mathbb{V}^{n+1})$. Nelle ipotesi fatte si dimostra facilmente (il lettore lo provi per esercizio!) che, per ogni fissata configurazione $\mathcal{C}_t$ al tempo t, considerando due sistemi di coordinate naturali locali definite attorno a tale configurazione, la matrice di coefficienti $\frac{\partial q'^k}{\partial q^h}|_{\mathcal{C}_t}$ è non singolare (ha determinante non nullo).
(3) Le equazioni di Eulero-Lagrange sono equazioni differenziali, per il momento definite localmente, in $j^1(\mathbb{V}^{n+1})$ che determinano, localmente, curve di classe C^1, $I \ni t \mapsto \gamma(t) \subset j^1(\mathbb{V}^{n+1})$, parametrizzate nella coordinata privilegiata data dal tempo assoluto t. In coordinate locali naturali si ha quindi:

$$I \ni t \mapsto (t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))$$

Questo tipo di curve, tenendo conto della struttura geometricodifferenziale di $j^1(\mathbb{V}^{n+1})$ si dicono anche **sezioni** (vedi il paragrafo 6.3).

Si ha seguente proposizione sull'invarianza delle equazioni di Eulero-Lagrange.

Proposizione 6.3. *Nelle ipotesi del teorema 6.1, siano $U, V \subset \mathbb{V}^{n+1}$ sono insiemi aperti, dotati di coordinate naturali $t, q^1, \ldots, q^n$ e $t', q'^1, \ldots, q'^n$ rispettivamente e con $U \cap V \neq \emptyset$, e pertanto valgano le leggi di trasformazione (6.32),(6.31), (6.33). In tal caso, fissato un riferimento $\mathscr{I}$, con ovvie notazioni valgono le relazioni su $U \cap V$:*

$$\mathscr{T}'|_{\mathscr{I}}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n) = \mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)\,, \tag{6.39}$$

inoltre:

$$\mathscr{Q}'_k|_{\mathscr{I}}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n) = \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} \mathscr{Q}_h|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)\,. \tag{6.40}$$

Infine, su ogni curva $q^k = q^k(t)$ (per $k = 1, \ldots, n$) non necessariamente soddisfacente le equazioni di Eulero-Lagrange, vale ad ogni istante:

$$\frac{d}{dt'} \frac{\partial \mathscr{T}'|_{\mathscr{I}}}{\partial \dot{q}'^k} - \frac{\partial \mathscr{T}'|_{\mathscr{I}}}{\partial q'^k} = \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} \left[\frac{d}{dt} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} - \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^h} \right]\,, \tag{6.41}$$

quando si assuma che $\dot{q}'^k = \frac{dq'^k}{dt}$ e $\dot{q}^h = \frac{dq^h}{dt}$. $\diamondsuit$

Dimostrazione. Le prime due identità si dimostrano per computo diretto usando le definizioni delle grandezze coinvolte e la loro verifica non presenta alcuna difficoltà . La terza identità (6.41) presenta calcoli complicati se non eseguiti opportunamente. Dimostriamola. Il secondo membro si decompone in due parti:

$$-\sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^h} \tag{6.42}$$

e

$$\sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} \frac{d}{dt} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} = \sum_{h=1}^{n} \frac{d}{dt} \left(\frac{\partial q^h}{\partial q'^k} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} \right) - \sum_{h=1}^{n} \left(\frac{d}{dt} \frac{\partial q^h}{\partial q'^k} \right) \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} .$$

Usando (6.34) si ha:

$$\sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} \frac{d}{dt} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} = \sum_{h=1}^{n} \frac{d}{dt} \left(\frac{\partial \dot{q}^h}{\partial \dot{q}'^k} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} \right) - \sum_{h=1}^{n} \left(\frac{d}{dt} \frac{\partial q^h}{\partial q'^k} \right) \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} .$$

Dato che vale (6.37), possiamo ancora scrivere:

$$\sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} \frac{d}{dt} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} = \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}'^k} - \sum_{h=1}^{n} \left(\frac{d}{dt} \frac{\partial q^h}{\partial q'^k} \right) \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} .$$

Sommando il risultato a (6.42), concludiamo che il secondo membro di (6.41) vale:

$$\frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}'^k} - \sum_{h=1}^{n} \left[\left(\frac{d}{dt} \frac{\partial q^h}{\partial q'^k} \right) \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} + \frac{\partial q^h}{\partial q'^k} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^h} \right] .$$

Su ogni curva $q^h = q^h(t)$, pensando le variabili puntate come le derivate temporali di quelle non puntate si può usare (6.35) ottenendo che il secondo membro di (6.41) risulta essere:

$$\frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}'^k} - \sum_{h=1}^{n} \left[\frac{\partial \dot{q}^h}{\partial q'^k} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^h} + \frac{\partial q^h}{\partial q'^k} \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^h} \right] .$$

Dato che, infine, vale (6.38), concludiamo che il secondo membro di (6.41) vale:

$$\frac{d}{dt'} \frac{\partial \mathscr{T}'|_{\mathscr{I}}}{\partial \dot{q}'^k} - \frac{\partial \mathscr{T}'|_{\mathscr{I}}}{\partial q'^k} ,$$

dove si è tenuto conto che la derivazione in t equivale a quella in t' dato che le due variabili differiscono per una semplice costante additiva, per cui (6.41) è verificata. □

Abbiamo l'immediato corollario:

Corollario. *Nelle ipotesi della proposizione 6.3 si consideri la curva* $I \ni t \mapsto \gamma(t) \subset j^1(\mathbb{V}^{n+1})$ *di classe* C^2 *su espressa, in coordinate locali naturali, come:*

$$I \ni t \mapsto (t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))$$

(dove $I \subset \mathbb{R}$ *è un intervallo aperto) che risolve le equazioni di Eulero Lagrange in una carta locale di* $j^1(\mathbb{V}^{n+1})$ *rispetto alle funzioni* $\mathscr{T}|_{\mathscr{I}}, \mathscr{Q}|_{\mathscr{I}}$. *La stessa curva espressa in coordinate* $(t', q'^1, \ldots, q'^n,\ \dot{q}'^1, \ldots,\ \dot{q}'^n)$ *di un'altra carta locale naturale che che include completamente la curva, risolverà ancora le equazioni di Eulero-Lagrange rispetto alle funzioni* $\mathscr{T}'|_{\mathscr{I}}, \mathscr{Q}'|_{\mathscr{I}}$ *sopra indicate.* $\diamondsuit$

Osservazioni 6.6. * Il risultato ha un'importante conseguenza. Noi abbiamo fino ad ora costruito la teoria delle equazioni di Eulero-Lagrange in un sistema di coordinate locali (naturali), sulla varietà differenziabile $j^1(\mathbb{V}^{n+1})$, scelto arbitrariamente. Tuttavia, in generale $j^1(\mathbb{V}^{n+1})$ non è ricopribile con una sola carta locale e considerando una soluzione delle equazioni di E-L in una carta locale (U, ψ), ci aspettiamo che tale soluzione possa arrivare fino al bordo ∂U di tale carta. Dal punto di vista fisico ci si aspetta che la soluzione sia estendibile fuori da tale carta, dato che la carta scelta non ha in generale alcun significato fisico privilegiato. Possiamo provare ad usare diversi sistemi di coordinate incollando di volta in volta le varie soluzioni, ma a priori non è detto che la teoria complessiva risulti coerente. La proposizione dimostrata ci dice che i vari ingredienti usati nella formulazione delle equazioni di E-L, cambiando coordinate, si comportano in modo coerente. Se infatti (V, ϕ) è un'altra carta locale con $V \cap U \neq \emptyset$, la soluzione γ_U suddetta, raggiungerà un punto $\gamma_U(t_1) \in U \cap V$. Possiamo allora reimpostare il problema di Cauchy delle equazioni di E-L ora *trascritte nelle coordinate in* V, usando come condizioni iniziali al tempo t_1, $\gamma_U(t_1)$ e $\dot{\gamma}_U(t_1)$. L'unica soluzione γ_V che otteniamo si estenderà in generale su V *anche fuori da* U. Le curve γ_U e γ_V, per costruzione si raccorderanno differenziabilmente in $U \cap V$. É chiaro che per questa via si ottiene alla fine una soluzione delle equazioni di Eulero-Lagrange definita, potenzialmente, su *tutto* $j^1(\mathbb{V}^{n+1})$.

6.2.3 Lagrangiane.

Supponiamo che le forze attive agenti sul sistema di punti materiali $\mathfrak{S}$ (vincolato da vincoli olonomi ideali) siano tutte conservative nel riferimento $\mathscr{I}$. Esisterà pertanto una funzione energia potenziale:

$$\mathscr{U}|_{\mathscr{I}} = \mathscr{U}|_{\mathscr{I}}(P_1, \ldots, P_N)$$

e varrà

$$\mathbf{F}_{i\mathscr{I}} = -\nabla_{P_i}\mathscr{U}(P_1, \ldots, P_N)\,.$$

Possiamo indebolire la richiesta, supponendo più semplicemente che esista una funzione, detta **potenziale delle forze**, $\mathscr{V}|_{\mathscr{I}} = \mathscr{V}|_{\mathscr{I}}(t, P_1, \ldots, P_N)$, in generale dipendente dal tempo (per cui non vale il teorema di conservazione dell'energia meccanica), per cui valga

$$\mathbf{F}_{i\mathscr{I}} = \nabla_{P_i}\mathscr{V}(t, P_1, \ldots, P_N)\,.$$

In riferimento a coordinate naturali $t, q^1, \ldots, q^n$ su $\mathbb{V}^{n+1}$, e fissata un'origine O solidale con $\mathscr{I}$ in modo che $P_i = \mathbf{x}_i + O$, le componenti lagrangiane delle forze attive avranno forma

$$\mathscr{Q}_k(t, q^1, \ldots, q^n) = \sum_{n=1}^{N} \frac{\partial \mathbf{x}_i}{\partial q^k} \nabla_{\mathbf{x}_i} \mathscr{V}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{\partial}{\partial q^k} \mathscr{V}(t, q^1, \ldots, q^n)\,.$$

Inserendo questa espressione per le componenti lagrangiane delle forze attive nelle equazioni di Eulero-Lagrange, si ha immediatamente che esse possono essere riscritte come:

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial(\mathscr{T}|_{\mathscr{I}} + \mathscr{V}|_{\mathscr{I}})}{\partial \dot{q}^k} - \dfrac{\partial(\mathscr{T}|_{\mathscr{I}} + \mathscr{V}|_{\mathscr{I}})}{\partial q^k} & = 0\,, \\ \dfrac{dq^k}{dt} & = \dot{q}^k(t) \end{cases} \quad \text{per } k = 1, \ldots, n,$$

dove si è tenuto conto del fatto che $\mathscr{V}|_{\mathscr{I}}$ *non* dipende dalle coordinate $\dot{q}^k$. Introducendo la **lagrangiana del sistema**:

$$\mathscr{L}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) + \mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n)\,, \tag{6.43}$$

le equazioni di Eulero-Lagrange prendono forma classica:

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial \mathscr{L}|_{\mathscr{I}}}{\partial \dot{q}^k} - \dfrac{\partial \mathscr{L}|_{\mathscr{I}}}{\partial q^k} & = 0\,, \\ \dfrac{dq^k}{dt} & = \dot{q}^k(t) \end{cases} \quad \text{per } k = 1, \ldots, n, \tag{6.44}$$

Esempi 6.3.
(1) Un caso elementare è in seguente. Consideriamo un sistema di N punti materiali, $P_1, \ldots, P_N$ di masse $m_1, \ldots, m_N$ rispettivamente, *non sottoposti a vincoli*, ma interagenti tramite forze dedotte da un'energia potenziale totale $\mathscr{U}|_{\mathscr{I}}$. $\mathscr{I}$ è un sistema di riferimento inerziale nel quale i punti dono individuati da vettori posizione $\mathbf{x}_1 := P_1 - O, \ldots, \mathbf{x}_n := P_N - O$. O è l'origine di un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ e le componenti di ogni $\mathbf{x}_i$ sono le componenti di $P - O$ in tale sistema di coordinate. La lagrangiana del sistema sarà allora data da, usando le componenti di tutti i vettori $\mathbf{x}_i$ come coordinate libere, con ovvie notazioni:

$$\mathscr{L}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_N) = \sum_{i=1}^{N} \frac{m_i}{2} \dot{\mathbf{x}}_i^2 - \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N)\,.$$

Le equazioni di Eulero-Lagrange non sono altro che le equazioni di Newton per il sistema di punti materiali considerati nel sistema di riferimento inerziale $\mathscr{I}$. Infatti, se $\mathbf{x}_i = \sum_{k=1}^{3} x_i^k\, \mathbf{e}_k$, allora:

$$\left(\frac{d}{dt}\frac{\partial}{\partial \dot{x}_i^k} - \frac{\partial}{\partial x_i^k}\right)\left(\sum_{i=1}^{N} \frac{m_i}{2} \dot{\mathbf{x}}_i^2 - \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N)\right) = \frac{d}{dt}\frac{\partial \frac{m_i}{2}\dot{\mathbf{x}}_i^2}{\partial \dot{x}_i^k} + \frac{\partial}{\partial x_i^k}\mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$$

$$= m_i \frac{d\dot{x}_i^k}{dt} + \frac{\partial}{\partial x_i^k} \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,,$$

e quindi le equazioni di Eulero-Lagrange, scritte come equazioni del secondo ordine (usando $\dot{x}_i^k = dx_i^k/dt$) diventano quelle di Newton:

$$m_i \frac{d^2 x_i^k}{dt^2} = -\frac{\partial}{\partial x_i^k} \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,, \quad \text{per } i = 1, \ldots, N \text{ e } k = 1, 2, 3.$$

(2) Si consideri un sistema fisico costituito da tre punti materiali P, P_1, P_2 di masse m, m_1, m_2 rispettivamente. P_1 e P_2 sono connessi da un'asta rigida ideale di lunghezza $d > 0$ e priva di peso. Il punto P è connesso all'asta tramite una molla di costante elastica $K > 0$ attaccata al centro di massa G dell'asta (tenuto conto delle masse dei due punti). Si suppone che la molla abbia lunghezza nulla a riposo. Non sono presenti altre forze sui punti del sistema, escluse ovviamente le reazioni vincolari. Vogliamo scrivere le equazioni del moto (quindi equazioni pure di movimento) del sistema fisico, descrivendolo in un riferimento inerziale $\mathscr{I}$.
Fissiamo un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ con origine O e assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Un insieme di coordinate libere possono essere le seguenti 8 coordinate: le tre componenti x^1, x^2, x^3, rispetto agli assi detti, del vettore posizione $P - O$, le tre componenti X^1, X^2, X^3, rispetto agli assi detti, del vettore posizione $G - O$, i due angoli polari θ, ϕ del vettore posizione $P_2 - G$ per un sistema di coordinate cartesiane con origine G ed assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ (gli stessi di $\mathscr{I}$) individuanti il sistema di riferimento del centro di massa $\mathscr{I}_G$.
Dato che l'unica forza attiva è conservativa, dovuta alla molla, le equazioni di Eulero-Lagrange si potranno riferire alla lagrangiana $\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$. Usando il teorema di König (teorema 4.2) per valutare l'energia cinetica rispetto ad $\mathscr{I}$ del sistema dei due punti P_1, P_2 si ha:

$$\mathscr{T}|_{\mathscr{I}} = \frac{1}{2} \sum_{j=1}^{3} m(\dot{x}^j)^2 + \frac{1}{2} \sum_{j=1}^{3} M(\dot{X}^j)^2 + \mathscr{T}|_{\mathscr{I}_G} \,,$$

dove $M := m_1 + m_2$. L'energia cinetica riferita al sistema del centro di massa $\mathscr{I}_G$ dei due punti è semplicemente la somma delle energie cinetiche di ciascun punto valutate in $\mathscr{I}_G$. Indichiamo con d_1 e d_2 le distanze di P_1 e P_2 da G, unicamente determinate dalle relazioni $m_1 d_1 = m_2 d_2$ e $d_1 + d_2 = d$. Consideriamo il punto P_2. Il suo vettore posizione è $P - G = d_1\, \mathbf{e}_r$ dove $\mathbf{e}_r$ è il versore radiale delle coordinate polari centrate in G e riferite agli assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Di conseguenza, usando l'esercizio 2 negli esercizi 1.2:

$$\mathbf{v}_{P_2}|_{\mathscr{I}_g} = d_1\, \dot{\mathbf{e}}_r = d_1(\dot{\theta}\, \mathbf{e}_\theta + \dot{\phi} \sin\theta\, \mathbf{e}_\phi) \,.$$

Dato che $P_1 - G = -\frac{d_1}{d_2}(P_2 - G)$ avremo anche che

$$\mathbf{v}_{P_2}|_{\mathscr{I}_g} = -d_2\, \dot{\mathbf{e}}_r = -d_2(\dot{\theta}\, \mathbf{e}_\theta + \dot{\phi} \sin\theta\, \mathbf{e}_\phi) \,.$$

In definitiva, dato che i versori $\mathbf{e}_\theta$ e $\mathbf{e}_\phi$ sono ortogonali, quadrando le espressioni di sopra troviamo che

$$(\mathbf{v}_{P_2}|_{\mathscr{I}_g})^2 = d_2^2(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta)$$

e

$$(\mathbf{v}_{P_1}|_{\mathscr{I}_g})^2 = d_1^2(\dot\theta^2 + \dot\phi^2 \sin^2\theta)\,.$$

Pertanto:

$$\mathscr{T}|_{\mathscr{I}} = \frac{m}{2}\sum_{j=1}^{3}(\dot x^j)^2 + \frac{M}{2}\sum_{j=1}^{3}(\dot X^j)^2 + \frac{1}{2}(m_1^2 d_1^2 + m_2 d_2^2)(\dot\theta^2 + \dot\phi^2\sin^2\theta)\,.$$

L'energia potenziale della molla (che in questo caso non dipende dal riferimento), vale

$$\mathscr{U} = \frac{K}{2}\sum_{j=1}^{3}(X^j - x^j)^2\,.$$

Concludiamo che la lagrangiana del sistema dei tre punti materiali riferita al sistema di riferimento inerziale $\mathscr{I}$ è :

$$\mathscr{L}|_{\mathscr{I}} = \frac{m}{2}\sum_{j=1}^{3}(\dot x^j)^2 + \frac{M}{2}\sum_{j=1}^{3}(\dot X^j)^2 + \frac{1}{2}(m_1^2 d_1^2 + m_2 d_2^2)(\dot\theta^2 + \dot\phi^2\sin^2\theta) - \frac{K}{2}\sum_{j=1}^{3}(X^j - x^j)^2\,.$$

Abbiamo quindi un sistema di equazioni di Eulero-Lagrange, dato da 8 equazioni che si ottengono da (6.44) per q^k dati rispettivamente da x^k, X^k con $k = 1,2,3$ e θ, ϕ. Le equazioni sono le seguenti.

$$\begin{cases} m\dfrac{d^2x^k}{dt^2} = -K(x^k - X^k)\,, \\ M\dfrac{d^2X^k}{dt^2} = -K(X^k - x^k)\,, \\ \dfrac{d^2\theta}{dt^2} = 2\left(\dfrac{d\phi}{dt}\right)^2 \sin\theta\cos\theta\,, \\ \dfrac{d}{dt}\left(\dfrac{d\phi}{dt}\sin^2\theta\right) = 0\,. \end{cases} \quad \text{per } k = 1,2,3,$$

Il sistema di equazioni è del secondo ordine e scrivibile in forma normale se $\sin\theta \neq 0$, quando ciò accade siamo in realtà fuori dal dominio delle coordinate polari. Dato che il secondo membro del sistema che si ottiene passando a forma normale è di classe C^∞, il sistema di equazioni ammette una sola soluzione, il moto, una volta fissate condizioni iniziali date $x^i(t_0)$, $X^i(t_0)$, $\theta(t_0)$ e $\phi(t_0)$ e le corrispondenti derivate $(dx^i/dt)(t_0)$, $(dX^i/dt)(t_0)$, $(d\theta/dt)(t_0)$ e $(d\phi/dt)(t_0)$.

(3) Studiando sistemi fisici generali $\mathfrak{S}$ composti da due (o più) sottosistemi $\mathfrak{S}_1, \mathfrak{S}_2$, la lagrangiana del sistema (riferita ad un sistema di riferimento $\mathscr{I}$ che sottointendiamo) ha spesso una struttura della forma:

$$\mathscr{L} = \mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_I\,,$$

dove $\mathscr{L}_1$ e $\mathscr{L}_2$ sono le lagrangiane dei due sistemi pensati come non interagenti tra di loro: ognuna di tali lagrangiane contiene *solo* coordinate riferite al corrispondente sottosistema, e $\mathscr{L}_I$

è la *lagrangiana d'interazione* che contiene le coordinate di *entrambi* i sottosistemi. L'esempio più semplice è dato da un sistema composto da due particelle di masse m_1 e m_2 rispettivamente, con coordinate naturali $(\mathbf{x}_1, \dot{\mathbf{x}}_1)$ e $(\mathbf{x}_2, \dot{\mathbf{x}}_2)$ usando le stesse convenzioni che nell'esempio (1) di sopra (rispetto al sistema di riferimento inerziale $\mathscr{I}$). In tal caso le lagrangiane libere sono, per esempio:

$$\mathscr{L}_1 := \frac{m\dot{\mathbf{x}}_1^2}{2}\,, \quad \mathscr{L}_2 := \frac{m\dot{\mathbf{x}}_2^2}{2}\,,$$

mentre una possibile lagrangiana d'interazione è quella data da una forza conservativa associata ad un'energia potenziale che dipende dalle posizioni dei due punti materiali:

$$\mathscr{L}_I := -\mathscr{U}(||\mathbf{x}_1 - \mathbf{x}_2||)\,.$$

Considerando sistemi composti da più sottosistemi la lagrangiana d'interazione può essere spesso decomposta in lagrangiane cosiddette: "a due corpi", "a tre corpi" e via di seguito. Per esempio con 3 particelle, si può avere:

$$\mathscr{L}_I = -\mathscr{U}^{(2)}(\mathbf{x}_1, \mathbf{x}_2) - \mathscr{U}^{(2)}(\mathbf{x}_2, \mathbf{x}_3) - \mathscr{U}^{(2)}(\mathbf{x}_1, \mathbf{x}_3) - \mathscr{U}^{(3)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\,,$$

dove $\mathscr{U}^{(2)}$ rappresenta l'interazione (conservativa in questo caso) a due corpi e $\mathscr{U}^{(3)}$ quella a tre corpi. Nelle applicazioni pratiche, per esempio in meccanica statistica, si usano a volte metodi approssimati in cui in prima approssimazione si trascura la lagrangiana d'interazione o solo una sua parte (es. quella a tre corpi) e se ne tiene conto solo nelle approssimazioni successive.

6.2.4 Cambiamento di riferimento inerziale e non unicità della lagrangiana.

Consideriamo un sistema fisico $\mathfrak{S}$ di N punti materiali con n gradi di libertà, descritto dalle componenti lagrangiane delle forze attive $\mathscr{Q}_k|_{\mathscr{I}}$ riferite ad un sistema inerziale $\mathscr{I}$ ed all'energia cinetica $\mathscr{T}|_{\mathscr{I}}$. Non facciamo alcuna ipotesi sulla natura delle coordinate $q^1, \ldots, q^n$ che possono non essere solidali con alcuno dei riferimenti considerati. Per costruzione l'energia cinetica $\mathscr{T}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots\ \dot{q}^n)$, il potenziale $\mathscr{V}|_{\mathscr{I}} = \mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots\ \dot{q}^n)$ e la lagrangiana $\mathscr{L}|_{\mathscr{I}} = \mathscr{L}|_{\mathscr{I}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots\ \dot{q}^n)$ devono essere pensati *campi scalari* su $j^1(\mathbb{V}^{n+1})$ quando il riferimento $\mathscr{I}$ è fissato, ma si cambiano le coordinate libere. Tenendo fisso $\mathscr{I}$, ma cambiando le coordinate naturali usate all'inizio $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots\ \dot{q}^n)$ e passando a nuove coordinate naturali (ovviamente locali ed adattate alle fibre), la funzione lagrangiana nelle nuove coordinate locali sarà quindi data semplicemente da, assumendo per semplicità $n = 1$, la generalizzazione essendo ovvia):

$$\mathscr{L}|_{\mathscr{I}}(t(t'), q(t', q'),\ \dot{q}(t', q'\ \dot{q}'))\,.$$

Valgono analoghi risultati per gli altri campi scalari menzionati sopra. Vediamo cosa succede. tenendo fisse le coordinate libere, ma cambiando il sistema di riferimento $\mathscr{I}$ e passando a $\mathscr{I}'$, limitandoci per ora al caso in cui $\mathscr{I}$ e $\mathscr{I}'$ siano inerziali. Ci aspettiamo che le nuove equazioni di Eulero-Lagrange siano le stesse ottenute lavorando rispetto a $\mathscr{I}$, dato che non vi è alcun motivo

fisico per preferire uno dei due riferimenti e le coordinate naturali $t, q^1, \ldots, q^n$ non sono state scelte in relazione particolare con alcuno dei due riferimenti. Le componenti lagrangiane delle forze attive rimarranno le stesse come osservato in (1) di Osservazioni 6.3, ma l'energia cinetica riferita a $\mathscr{I}'$, $\mathscr{T}|_{\mathscr{I}'}$, differirà invece da $\mathscr{T}|_{\mathscr{I}}$ che è riferita a $\mathscr{I}$. Tenendo conto che, per ogni punto P_i di $\mathfrak{S}$:

$$\mathbf{v}_{P_i}|_{\mathscr{I}} = \mathbf{v}_{P_i}|_{\mathscr{I}'} + \mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}} \ ,$$

dove $\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}}$ è costante nel tempo e nello spazio (essendo i due riferimenti in moto relativo rettilineo uniforme) si trova immediatamente che:

$$\Delta\mathscr{T} = \mathscr{T}|_{\mathscr{I}} - \mathscr{T}|_{\mathscr{I}'} = \frac{M}{2}(\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}})^2 + \mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}} \cdot \sum_{i=1}^{N} m_i \mathbf{v}_{P_i}|_{\mathscr{I}'} \ ,$$

dove $M := \sum_{i=1}^{N} m_i$. Se $\mathbf{x}'_i = \mathbf{x}'_i(t, q^1, \ldots, q^n)$ è il vettore posizione del punto P_i nel riferimento $\mathscr{I}$ vale:

$$\mathbf{v}_{P_i}|_{\mathscr{I}'} = \sum_{k=1}^{n} \frac{\partial \mathbf{x}_i}{\partial q^k}\frac{dq^k}{dt} + \frac{\partial q^k}{\partial t} \ ,$$

per cui, *sul moto*:

$$\Delta\mathscr{T} = \frac{d}{dt}\left(\frac{M}{2}(\mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}})^2 t + \mathbf{v}_{\mathscr{I}'}|_{\mathscr{I}} \cdot \sum_{i=1}^{N} m_i \mathbf{x}'_i(t, q^1(t), \ldots, q^n(t))\right) \ .$$

Lavorando su $j^1(\mathbb{V}^{n+1})$, non pensando più le $\dot{q}^k$ come le derivate delle q^k, possiamo scrivere che

$$\Delta\mathscr{T} := \sum_{k=1}^{n} \frac{\partial g}{\partial q^k}\dot{q}^k + \frac{\partial g}{\partial t} \ , \tag{6.45}$$

dove $g = g(t, q^1, \ldots, q^n)$ è la funzione su $\mathbb{V}^{n+1}$ definita all'interno delle parentesi tonde nel secondo membro dell'identità precedente. Si verifica immediatamente che, con $\Delta\mathscr{T}$ della forma (6.45):

$$\frac{d}{dt}\frac{\partial \Delta\mathscr{T}}{\partial \dot{q}^k} - \frac{\partial \Delta\mathscr{T}}{\partial q^k} = 0 \ .$$

La conseguenza immediata di questo risultato è che il termine $\Delta\mathscr{T}$ non fornisce alcun contributo alle equazioni di Eulero-Lagrange e pertanto, tenuto conto che le componenti lagrangiane delle forze (vere) attive non dipendono dal riferimento, possiamo concludere che: *le equazioni di Eulero-Lagrange associate al riferimento inerziale $\mathscr{I}$ coincidono con le equazioni di Eulero-Lagrange associate al riferimento inerziale $\mathscr{I}'$.*

Questo risultato ha come implicazione immediata il fatto che, se un sistema fisico ammette lagrangiana, questa non è univocamente determinata. Supponiamo infatti che le componenti lagrangiane delle forze attive siano date da un potenziale $\mathscr{V}|_{\mathscr{I}}$ e pertanto il sistema fisico ammetta lagrangiana, riferita al sistema inerziale $\mathscr{I}$:

$$\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) + \mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n) \ .$$

Se cambiamo riferimento inerziale passando a $\mathscr{I}$, le componenti lagrangiane delle forze attive in $\mathscr{I}'$ saranno ancora date da un potenziale $\mathscr{V}|_{\mathscr{I}'}(t, q^1, \ldots, q^n)$ che coincide con $\mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n)$ data l'invarianza delle componenti lagrangiane delle forze attive al variare del riferimento inerziale. Lasciamo la semplice prova di ciò al lettore. Le uniche differenze tra le due lagrangiane riguarderanno l'energia cinetica, pertanto:

$$\Delta\mathscr{L} = \mathscr{L}|_{\mathscr{I}} - \mathscr{L}|_{\mathscr{I}'} = \Delta\mathscr{T} \,.$$

Come già osservato sopra, nel caso in esame, il termine $\Delta\mathscr{T}$ non fornisce contributo alle equazioni di Eulero-Lagrange che risultano pertanto le stesse. Abbiamo ottenuto che, per un sistema fisico fissato, esistono almeno due lagrangiane che forniscono le stesse equazioni di Eulero-Lagrange. In questo caso le due lagrangiane hanno un preciso significato fisico essendo riferite a due distinti riferimenti inerziali. In ogni caso il risultato è del tutto generale: se aggiungiamo ad una lagrangiana $\mathscr{L}|_{\mathscr{I}}$ una funzione su $j^1(\mathbb{V}^{n+1})$ che ha una struttura del tipo:

$$\Delta\mathscr{L} := \sum_{k=1}^{n} \frac{\partial g}{\partial q^k}\dot{q}^k + \frac{\partial g}{\partial t} \,, \tag{6.46}$$

dove $g = g(t, q^1, \ldots, q^n)$ è una funzione arbitraria assegnata su $\mathbb{V}^{n+1}$, allora le equazioni prodotte da $\mathscr{L}' := \mathscr{L}|_{\mathscr{I}} + \Delta\mathscr{L}$ quando $\mathscr{L}'$ è inserita nelle (6.44), sono le stesse prodotte da $\mathscr{L}_{\mathscr{I}}$. La dimostrazione è la stessa di prima, basata sul risultato di verifica immediata che, con le definizioni date:

$$\frac{d}{dt}\frac{\partial\Delta\mathscr{L}}{\partial\dot{q}^k} - \frac{\partial\Delta\mathscr{L}}{\partial q^k} = 0 \,.$$

Osservazioni 6.7. Si trova scritto molto spesso in letteratura che il secondo membro di (6.46) è una *derivata totale*. In altre parole:

$$\sum_{k=1}^{n} \frac{\partial g}{\partial q^k}\dot{q}^k + \frac{\partial g}{\partial t} = \frac{dg}{dt} \,.$$

Tuttavia questa caratterizzazione, se presa alla lettera, è errata. Infatti *non è vero*, quando si scrive la (6.45), che le $\dot{q}^k$ siano le derivate delle q^k, ciò è vero solo quando è assegnata una soluzione delle equazioni di Eulero-Lagrange.

Esercizi 6.1.

1. Si considerino due punti materiali P e Q, entrambi di massa m, vincolati sulla curva liscia Γ di equazione $x = R\cos\phi$, $y = R\sin\phi$, $z = R\phi$ ($R > 0$ costante) con $\phi \in \mathbb{R}$, dove x, y, z sono coordinate cartesiane ortonormali solidali con un riferimento inerziale $\mathscr{I}$. Si supponga che i punti siano sottoposti, oltre alla reazione vincolare dovuta alla curva Γ, alla forza peso $-mg\,\mathbf{e}_z$ e che siano legati l'un l'altro attraverso una molla ideale di lunghezza nulla a riposo e costante elastica $\kappa > 0$. Infine si supponga che sul punto P agisca una forza di attrito viscoso $-\gamma\mathbf{v}_P|_{\mathscr{I}}$ con $\gamma \geq 0$ costante. Si risolvano i questiti seguenti.

(a) Dimostrare che se $\gamma = 0$ il sistema ammette lagrangiana $\mathscr{L}|_{\mathscr{I}}$ di cui si scriva la forma

esplicita. Si scriva il sistema delle equazioni pure di movimento dei due punti nel caso $\gamma = 0$ e nel caso $\gamma > 0$. Si usino, come coordinate libere, gli angoli ϕ_P e ϕ_Q.
(b) Scrivendo le equazioni del moto nelle nuove variabili libere $\theta := \phi_P - \phi_Q$ e $\tau := \phi_P + \phi_Q$, dimostrare che, se $\gamma = 0$ e $g = 0$, la grandezza:

$$I := \frac{\partial \mathscr{L}|_{\mathscr{I}}}{\partial \dot{\tau}}$$

si conserva sui moti del sistema e spiegare il significato fisico di tale grandezza.

2. Si considerino due punti materiali P_1 e P_2 di massa m vincolati alla superficie conica liscia C di equazione $z = \sqrt{x^2 + y^2}$ in riferimento a coordinate cartesiane ortogonali di origine O solidali con un riferimento inerziale $\mathscr{I}$. I due punti sono connessi da una molla ideale di lunghezza nulla a riposo e costante elastica $\kappa > 0$ e sono sottoposti alla forza di gravità $-mg\,\mathbf{e}_z$. Si risolvano i quesiti seguenti.
(a) Si scrivano le equazioni di Eulero-Lagrange per il sistema dei due punti. Si descrivano i punti tramite coordinate polari piane (r_1, φ_1) e (r_2, φ_2) delle proiezioni di P_1 e P_2 rispettivamente sul piano $z = 0$.
(b) Scrivendo le equazioni del moto nelle nuove variabili libere $\Phi := \varphi_1 - \varphi_2$ e $\Theta := \varphi_1 + \varphi_2$, dimostrare che la grandezza:

$$I := \frac{\partial \mathscr{L}|_{\mathscr{I}}}{\partial \dot{\Theta}}$$

si conserva sui moti del sistema e spiegare il significato fisico di tale grandezza.

3. Si consideri un sistema di riferimento inerziale $\hat{\mathscr{I}}$ e in esso un sistema di coordinate cartesiane ortonormali solidali $\hat{x}, \hat{y}, \hat{z}$ di origine $\hat{O}$. Un secondo sistema di riferimento, non inerziale, $\mathscr{I}$ è individuato da un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ di origine $O \equiv \hat{O}$ ed assi x, y, z. L'asse z coincide con $\hat{z}$ istante per istante mentre x e y ruotano nel piano $z = 0$ in modo tale che $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = \Omega\,\mathbf{e}_z$ con $\Omega > 0$ costante assegnata.
Un punto materiale P di massa $m > 0$ è vincolato a muoversi sulla curva liscia Γ, ferma in $\mathscr{I}$, di equazione $z = \sinh x$. Oltre alla reazione vincolare $\boldsymbol{\phi}$, il punto P è sottoposto alla forza di gravità $-mg\,\mathbf{e}_z$ ed alla forza di una molla ideale (lunghezza nulla a riposo) di costante elastica κ attaccata a P ed al punto Q sull'asse z che si trova sempre alla stessa quota di P. Si risolvano i seguenti quesiti.
(a) Usando come coordinata libera di P la coordinata x del punto P, si scriva l'equazione differenziale del moto di P. Si lavori con la lagrangiana del punto P definita rispetto al riferimento inerziale $\hat{\mathscr{I}}$.
(b) Si dimostri che la grandezza:

$$\mathscr{H}(x, \dot{x}) := \dot{x}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial \dot{x}} - \mathscr{L}|_{\hat{\mathscr{I}}}$$

è un integrale primo del sistema e se ne spieghi il significato fisico.
(c) Si scriva la componente lungo l'asse x della reazione vincolare $\boldsymbol{\phi}$ in funzione di x e $\dot{x}$.

(d) Considerando il moto con condizioni iniziali $x(0) = 0$ e $\dot{x}(0) = v > 0$ assegnata, si esprima, su tale moto, la componente lungo l'asse x della reazione vincolare $\boldsymbol{\phi}$ in funzione della sola variabile x.

4. Nel riferimento $\mathscr{I}$ si definisca un sistema di coordinate cartesiane ortonormali solidali ad $\mathscr{I}$ con origine O ed assi $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. Si consideri poi il seguente sistema costituito da due punti materiali P e Q entrambi di massa $m > 0$: P è connesso ad O tramite un'asta rigida senza massa di lunghezza $L > 0$, Q è connesso a P tramite una seconda asta rigida senza massa e di lunghezza L. Nei punti di giunzione vi sono vincoli ideali che permettono alle aste di ruotare liberamente attorno ad O e P rispettivamente rimanendo nel piano $y = 0$. Una molla ideale (di lunghezza nulla a riposo) di costante elastica $k > 0$ unisce Q ad O. Il riferimento $\mathscr{I}$ ruota attorno all'asse $\mathbf{e}_z$ rispetto ad un riferimento inerziale $\hat{\mathscr{I}}$ con un vettore $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = \Omega\, \mathbf{e}_z$ dove $\Omega > 0$ è una costante.
I due punti materiali sono soggetti, oltre alle reazioni vincolari ed alle forze inerziali, alla forza peso $-mg\, \mathbf{e}_z$.
Descrivendo le configurazioni del sistema tramite i due angoli: ϕ che l'asta $P - O$ forma con la retta x e ψ che l'asta $Q - P$ forma con la retta x' parallela ad x e passante per P (si fissi l'orientazione positiva di entrambi gli angoli in senso antiorario rispetto a $-\mathbf{e}_y$) si risolvano i seguenti quesiti.
(a) Usando la Lagrangiana $\mathscr{L} = \mathscr{L}|_{\hat{\mathscr{I}}}$ del sistema riferita a $\hat{\mathscr{I}}$, si scrivano le equazioni del moto di Eulero-Lagrange per il sistema dei due punti.
(b) Si scrivano esplicitamente tali equazioni in forma normale.
(c) Se $q^1 = \phi$ e $q^2 = \psi$, si dimostri che l'Hamiltoniana

$$\mathscr{H} := \sum_{k=1,2} \dot{q}^k \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k} - \mathscr{L}$$

associata a $\mathscr{L}$ ed al sistema di coordinate libere usate è un integrale primo e se ne discuta il significato fisico.
(d) Dimostrare che le 4 configurazioni $(\phi_\pm, \psi_\pm) = (\pm\pi/2, \pm\pi/2)$ sono *configurazioni di equilibrio* nel riferimento $\mathscr{I}$: l'unica soluzione delle equazioni del moto che ammette separatamente ciascuna delle 4 configurazioni dette come condizione iniziale $(\phi(0), \psi(0))$, unitamente alle condizioni $(d\phi/dt(0), d\psi/dt(0)) = (0, 0)$, è quella di quiete nel riferimento $\mathscr{I}$: $(\phi(t), \psi(t)) = (\phi(0), \psi(0))$ per ogni $t \in \mathbb{R}$.

6.3 *Formulazione geometrico differenziale globale delle equazioni di Eulero-Lagrange.

6.3.1 La struttura di varietà fibrata di $\mathbb{V}^{n+1}$ e di $j^1(\mathbb{V}^{n+1})$.

Chiariamo qui la definizione 6.3. Si può dare un significato preciso a $\mathbb{V}^{n+1}$ e $j^1(\mathbb{V}^{n+1})$ nella teoria delle *varietà fibrate* (vedi sezione A.1.9).

In quest'ottica si assume che $\mathbb{V}^{n+1}$ non sia solo una varietà fogliata, ma anche una *varietà fibrata* (cfr definizione A.11) in cui la *base* è data dallo spazio euclideo $\mathbb{E}^1$ che definisce l'asse del tempo assoluto, le fibre nei vari punti di base t sono gli spazi $\mathbb{Q}_t$ e la *proiezione canonica* è banalmente data da $\Pi : \mathbb{V}^{n+1} \to \mathbb{E}^1$. Si osservi che per definizione di varietà fibrata è richisto che le fibre $\mathbb{Q}_t$ siano tutte diffeomorfe tra di esse (in modo non canonico in generale). Questa richiesta è un pò restrittiva dato che non è difficile immaginare situazioni fisiche in cui lo spazio delle configurazioni vari la sua topologia al variare del tempo. Ogni sistema di coordinate locali $(t, q^1, \ldots, q^n)$ (in cui, lo ricordiamo, la coordinata t misura il tempo assoluto a meno di una costante additiva) è un sistema di coordinate locali *adattate alle fibre* sul fibrato $\mathbb{V}^{n+1}$: t è la coordinata che corre localmente sulla base e che individua il tempo assoluto a meno di costanti additive, mentre $(q^1, \ldots, q^n)$, per ogni fissato t, individuano coordinate locali sulla fibra $\mathbb{Q}_t$.
Una **sezione** del fibrato $\mathbb{V}^{n+1}$ è per definizione una funzione differenziabile $\mathbb{E}^1 \ni u \mapsto \gamma(u) \in \mathbb{V}^{n+1}$ tale che $\Pi(\gamma(u)) = u$. In altre parole una sezione è una curva differenziabile in $\mathbb{V}^{n+1}$ che può essere parametrizzata usando il tempo assoluto t e pertanto con il possibile significato di moto del sistema.
Lo spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$ si costruisce ora come segue tenendo conto del fatto che $\mathbb{V}^{n+1}$ ammette la funzione privilegiata $T : \mathbb{V}^{n+1} \to \mathbb{R}$ data dal tempo assoluto. Per prima cosa si considera la varietà *fibrato tangente* (vedi la sezione A.1.9): $T\mathbb{V}^{n+1}$. La funzione differenziabile tempo assoluto $T : \mathbb{V}^{n+1} \to \mathbb{R}$ si estende naturalmente a $T\mathbb{V}^{n+1}$ in una funzione $\tilde{T}$ tale che

$$\tilde{T} : T\mathbb{V}^{n+1} \ni (p, v_p) \mapsto T(p) \, .$$

Indicheremo d'ora in poi $\tilde{T}$ con la stessa lettera T. Un sistema di coordinate locali naturali $(t, q^1, \ldots, q^n)$ su $\mathbb{V}^{n+1}$ individua in modo canonico un sistema locale di coordinate naturali su $T\mathbb{V}^{n+1}$: $t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n$ dove $\dot{q}^1, \ldots, \dot{q}^n$ sono le componenti sulla base $\partial/\partial t, \partial/\partial q^1, \ldots \partial/\partial q^n$ di un vettore di $T_{(t,q^1,\ldots,q^n)}\mathbb{V}^{n+1}$. Un possibile moto del sistema, cioé una sezione $\mathbb{R} \ni t \mapsto \gamma(t) \in \mathbb{V}^{n+1}$ di classe C^2, parametrizzata nel tempo assoluto, individua a sua volta un'analoga curva su $T\mathbb{V}^{n+1}$ di classe C^1, detta **rialzamento** della precedente, data da: $\mathbb{R} \ni t \mapsto (\gamma(t), \dot{\gamma}(t))$. In coordinate naturali locali di $\mathbb{V}^{n+1}$ e $T\mathbb{V}^{n+1}$ abbiamo allora l'espressione esplicita per le due curve:

$$t \mapsto (t, q^1(t), \ldots, q^n(t)) \, , \quad \text{e} \quad t \mapsto \left(t, q^1(t), \ldots, q^n(t), 1, \frac{dq^1}{dt}, \ldots, \frac{dq^n}{dt}\right) \, .$$

Si osservi che la componente del vettore tangente $\dot{\gamma}$ relativa alla coordinata t vale sempre 1 per costruzione (anche se alteriamo la costante additiva insita nella definizione di tempo assoluto). Ciò deve accadere per ogni moto del sistema. Questa condizione può essere scritta in modo intrinseco usando le notazioni introdotte nell'appendice A, richiedendo che:

$$\langle \gamma(t), dT_{\gamma(t)} \rangle = 1 \, .$$

Se nella varietà $T\mathbb{V}^{n+1}$ ci restringiamo a considerare solo i punti (p, v_p) tali che i vettori v_p soddisfino il requisito $\langle v_p, dT_p \rangle = 1$ otteniamo una sottovarietà embedded indicata con $j_1(\mathbb{V}^{n+1})$. Si osservi che in tale varietà sono contenuti tutte le coppie (p, v_p) in cui p è attraversato da

un possibile moto del sistema con vettore tangente v_p. $j^1(\mathbb{V}^{n+1})$ può essere visto come una varietà fibrata rispetto a due fibrazioni distinte:

(a) base data da $\mathbb{V}^{n+1}$ e fibre date dai sottoinsiemi di ogni spazio tangente $T_p\mathbb{V}^{n+1}$:

$$\left\{(p, v_p) \in T_p\mathbb{V}^{n+1} \mid \langle v_p, dT_p\rangle = 1\right\} .$$

(b) base data dall'asse dei tempi $\mathbb{E}^1$ e fibre date dagli *spazi degli atti di moto* $\mathbb{A}_t$, per ogni tempo $t \in \mathbb{R}$,

$$\mathbb{A}_t = \left\{(c, v_{(t,c)}) \in T_{(t,c)}\mathbb{V}^{n+1} \mid c \in \mathbb{Q}_t\,, \quad \langle v_{(t,c)}, dTv_{(t,c)}\rangle = 1\right\} .$$

Per ogni sistema di coordinate locali naturali $(t, q^1, \ldots, q^n)$ sullo spaziotempo delle configurazioni $\mathbb{V}^{n+1}$, le coordinate locali $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ sono allora coordinate locali sullo spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$ adattate alle fibre di entrambe le fibrazioni in senso ovvio: $(t, q^1, \ldots, q^n)$ sono coordinate sulla base della prima fibrazione e le rimanenti $(\dot{q}^1, \ldots, \dot{q}^n)$ sono coordinate su ciascuna fibra rispetto alla prima fibrazione. Inoltre t è una coordinata sulla base della seconda fibrazione mentre le rimanenti $(q^1, \ldots q^n, \dot{q}^1, \ldots, \dot{q}^n)$ sono coordinate su ciascuna fibra rispetto alla seconda fibrazione. Tali sistemi di coordinate locali adattati alle fibre di $j^1(\mathbb{V}^{n+1})$ formano un atlante per costruzione. La legge di trasformazione tra coordinate di questo genere è proprio espressa dalle (6.31),(6.32), (6.33).

Osservazioni 6.8. La notazione $j^1(\mathbb{V}^{n+1})$ deriva dalla teoria del *jet bundles* della quale non ci occuperemo. $j^1(\mathbb{V}^{n+1})$ è il primo (questo spiega l'indice ^{1}) jet bundle del fibrato $\mathbb{V}^{n+1}$.

6.3.2 Il campo vettoriale dinamico Z.

Il corollario alla proposizione 6.3 prova che le soluzioni alle equazioni di Eulero-Lagrange (6.19) su $j^1(\mathbb{V}^{n+1})$ sono ben definite su tale varietà anche se le equazioni (6.19) sono scritte in forma locale su ogni sistema di coordinate locali naturali adattate alle fibre. Dato che le equazioni di Eulero-Lagrange sono equazioni differenziale del prim'ordine su $j^1(\mathbb{V}^{n+1})$ le soluzioni saranno date dalle curve integrali di un particolare campo vettoriale Z definito sulla varietà $j^1(\mathbb{V}^{n+1})$ come discusso nella sezione 3.5. Tale campo per definizione è esteso a tutto $j^1(\mathbb{V}^{n+1})$. Vogliamo vedere come esso è scritto in coordinate naturali di $j^1(\mathbb{V}^{n+1})$. Se $t \mapsto (t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots, dq^n(t))$ soddisfa le equazioni di Eulero-Lagrange, direttamente dalle (6.19) abbiamo che il suo vettore tangente ha la forma:

$$\frac{\partial}{\partial t} + \sum_{k=1}^{n} \dot{q}^k(t)\frac{\partial}{\partial q^k} + \sum_{k=1}^{n} z^k(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots, dq^n(t))\,\frac{\partial}{\partial\,\dot{q}^k} .$$

L'espressione esplicita delle n funzioni $z^k(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots, dq^n(t))$ si trova scrivendo in forma normale le equazioni di Eulero-Lagrange, come provato nella proposizione 6.2. In definitiva, il vettore tangente alle soluzioni delle equazioni di Eulero-Lagrange ha struttura, in

ogni sistema di coordinate locali naturali,

$$Z(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \frac{\partial}{\partial t} + \sum_{k=1}^{n} \dot{q}^k \frac{\partial}{\partial q^k} + \sum_{k=1}^{n} z^k(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)\, \frac{\partial}{\partial\, \dot{q}^k}\,. \quad (6.47)$$

dove le funzioni z^k definite localmente su $j^1(\mathbb{V}^{n+1})$ si ottengono scrivendo in forma normale le equazioni di Eulero-Lagrange. Indipendentemente dal fatto che abbiamo dato sopra l'espressione si Z in coordinate locali, si deve tener presente che il campo vettoriale Z è ben definito su tutto $j^1(\mathbb{V}^{n+1})$. In quest'ottica le equazioni di Eulero-Lagrange (6.19) devono essere viste come equazioni che determinano il campo vettoriale Z detto **vettore dinamico** le cui curve integrali sono le soluzioni del moto del sistema. Le equazioni di Eulero-Lagrange, in questo contesto si riscrivono in coordinate locali:

$$Z\left(\frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial\, \dot{q}^k}\right) - \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^k} = \mathscr{Q}_k|_{\mathscr{I}}\,, \quad \text{per ogni } k = 1, \ldots, n\,. \quad (6.48)$$

Queste equazioni sono in realtà equazioni *algebriche*, che servono a determinare le componenti z^k (funzioni degli atti di moto su $j^1(\mathbb{V}^{n+1})$) del vettore Z in funzione dei termini noti. Le curve integrali delle equazioni di Eulero-Lagrange, cioé i moti del sistema, sono tutte e sole le soluzioni (massimali) $\gamma : I \to j^1(\mathbb{V}^{n+1})$ del sistema di equazioni differenziali:

$$\frac{d\gamma}{dt} = Z(\gamma(t))\,, \quad (6.49)$$

assumendo, ovviamente il campo Z sufficientemente regolare (C^1 è largamente sufficiente). Anche le (6.49) possono pensarsi come equazioni di Eulero-Lagrange in senso *differenziale*.

Osservazioni 6.9. Bisogna sottolineare che il campo vettoriale dinamico, al contrario della lagrangiana (che oltre a non essere unica può anche non esistere se le componenti lagrangiane delle forze attive non hanno forma opportuna) è univocamente determinato in quanto è associato direttamente alle soluzioni delle equazioni di Eulero-Lagrange.

Capitolo 7

Alcuni argomenti più avanzati di Meccanica Lagrangiana.

In questo capitolo e nel prossimo consideriamo argomenti più avanzati di Meccanica Lagrangiana che hanno avuto delle fondamentali conseguenze nei successivi sviluppi della fisica matematica e teorica, anche in ambiti molto lontani dalla meccanica classica.
In questo capitolo introdurremo in particolare la formulazione variazionale delle equazioni di Eulero-Lagrange, la nozione di potenziale generalizzato e alcune definizioni e risultati di teoria della stabilità .

7.1 Il cosiddetto "Principio di Minima Azione" per sistemi che ammettono lagrangiana.

Vogliamo mostrare come le equazioni di Eulero-Lagrange si possano ottenere da un principio variazionale. In altre parole le soluzioni delle equazioni di Eulero-Lagrange, per sistemi che ammettono lagrangiana, risultano essere *punti di stazionarietà* di un certo funzionale detto *azione*. Questo risultato ha avuto un notevole sviluppo in fisica teorica, quando si passa dalla meccanica classica a quella quantistica, avendo portando alla formulazione della meccanica quantistica (e della teoria dei campi quantistica) basata sul cosiddetto *integrale di Feynman* ed a conseguenti metodi di approssimazione semiclassici.

7.1.1 Primi rudimenti di calcolo delle variazioni.

Partiamo da qualche nozione generale. Se $f : \Omega \to \mathbb{R}$ è una funzione definita sull'aperto $\Omega \subset \mathbb{R}^n$, un punto $p \in \Omega$ è detto *punto di stazionarietà* se la funzione f è differenziabile in p ed il gradiente di f in p è nullo. Ogni punto di minimo, di massimo o di sella di f è un punto di stazionarietà , per cui la condizione di stazionarietà è una condizione necessaria per i punti di estremo.
Vogliamo generalizzare questa nozione al caso di *funzionali*, cioè funzioni che hanno come do-

minio un insieme di funzioni e prendono valori reali. Vediamo qualche esempio di funzionale. Consideriamo un applicazione $F : D \to \mathbb{R}^n$ dove D è un insieme di funzioni definite su un intervallo $I = [a, b]$ a valori in $\mathbb{R}^n$. Diremo tale applicazione **funzionale su** D.

Esempi 7.1. Ecco un esempio abbastanza ovvio di funzionale:

$$F[\gamma] := \gamma(a)\,, \quad \text{per ogni } \gamma \in D\,.$$

Assumendo che le curve in D siano differenziabili, se $c \in (a, b)$ è un punto fissato, un altro semplice esempio è :

$$F[\gamma] := \frac{d\gamma}{dt}|_c\,, \quad \text{per ogni } \gamma \in D\,.$$

Un esempio meno banale, se $n = 1$ e le curve di D sono continue con derivata continua è :

$$F[\gamma] := \int_a^b \left(\gamma(u) + \frac{d\gamma}{du}\right) du\,, \quad \text{per ogni } \gamma \in D\,.$$

Torniamo alla nozione di punto di stazionarietà per funzionali. La definizione basata sull'annullarsi del gradiente è fin troppo complessa in quanto necessita di una nozione di gradiente in uno spazio *infinito dimensionale*. Tale definizione può essere, di fatto, data e si chiama *derivata di Fréchet*. Tuttavia sarebbe necessario introdurre nuove nozioni topologiche che non sono rilevanti nel nostro contesto elementare. Esiste una via più semplice basata sulla nozione di derivata direzionale, e quindi derivata in una sola variabile (tecnicamente, con le debite precisazioni corrisponde alla cosiddetta *derivata debole di Gateaux*). Per arrivare a tale definizione notiamo che un modo equivalente di dire che p è un punto di stazionarietà per f è affermare che tutte le derivate direzionali si annullano:

$$\exists\, \nabla f|_p \cdot \mathbf{u} = 0\,, \quad \text{per ogni vettore } \mathbf{u}.$$

Quest'ultima richiesta può infine essere trascritta, in termini di derivate di funzioni di una variabile, piuttosto che di gradiente di una funzione di più variabile, come:

$$\exists\ \left.\frac{d}{d\alpha}\right|_{\alpha=0} f(p + \alpha\mathbf{u}) = 0\,, \quad \text{per ogni vettore } \mathbf{u}.$$

Tornando al funzionale $F : D \to \mathbb{R}^n$ dove D è un insieme di funzioni definite su un intervallo $I = [a, b]$ a valori in $\mathbb{R}^n$, diremo che $\gamma_0 \in D$ è un punto di stazionarietà per F se:

$$\begin{cases} \exists\ \frac{d}{d\alpha}\big|_{\alpha=0} F[\gamma_0 + \alpha\eta] = 0\,, \\ \text{per ogni } \eta : I \to \mathbb{R}^n \text{ tale che } \gamma_0 + \alpha\eta \in D \text{ per } \alpha \in (-\delta_\eta, \delta_\eta). \end{cases} \tag{7.1}$$

Sopra $\delta_\eta > 0$ in generale dipende da η. Chiaramente se F assume massimo o minimo in γ_0, allora la condizione (7.1) è vera per ogni η, ogni volta che la funzione $\alpha \mapsto F[\gamma_0 + \alpha\eta]$ sia ben definita in un intorno di $\alpha = 0$ e differenziabile in tale punto.
Quello che ora vogliamo mostrare è che *le soluzioni delle equazioni di Eulero-Lagrange sono punti di stazionarietà di un certo funzionale definito su un opportuno spazio di curve.*

7.1.2 Il principio di minima azione.

Consideriamo un sistema fisico $\mathfrak{S}$ di punti materiali descritto su $\mathbb{V}^{n+1}$, fissiamo un sistema di coordinate naturali in $\mathbb{V}^{n+1}$: $t, q^1, \ldots, q^n$ che assumeremo, senza perdere generalità , essere variabili in $I \times \Omega$ dove $I = [a, b]$ con $a < b$ e $\Omega \subset \mathbb{R}^n$ è aperto (non vuoto). Supponiamo che il sistema fisico $\mathfrak{S}$ sia descritto da una lagrangiana, rappresentata, nelle coordinate dette dalla funzione $\mathscr{L} : I \times \Omega \times \cdots \times \Omega \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ di classe C^1 congiuntamente in tutte le sue variabile e definita rispetto ad un riferimento $\mathscr{I}$ che sottointenderemo d'ora in poi. Fissiamo quindi due punti $Q_a, Q_b \in \Omega$ e Consideriamo infine il funzionale, detto **azione**:

$$I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma] := \int_a^b \mathscr{L}\left(t, q^1(t), \ldots, q^n(t), \frac{dq^1}{dt}, \ldots, \frac{dq^n}{dt}\right) dt \,, \tag{7.2}$$

dove le curve $\gamma : I \ni t \mapsto (q^1(t), \ldots q^n(t))$ appartengono al dominio:

$$D_{Q_a,Q_b} := \left\{\gamma \in C^2(I) \;\middle|\; \gamma(t) \in \Omega \text{ per ogni } t \in I \text{ con } \gamma(a) = Q_a \text{ e } \gamma(b) = Q_b\right\} \,.$$

Tenendo conto che Ω è aperto, si prova facilmente che ogni curva $\gamma \in D_{Q_a,Q_b}$ ammette un intorno aperto di $\gamma(I)$ tutto contenuto in Ω; usando tale intorno è possibile definire curve $\eta : I \to \Omega$ tali che $\gamma + \alpha\eta \in D$ per $\alpha \in (-\delta_\eta, \delta_\eta)$ con $\delta_\eta > 0$ sufficientemente piccolo. Pertanto ha senso applicare la definizione di punto di stazionarietà per il funzionale azione producendo il seguente notevole teorema che cade sotto il nome, tecnicamente erroneo, di *principio di minima azione.*

Teorema 7.1. (**Principio di minima azione**.) *Con le definizioni date sopra, una curva $\gamma : I \to q^1(t), \ldots, q^n(t)$ che appartiene a D_{Q_a,Q_b} (e quindi vale $\gamma(a) = Q_a$, $\gamma(b) = Q_b$) soddisfa le equazioni di Eulero-Lagrange (6.44) rispetto a $\mathscr{L} : I \times \Omega \times \cdots \times \Omega \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ di classe C^1 congiuntamente in tutte le sue variabile, se e solo se γ è un punto di stazionarietà del funzionale azione $I^{(\mathscr{L})}_{Q_a,Q_b}$ (7.2).* $\diamondsuit$

Dimostrazione. Per costruzione e con ovvie notazioni:

$$I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma + \alpha\eta] := \int_a^b \mathscr{L}\left(t, \gamma(t) + \alpha\eta(t), \frac{d\gamma_0}{dt} + \alpha\frac{d\eta}{dt}\right) dt \,.$$

Otteniamo allora, con ovvie notazioni, e passando la derivata in α sotto il segno di integrale[1], che:

$$\frac{d}{d\alpha} I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma + \alpha\eta] = \int_a^b \left\{\eta \cdot \nabla_\gamma \mathscr{L}\left(t, \gamma + \alpha\eta, \dot\gamma + \alpha\dot\eta\right) + \dot\eta \cdot \nabla_{\dot\gamma} \mathscr{L}\left(t, \gamma + \alpha\eta, \dot\gamma + \alpha\dot\eta\right)\right\} dt \,.$$

[1] Si fissino γ e η. La derivata in α della funzione integranda del funzionale $I^{(\mathscr{L})}_{Q_a,Q_b}$ è limitata in un intorno J di ogni punto fissato α_0, essendo continua in α e t, e variando t in un compatto. Di conseguenza esiste una costante K positiva che maggiora, uniformemente in α in J, il modulo della la derivata in α dell'integrando. Il teorema di Lagrange assicura quindi che i rapporti incrementali in α della funzione integranda siano similmente maggiorati. Infine il teorema della convergenza dominata di Lebesgue (notando che la funzione costante K è integrabile su I essendo I di misura finita in quanto compatto) assicura la possibilità dello scambio tra simbolo di derivata in α e quello di integrale in t.

Per $\alpha = 0$ e $\gamma = \gamma_0$ troviamo:

$$\frac{d}{d\alpha}\bigg|_{\alpha=0} I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma_0 + \alpha\eta] = \int_a^b \left\{\eta \cdot \nabla_\gamma \mathscr{L}\left(t,\gamma,\dot\gamma\right) + \dot\eta \cdot \nabla_{\dot\gamma}\mathscr{L}\left(t,\gamma,\dot\gamma\right)\right\}\big|_{(t,\gamma_0,\dot\gamma_0)}\, dt\,.$$

Ricordando che $\dot\eta = d\eta/dt$ ed integrando per parti, abbiamo ancora che:

$$\frac{d}{d\alpha}\bigg|_{\alpha=0} I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma_0 + \alpha\eta] = \int_a^b \eta \cdot \left\{\nabla_\gamma \mathscr{L}\left(t,\gamma,\dot\gamma\right) - \frac{d}{dt}\nabla_{\dot\gamma}\mathscr{L}\left(t,\gamma,\dot\gamma\right)\right\}\bigg|_{(t,\gamma_0,\dot\gamma_0)} dt\,.$$

Abbiamo omesso un termine che deriva dell'integrazione per parti ma non fornisce contributo. Infatti tale termine sarebbe:

$$\int_a^b \frac{d}{dt}\left(\eta \cdot \nabla_\gamma \mathscr{L}\right)\bigg|_{(t,\gamma_0,\dot\gamma_0)} dt = \eta(b)\cdot\nabla_\gamma\mathscr{L}|_{(b,\gamma_0,\dot\gamma_0)} - \eta(a)\cdot\nabla_\gamma\mathscr{L}|_{(a,\gamma_0,\dot\gamma_0)}\,. \tag{7.3}$$

Dato che stiamo considerando variazioni $\gamma_0 + \alpha\eta$ della curva γ_0 tenendone fissi gli estremi, dovrà valere $\gamma_0(a) + \alpha\eta(a) = Q_a$ e $\gamma_0(b) + \alpha\eta(b) = Q_b$ per ogni α intorno a 0. Concludiamo che deve essere $\eta(a) = \eta(b) = 0$ da cui l'annullarsi del primo membro di (7.3)
Abbiamo ottenuto che, tornando il componenti

$$\frac{d}{d\alpha}\bigg|_{\alpha=0} I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma_0 + \alpha\eta] = \int_a^b \sum_{k=1}^n \eta^k \left[\frac{\partial\mathscr{L}}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial\mathscr{L}}{\partial\dot q^k}\right)\right]\bigg|_{(t,\gamma_0,\dot\gamma_0)} dt\,, \tag{7.4}$$

dove è sottointeso che $\dot q^k = dq^k/dt$. L'identità ottenuta mostra immediatamente che se $\gamma_0 : t \mapsto (q^1(t),\ldots,q^n(t))$ risolve le equazioni di Eulero-Lagrange allora γ_0 stesso è un punto di stazionarietà di $I^{(\mathscr{L})}_{Q_a,Q_b}$ dato che soddisfa, per (7.4):

$$\frac{d}{d\alpha}\bigg|_{\alpha=0} I^{(\mathscr{L})}_{Q_a,Q_b}[\gamma_0 + \alpha\eta] = 0\,, \tag{7.5}$$

per ogni $\eta : I \to \mathbb{R}^n$ tale che $\gamma_0 + \alpha\eta \in D_{Q_a,Q_b}$ per α variabile in qualche intervallo $(-\delta_\eta, \delta_\eta)$.
Supponiamo viceversa che valga la (7.5) per ogni $\eta : I \to \mathbb{R}^n$ suddetto. In virtù di (7.4) deve essere:

$$\int_a^b \sum_{k=1}^n \eta^k \left[\frac{\partial\mathscr{L}}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial\mathscr{L}}{\partial\dot q^k}\right)\right]\bigg|_{(t,\gamma_0,\dot\gamma_0)} dt = 0\,, \tag{7.6}$$

per ogni $\eta : I \to \mathbb{R}^n$ tale che $\gamma_0 + \alpha\eta \in D_{Q_a,Q_b}$ per α variabile in qualche intervallo $(-\delta_\eta, \delta_\eta)$.
Mostriamo che tale risultato implica in realtà , data l'arbitrarietà delle funzioni η^k, che γ_0 soddisfi le equazioni di Eulero-Lagrange. Supponiamo, per assurdo, che esista un punto $t_0 \in I$ per cui, per un fissato k, *non siano valide le equazioni di Eulero-Lagrange.* In altre parole:

$$\left[\frac{\partial\mathscr{L}}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial\mathscr{L}}{\partial\dot q^k}\right)\right]\bigg|_{(t_0,\gamma_0,\dot\gamma_0)} \neq 0$$

Senza perdere generalità possiamo assumere $k = 1$ e $t \in (a, b)$ (per continuità si ottengono, dal ragionamento che segue, i casi $t = a$ e $t = b$). Dato che la funzione

$$I \ni t \mapsto \left[\frac{\partial \mathscr{L}}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right)\right]\Bigg|_{(t_0,\gamma_0,\dot{\gamma}_0)}$$

è sicuramente continua in t, assumerà un segno fissato in un intorno aperto $I' = (c, d)$ di t_0. Supponiamo tale segno positivo (se fosse negativo, il ragionamento sarebbe analogo). Stringendo un poco I', possiamo fare in modo che tale funzione sia positiva in $[c, d]$. Quindi abbiamo che

$$\left[\frac{\partial \mathscr{L}}{\partial q^1} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^1}\right)\right]\Bigg|_{(t,\gamma_0,\dot{\gamma}_0)} \geq C > 0$$

per ogni $t \in [c, d]$ e per qualche costante C. Per concludere possiamo scegliere tutte le funzioni η^k nulle eccetto η^1 che sia nulla ovunque in I eccetto che in $[c, d]$ in cui è non negativa con integrale pari a qualche $L > 0$. Per questa scelta abbiamo che

$$\int_a^b \sum_{k=1}^n \eta^k \left[\frac{\partial \mathscr{L}}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right)\right]\Bigg|_{(t,\gamma_0,\dot{\gamma}_0)} dt = \int_c^d \eta^1 \left[\frac{\partial \mathscr{L}}{\partial q^1} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^1}\right)\right]\Bigg|_{(t,\gamma_0,\dot{\gamma}_0)} dt \geq C \int_c^d \eta^1 dt \geq LC > 0,$$

che contraddice (7.6). Concludiamo che se γ_0 è un punto di stazionarietà di I_{Q_a,Q_b} allora deve essere vero che, per ogni $t \in I$:

$$\left[\frac{\partial \mathscr{L}}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right)\right]\Bigg|_{(t_0,\gamma_0,\dot{\gamma}_0)} = 0$$

e pertanto γ_0 è anche soluzione delle equazioni di Eulero-Lagrange. □

Osservazioni 7.1.
(**1**) Bisogna sottolineare che *non è detto che esista una soluzione delle equazioni di Eulero-Lagrange* $\gamma : I \to q^1(t), \ldots, q^n(t)$ *quando sono assegnati i punti estremi* $\gamma(a) = Q_a$, $\gamma(b) = Q_b$. E, se esiste, *non è detto che sia unica.* Il teorema di esistenza garantisce infatti l'esistenza di una soluzione *quando sono assegnate condizioni iniziali (di Cauchy)* e non *condizioni al contorno*, cioé i punti estremi per i quali passa la curva cercata.
In questo senso il principio variazionale per le equazioni di Eulero-Lagrange dovrebbe essere visto come una procedura per determinare le equazioni del moto più che le soluzioni di esse.
Si può in realtà provare [Arnold92] che il problema con condizioni al contorno assegnate ammette sempre soluzione quando i punti estremi sono sufficientemente vicini.
(**2**) Ci si può chiedere se le soluzioni delle equazioni di Eulero-Lagrange siano in realtà massimi, minimi o punti di sella del funzionale azione. Si riesce a dimostrare [Arnold92] che per punti estremi sufficientemente vicini, tali soluzioni individuino minimi dell'azione, ma nel caso generale questo risultato non è valido.

(3) Se consideriamo du sistemi di riferimento inerziali $\mathscr{I}$ e $\mathscr{I}'$, le lagrangiane associate ad uno stesso sistema fisico differiscono per un termine di "derivata totale" (vedi la Sezione 6.2.4)

$$\Delta\mathscr{L} := \sum_{k=1}^{n} \frac{\partial g}{\partial q^k}\dot{q}^k + \frac{\partial g}{\partial t} \,. \tag{7.7}$$

Passando all'approccio variazionale nel quale si considerano curve nelle variabili q^k e le $\dot{q}^k$ sono interpretate come le derivate delle q^k, il termine detto è realmente una derivata totale nel tempo. Ogni termine di questo tipo fornisce un contributo *costante* al funzionale azione (7.2), dato che:

$$\int_a^b \Delta\mathscr{L} dt = \int_a^b \left(\sum_{k=1}^{n} \frac{\partial g}{\partial q^k}\frac{dq^k}{dt} + \frac{\partial g}{\partial t} \right) dt = g(b, q^1(b), \ldots, q^n(b)) - g(a, q^1(a), \ldots, q^n(a)) \,.$$

Al variare della curva $\gamma : [a, b] \ni t \mapsto (q^1(1), \ldots, q^n(t))$ nel dominio D_{Q_a,Q_b}, il termine

$$g(b, q^1(b), \ldots, q^n(b)) - g(a, q^1(a), \ldots, q^n(a))$$

è costante per costruzione, dato che tutte le curve del dominio detto hanno gli stessi estremi: $(q^1(a), \ldots, q^n(a))$ e $(q^1(b), \ldots, q^n(b))$ per definizione. Risulta allora evidente che i due funzionali azione, quello associato al riferimento inerziale $\mathscr{I}$ e quello associato al riferimento $\mathscr{I}'$, differendo per una costante, hanno gli stessi punti di stazionarietà . In altre parole determinano le stesse equazioni di Eulero-Lagrange, come abbiamo già osservato nella Sezione 6.2.4 indipendentemente dall'approccio variazionale.

7.2 I potenziali generalizzati.

In meccanica analitica nasce naturalmente un'estensione del concetto di potenziale che include potenziali che dipendono dalle "coordinate puntate" usati per calcolare le componenti lagrangiane di una forza.

7.2.1 Il caso della forza di Lorentz.

Prima di dare le debite definizioni consideriamo un caso di grande rilevanza fisica. Consideriamo una particella P di massa m, non sottoposta a vincoli, dotata di carica elettrica e e sottoposta, nel riferimento inerziale $\mathscr{I}$, ai campi elettrico $\mathbf{E}(t, \mathbf{x})$ e magnetico $\mathbf{B}(t, \mathbf{x})$ assegnati. Nel seguito $\mathbf{x} := P - O$ denota il vettore posizione della particella nel riferimento $\mathscr{I}$ rispetto all'origine O e le componenti cartesiane ortonormali solidali con $\mathscr{I}$ di $\mathbf{x}$ saranno usate come coordinate libere della particella. $\mathbf{v}$ denoterà la velocità della particella rispetto a $\mathscr{I}$ Sulla particella agisce al forza di Lorentz

$$\mathbf{F}_L(t, \mathbf{x}, \mathbf{v}) := e\mathbf{E}(t, \mathbf{x}) + \frac{e}{c}\mathbf{v} \wedge \mathbf{B}(t, \mathbf{x}) \,. \tag{7.8}$$

c è la velocità della luce. Per descrivere i campi elettrico e magnetico introduciamo il *potenziale scalare* $\varphi(t,\mathbf{x})$ ed il *potenziale vettore* $\mathbf{A}(t,\mathbf{x})$ in modo tale che, se $\nabla_{\mathbf{x}}$ indica il gradiente valutato in $\mathbf{x}$:

$$\mathbf{E}(t,\mathbf{x}) = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}(t,\mathbf{x}) - \nabla_{\mathbf{x}}\varphi(t,\mathbf{x}) \,, \tag{7.9}$$

$$\mathbf{B}(t,\mathbf{x}) = \nabla_{\mathbf{x}} \wedge \mathbf{A}(t,\mathbf{x}) \,. \tag{7.10}$$

In termini dei due potenziali, l'espressione (7.8) per la forza di Lorentz si trascrive immediatamente:

$$\mathbf{F}_L(t,\mathbf{x},\mathbf{v}) = -\frac{e}{c}\frac{\partial \mathbf{A}}{\partial t}(t,\mathbf{x}) - e\nabla_x\varphi(t,\mathbf{x}) + \frac{e}{c}\mathbf{v} \wedge (\nabla_{\mathbf{x}} \wedge \mathbf{A}(t,\mathbf{x})) \,. \tag{7.11}$$

Eseguendo l'ultimo prodotto vettoriale possiamo infine scrivere:

$$\mathbf{F}_L(t,\mathbf{x},\mathbf{v}) = -\frac{e}{c}\frac{\partial \mathbf{A}}{\partial t}(t,\mathbf{x}) - e\nabla_x\varphi(t,\mathbf{x}) - \frac{e}{c}\left(\mathbf{v}\cdot\nabla_{\mathbf{x}}\right)\mathbf{A}(t,\mathbf{x}) + \frac{e}{c}\nabla_{\mathbf{x}}\left(\mathbf{v}\cdot\mathbf{A}(t,\mathbf{x})\right) \,. \tag{7.12}$$

Vogliamo ora provare che l'espressione ottenuta (7.12) per la forza di Lorentz si può scrivere come, indicando $\mathbf{v}$ con la terna di coordinate puntate $\dot{\mathbf{x}}$,

$$\mathbf{F}_L(t,\mathbf{x},\mathbf{v}) = \left(\nabla_{\mathbf{x}} - \frac{d}{dt}\nabla_{\dot{\mathbf{x}}}\right)\mathscr{V}(t,\mathbf{x},\dot{\mathbf{x}}) \tag{7.13}$$

dove abbiamo introdotto il potenziale elettromagnetico generalizzato:

$$\mathscr{V}(t,\mathbf{x},\dot{\mathbf{x}}) := -e\varphi(t,\mathbf{x}) + \frac{e}{c}\mathbf{A}(t,\mathbf{x})\cdot\dot{\mathbf{x}} \,. \tag{7.14}$$

In effetti il calcolo diretto mostra che:

$$\begin{aligned}\left(\nabla_{\mathbf{x}} - \frac{d}{dt}\nabla_{\dot{\mathbf{x}}}\right)\mathscr{V}(t,\mathbf{x},\dot{\mathbf{x}}) &= -e\nabla_x\varphi(t,\mathbf{x}) + \frac{e}{c}\nabla_{\mathbf{x}}\left(\dot{\mathbf{x}}\cdot\mathbf{A}(t,\mathbf{x})\right) - \frac{d}{dt}\left(\frac{e}{c}\mathbf{A}(t,\mathbf{x})\right) \\ &= -e\nabla_x\varphi(t,\mathbf{x}) + \frac{e}{c}\nabla_{\mathbf{x}}\left(\dot{\mathbf{x}}\cdot\mathbf{A}(t,\mathbf{x})\right) - \frac{e}{c}\frac{\partial \mathbf{A}}{\partial t}(t,\mathbf{x}) - \frac{e}{c}\dot{\mathbf{x}}\cdot\nabla_{\mathbf{x}}\mathbf{A}(t,\mathbf{x}) \\ &= -\frac{e}{c}\frac{\partial \mathbf{A}}{\partial t}(t,\mathbf{x}) - e\nabla_x\varphi(t,\mathbf{x}) - \frac{e}{c}\left(\mathbf{v}\cdot\nabla_{\mathbf{x}}\right)\mathbf{A}(t,\mathbf{x}) + \frac{e}{c}\nabla_{\mathbf{x}}\left(\mathbf{v}\cdot\mathbf{A}(t,\mathbf{x})\right) = \mathbf{F}_L(t,\mathbf{x},\mathbf{v}) \,.\end{aligned}$$

Questo risultato implica immediatamente che le equazioni del moto della carica si possono scrivere come equazioni di Eulero-Lagrange purché si usi la lagrangiana

$$\mathscr{L}(t,\mathbf{x},\dot{\mathbf{x}}) = \frac{m}{2}\dot{\mathbf{x}}^2 + \mathscr{V}(t,\mathbf{x},\dot{\mathbf{x}}) \tag{7.15}$$

dove la funzione $\mathscr{V} : j^1(\mathbb{V}^{3+1}) \to \mathbb{R}$ è stata definita in (7.14). Infatti, le equazioni di Eulero-Lagrange si scrivono in questo caso:

$$\left(\frac{d}{dt}\nabla_{\dot{\mathbf{x}}} - \nabla_{\mathbf{x}}\right)\left(\frac{m}{2}\dot{\mathbf{x}}^2 + \mathscr{V}(t,\mathbf{x},\dot{\mathbf{x}})\right) = 0 \,, \quad \dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} \,,$$

che possiamo riscrivere come:

$$\left(\frac{d}{dt}\nabla_{\dot{\mathbf{x}}} - \nabla_{\mathbf{x}}\right)\frac{m}{2}\dot{\mathbf{x}}^2 = \left(\nabla_{\mathbf{x}} - \frac{d}{dt}\nabla_{\dot{\mathbf{x}}}\right)\mathscr{V}(t,\mathbf{x},\dot{\mathbf{x}})\,, \quad \dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}\,,$$

e quindi, tenendo conto della (7.13) ed eseguendo esplicitamente il calcolo a primo membro della prima uguaglianza:

$$m\frac{d\dot{\mathbf{x}}}{dt} = \mathbf{F}_L(t,\mathbf{x},\dot{\mathbf{x}})\,, \quad \dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}\,,$$

da cui finalmente:

$$m\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F}_L(t,\mathbf{x},\dot{\mathbf{x}})\,.$$

Concludiamo che la lagrangiana per una particella con carica e sottoposta ad un campo elettromagnetico assegnato nel riferimento inerziale $\mathscr{I}$, in termini del potenziale elettrico φ e di quello magnetico vettore $\mathbf{A}$, è datata da:

$$\mathscr{L}(t,\mathbf{x},\dot{\mathbf{x}}) = \frac{m}{2}\dot{\mathbf{x}}^2 - e\varphi(t,\mathbf{x}) + \frac{e}{c}\mathbf{A}(t,\mathbf{x})\cdot\dot{\mathbf{x}}\,. \tag{7.16}$$

Osservazioni 7.2. I campi $\mathbf{E}$ e $\mathbf{B}$ dati in (7.9) e (7.10) non sono alterati se cambiamo i potenziali attraverso una *trasformazione di gauge*:

$$\varphi(t,\mathbf{x}) \to \varphi'(t,\mathbf{x}) := \varphi(t,\mathbf{x}) - \frac{1}{c}\frac{\partial\chi}{\partial t}(t,\mathbf{x})\,, \qquad \mathbf{A}(t,\mathbf{x}) \to \mathbf{A}'(t,\mathbf{x}) := \mathbf{A}(t,\mathbf{x}) + \nabla_{\mathbf{x}}\chi(t,\mathbf{x})\,, \tag{7.17}$$

dove χ è una qualsiasi funzione di classe C^2. Sotto trasformazioni di gauge la lagrangiana (7.16) *non* è invariante. Tuttavia, il calcolo esplicito mostra immediatamente che sotto una trasformazione di gauge risulta:

$$\mathscr{L}(t,\mathbf{x},\dot{\mathbf{x}}) \to \mathscr{L}'(t,\mathbf{x},\dot{\mathbf{x}}) := \mathscr{L}(t,\mathbf{x},\dot{\mathbf{x}}) + \frac{d}{dt}\frac{e}{c}\chi(t,\mathbf{x})\,,$$

pertanto le equazioni del moto risultano essere inalterate dato che la nuova lagrangiana differisce dalla vecchia per mezzo di una sola "derivata totale".

7.2.2 Generalizzazione della nozione di potenziale.

L'esempio precedente mostra che in alcuni casi le componenti lagrangiane $\mathscr{Q}_k$ di una forza si possono ottenere applicando l'operatore di Eulero-Lagrange $\frac{\partial}{\partial q^k} - \frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}$ su una funzione $\mathscr{V} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ che dipende sia dalle coordinate q^k (e dal tempo) e anche *dalle coordinate* $\dot{q}^k$. Il caso di componenti lagrangiane che si ottengono da un potenziale (funzione dt t e delle q^k solamente) è un semplice sottocaso. Osserviamo che, nel caso della forza di Lorentz studiato sopra, la funzione $\mathscr{V}$ dipendeva *linearmente* dalle coordinate puntate. In questo modo le componenti lagrangiane delle forze associate a $\mathscr{V}$ risultavano dipendere dalle *derivate prime nel tempo* delle coordinate q^k (cioé dalle coordinate $\dot{q}^k$) e non dalle *derivate seconde nel tempo* delle $\dot{q}^k$ (cioé dalle derivate prime delle coordinate $\dot{q}^k$). La dipendenza dalle derivate seconde nel tempo

delle coordinate q^k è interamente dovuta, nelle equazioni di Eulero-Lagrange, al termine associato all'energia cinetica del sistema. Sappiamo che in virtù di questa dipendenza le equazioni di Eulero-Lagrange possono scriversi in forma normale come dimostrato nella proposizione 6.2. Questo è un punto molto importante, sia dal punto di vista fisico che matematico, sul quale ci si deve soffermare. Assumendo di avere componenti lagrangiane delle forze $\mathscr{Q}_k$ che si ottengono applicando l'operatore di Eulero-Lagrange $\frac{\partial}{\partial q^k} - \frac{d}{dt}\frac{\partial}{\partial \dot{q}^k}$ su una funzione $\mathscr{V} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ che dipende sia dalle coordinate q^k (e dal tempo) e anche *dalle coordinate* $\dot{q}^k$, è ragionevole limitare la possibile dipendenza di $\mathscr{V}$ dalle $\dot{q}^k$ al caso *lineare*. Se la dipendenza di $\mathscr{V}$ dalle $\dot{q}^k$ fosse di tipo arbitrario, potremmo avere componenti lagrangiane delle forze, tenendo conto della richiesta $dq^k = dq^k/dt$, che contengono le *derivate temporali seconde* delle coordinate q^k. In questo caso saremmo fuori dalle ipotesi della proposizione 6.2 e non sarebbe più assicurata la possibilità di trascrivere in forma normale le equazioni di Eulero-Lagrange perdendo, in generale, l'esistenza e l'unicità della soluzione, cioé il *determinismo* dal punto di vista fisico. Tenuto conto di ciò diamo la seguente definizione.

Definizione 7.1. **(Potenziale generalizzato)** Si consideri un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$ sottoposto a vincoli olonomi ideali e descritto su $j^1(\mathbb{V}^{n+1})$. Si consideri quindi una distribuzione di forze $\mathbf{F}_i|_{\mathscr{I}}$, $i = 1, \ldots, N$, con $\mathbf{F}_i|_{\mathscr{I}}$ agente su P_i (la dipendenza dal riferimento $\mathscr{I}$ sussiste solamente nel caso in cui le forze siano inerziali). Tale distribuzione di forze ammette $\mathscr{V} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ come **potenziale generalizzato**, se, in ogni sistema di coordinate locali naturali adattate alle fibre di $j^1(\mathbb{V}^{n+1})$, le componenti lagrangiane $\mathscr{Q}_k|_{\mathscr{I}}$ sono ottenibili come:

$$\mathscr{Q}_k|_{\mathscr{I}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \frac{\partial \mathscr{V}|_{\mathscr{I}}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}|_{\mathscr{I}}}{\partial\, \dot{q}^k}\,, \tag{7.18}$$

dove la funzione $\mathscr{V}|_{\mathscr{I}}$ dipende al più linearmente dalle coordinate $\dot{q}^k$, e l'operatore d/dt è definito come

$$\frac{d}{dt} f(t, q^1, \ldots, q^n) := \frac{\partial f}{\partial t} + \sum_{k=1}^{n} \dot{q}^k \frac{\partial f}{\partial q^k}$$

per ogni funzione f delle coordinate locali $(t, q^1, \ldots, q^n)$. $\diamondsuit$

Osservazioni 7.3.
(1) Dato che nella (7.18) non è supposto alcun legame tra le coordinate q^k e le coordinate $\dot{q}^k$, in particolare non è fissata alcuna curva parametrizzata in t, la definizione di d/dt è necessaria. Tale definizione risulta essere quella standard quando l'espressione (7.18) è usata nelle equazioni di Eulero-Lagrange, in cui si assume esplicitamente $\dot{q}^k = dq^k/dt$.
(2) Se si assume che $\mathscr{V}|_{\mathscr{I}}$ sia un campo scalare su $j^1(\mathbb{V}^{n+1})$, si dimostra con al stessa procedura usata per provare la proposizione 6.3 si verifica che, cambiando coordinate naturali, valgono le relazioni:

$$\frac{\partial \mathscr{V}|_{\mathscr{I}}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}|_{\mathscr{I}}}{\partial\, \dot{q}^k} = \sum_{j=1}^{n} \frac{\partial q'^j}{\partial q^k}\left(\frac{\partial \mathscr{V}|_{\mathscr{I}}}{\partial q'^j} - \frac{d}{dt}\frac{\partial \mathscr{V}|_{\mathscr{I}}}{\partial\, \dot{q}'^j}\right)\,.$$

Queste relazioni sono compatibili con le relazioni valide per le componenti lagrangiane sotto cambiamento di coordinate locali naturali adattate alle fibre:

$$\mathscr{Q}_k|_{\mathscr{I}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \sum_{j=1}^{n} \frac{\partial q'^j}{\partial q^k} \mathscr{Q}'_j|_{\mathscr{I}}(t', q'^1, \ldots, q'^n,\ \dot{q}'^1, \ldots,\ \dot{q}'^n)\,.$$

(3) Tenuto conto della linearità rispetto alle coordinate puntate, in un sistema di coordinate naturali fissate, l'espressione esplicita di $\mathscr{V}$ può solo essere della forma

$$\mathscr{V}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \sum_{k=1}^{n} A_h(t, q^1, \ldots, q^n)\, \dot{q}^h + B(t, q^1, \ldots, q^n)\,. \tag{7.19}$$

Si verifica subito che cambiando coordinate naturali la struttura di sopra è effettivamente preservata e rimane lineare nelle coordinate puntate. In ogni sistema di coordinate locali naturali adattate alle fibre di $j^1(\mathbb{V}^{n+1})$ il calcolo esplicito di $\mathscr{Q}_k|_{\mathscr{I}}$ tenendo conto di (7.19) e (7.18) fornisce immediatamente:

$$\mathscr{Q}_k|_{\mathscr{I}} = \sum_{h=1}^{n} \left(\frac{\partial A_h}{\partial q^k} - \frac{\partial A_k}{\partial q^h} \right) \dot{q}^h - \frac{\partial A_k}{\partial t} + \frac{\partial B}{\partial q^k}\,. \tag{7.20}$$

Risulta allora evidente che anche le componenti lagrangiane delle forze, e non solo il potenziale generalizzato che le genera, devono essere funzioni *lineari* delle coordinate puntate $\dot{q}^k$.

7.2.3 Condizioni per l'esistenza del potenziale generalizzato.

Se è assegnata la forma funzionale di componenti lagrangiane di sistemi di forze: $\mathscr{Q}_k|_{\mathscr{I}} = \mathscr{V}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ non è un compito banale stabilire se esse possono essere generate da un potenziale generalizzato. Esistono però delle condizioni necessarie e localmente sufficienti che ora studieremo. La situazione è simile a quella in cui ci si chiede se una determinata forza sia ottenibile come il gradiente di un campo scalare. (In quel caso condizioni necessarie e localmente sufficienti, se il campo di forze è posizionale e di classe C^2, sono quelle di irrotazionalità (4.12) come sancito nel teorema 4.4). Nel seguito ometteremo la specificazione $|_{\mathscr{I}}$ per semplicità notazionale.
Cerchiamo dunque delle identità che coinvolgano solo le $\mathscr{Q}_k$, ma non $\mathscr{V}$ e nemmeno le funzioni A_k e B, e che valgano ogni qual volta le componenti lagrangiane $\mathscr{Q}_k$ ammetto potenziale generalizzato per cui è verificata la (7.19).
Da (7.20) abbiamo immediatamente che:

$$\frac{\partial \mathscr{Q}_k}{\partial\, \dot{q}^h} = \frac{\partial A_h}{\partial q^k} - \frac{\partial A_k}{\partial q^h}$$

e di conseguenza deve valere l'identità :

$$\frac{\partial \mathscr{Q}_k}{\partial\, \dot{q}^h} + \frac{\partial \mathscr{Q}_h}{\partial\, \dot{q}^k} = 0\,, \quad \text{ovunque e per ogni } k, h = 1, \ldots, n.$$

Nello stesso modo, partendo da (7.20) si verifica facilmente, confrontando i due membri, che:

$$\frac{\partial \mathscr{Q}_k}{\partial q^h} - \frac{\partial \mathscr{Q}_h}{\partial q^k} = \frac{d}{dt}\left(\frac{\partial \mathscr{Q}_h}{\partial \dot{q}^k}\right), \quad \text{ovunque e per ogni } k, h = 1, \ldots, n.$$

Vale il seguente notevole risultato.

Teorema 7.2. *In riferimento alla definizione 7.1 valgono i seguenti fatti.*
(a) *Se un sistema di forze ammette potenziale generalizzato allora componenti lagrangiane associate $\mathscr{Q}_k$, pensate come funzioni di classe C^2 su $j^1(\mathbb{V}^{n+1})$ soddisfano in ogni sistema di coordinate locali naturali adattate alle fibre di $j^1(\mathbb{V}^{n+1})$, le identità :*

$$\frac{\partial \mathscr{Q}_k}{\partial \dot{q}^h} + \frac{\partial \mathscr{Q}_h}{\partial \dot{q}^k} = 0\,, \quad \textit{ovunque e per ogni } k, h = 1, \ldots, n. \tag{7.21}$$

$$\frac{\partial \mathscr{Q}_k}{\partial q^h} - \frac{\partial \mathscr{Q}_h}{\partial q^k} = \frac{d}{dt}\left(\frac{\partial \mathscr{Q}_h}{\partial \dot{q}^k}\right), \quad \textit{ovunque e per ogni } k, h = 1, \ldots, n. \tag{7.22}$$

(b) *Viceversa le componenti lagrangiane $\mathscr{Q}_k$, pensate come funzioni di classe C^2 su $j^1(\mathbb{V}^{n+1})$ e lineari nelle coordinate $\dot{q}^k$, ammettono potenziale generalizzato in un intorno sufficientemente piccolo di ogni punto dato in coordinate da $(t_0, q_0^1, \ldots, q_0^n, \dot{q}_0^1, \ldots \dot{q}_0^n)$ di $j^1(\mathbb{V}^{n+1})$ se soddisfano le condizioni (7.21) e (7.22). L'espressione esplicita del potenziale generalizzato è in tal caso:*

$$\mathscr{V}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \int_0^1 \sum_{k=1}^n q^k \mathscr{Q}_k(t, q_0 + sq^1, \ldots, q_0 + sq^n, \dot{q}_0^1 + s\,\dot{q}^1, \ldots, \dot{q}^n + s\,\dot{q}^n) ds\,. \tag{7.23}$$

$\diamondsuit$

Dimostrazione. La dimostrazione di (a) è stata data sopra. Passiamo a (b). Consideriamo un punto di $j^1(\mathbb{V}^{n+1})$ di coordinate $(t_0, q_0^1, \ldots, q_0^n, \dot{q}_0^1, \ldots, \dot{q}_0^n) = (t_0, 0, \ldots, 0, 0, \ldots, 0)$. Questa non è una restrizione dato che possiamo sempre restringerci a tale caso traslando l'origine delle coordinate. Restringiamoci ad un intorno U di tale punto per cui le coordinate $(t, sq^1, \ldots, sq^n, s\,\dot{q}^1, \ldots, s\,\dot{q}^n)$ per $s \in [0, 1]$ siano associate a punti di $j^1(\mathbb{V}^{n+1})$ in U quando vale $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \in U$ (lasciamo al lettore la prova dell'esistenza di tale intorno basata sul fatto che le palle aperte sufficientemente piccole definite in coordinate sono una base della topologia di ogni varietà differenziabile ed in particolare di $j^1(\mathbb{V}^{n+1})$). Vogliamo provare che, nelle ipotesi fatte, da (7.23) si ricava

$$\frac{\partial \mathscr{V}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}}{\partial \dot{q}^k} = \mathscr{Q}_k(t, q^1, \ldots q^n, \dot{q}^1, \ldots, \dot{q}^n)\,.$$

Lavorando in U l'espressione (7.23) è ben definita e fornisce le seguenti due identità , con ovvie notazioni:

$$\frac{\partial \mathscr{V}}{\partial q^k} = \int_0^1 ds \mathscr{Q}_k(t, sq, s\,\dot{q}) + \int_0^1 ds \sum_h sq^h \frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial q^k}, \quad \frac{\partial \mathscr{V}}{\partial \dot{q}^k} = \int_0^1 ds \sum_h sq^h \frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial \dot{q}^k}\,.$$

Si tenga presente nel seguito che, per ipotesi, $\partial \mathscr{Q}_h / \partial\, \dot{q}^k$ non dipende dalle $\dot{q}$. Usando le identità appena ottenute si ricava subito che

$$\frac{\partial \mathscr{V}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}}{\partial\, \dot{q}^k} = \int_0^1 ds \mathscr{Q}_k(t, sq, s\,\dot{q}) + \int_0^1 ds \sum_h sq^h \frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial q^k}$$

$$- \int_0^1 ds \sum_h s\dot{q}^h \frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial\, \dot{q}^k} - \int_0^1 ds \sum_{h,p} sq^h \frac{\partial^2 \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial\, \dot{q}^k \partial q^p} sq^p - \int_0^1 ds \sum_h sq^h \frac{\partial^2 \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial\, \dot{q}^k \partial t}.$$

Quindi, riordinando il secondo membro:

$$\frac{\partial \mathscr{V}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}}{\partial\, \dot{q}^k} = \int_0^1 ds \mathscr{Q}_k(t, sq, s\,\dot{q}) - \int_0^1 ds \sum_h s\,\dot{q}^h \frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial\, \dot{q}^k}$$

$$+ \int_0^1 ds \sum_h sq^h \left[\frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{Q}_h(t, sq, s\,\dot{q})}{\partial\, \dot{q}^k} \right].$$

Facendo infine uso delle (7.21) e (7.22) nel secondo membro, si ha l'identità finale:

$$\frac{\partial \mathscr{V}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}}{\partial\, \dot{q}^k} = \int_0^1 ds \mathscr{Q}_k(t, sq, s\,\dot{q}) + \int_0^1 ds \sum_h sq^h \frac{\partial \mathscr{Q}_k(t, sq, s\,\dot{q})}{\partial q^h} + \int_0^1 ds \sum_h s\,\dot{q}^h \frac{\partial \mathscr{Q}_k(t, sq, s\,\dot{q})}{\partial\, \dot{q}^h}$$

cioé :

$$\frac{\partial \mathscr{V}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}}{\partial\, \dot{q}^k} = \int_0^1 ds \frac{d}{ds} s \mathscr{Q}_k(t, sq, s\,\dot{q}),$$

da cui infine:

$$\frac{\partial \mathscr{V}}{\partial q^k} - \frac{d}{dt}\frac{\partial \mathscr{V}}{\partial\, \dot{q}^k} = \mathscr{Q}_k(t, q,\ \dot{q})$$

che è la nostra tesi. □

Osservazioni 7.4. Le condizioni date nel teorema assicurano l'esistenza locale di un potenziale generalizzato per un insieme di componenti lagrangiane. Non è però ovvio che i vari potenziali locali ottenuti al variare dell'intorno si raccordino definendo una struttura globale. Tale problema viene affrontato caso per caso.

7.2.4 Potenziali generalizzati delle forze inerziali.

Consideriamo il solito sistema $\mathfrak{S}$ di N punti materiali P_i con masse m_i. Supponiamo che il sistema sia sottoposto a $c = 3N - n$ vincoli olonomi ideali e quindi sia descritto su $j^1(\mathbb{V}^{n+1})$ dalle equazioni di Eulero-Lagrange in forma generica:

$$\frac{d}{dt}\frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial\, \dot{q}^k} - \frac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^k} = \mathscr{Q}_k, \quad \text{per } k = 1, \ldots, n, \tag{7.24}$$

$$\frac{dq^k}{dt} = \dot{q}^k, \quad \text{per } k = 1, \ldots, n, \tag{7.25}$$

dove $\mathscr{I}$ è un sistema di riferimento inerziale e le $\mathscr{Q}_k$ sono le componenti lagrangiane delle forze attive (vere) agenti sul sistema, e ci riferiamo a coordinate locali naturali adattate alle fibre di $j^1(\mathbb{V}^{n+1})$. Vogliamo confrontare queste equazioni con quelle relative ad un sistema di riferimento non inerziale $\mathscr{I}'$ di moto assegnato rispetto a $\mathscr{I}$. Continueremo ad usare le stesse coordinate locali $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$.

$$\frac{d}{dt}\frac{\partial \mathscr{T}|_{\mathscr{I}'}}{\partial\, \dot{q}^k} - \frac{\partial \mathscr{T}|_{\mathscr{I}'}}{\partial q^k} = \mathscr{Q}_k + \mathscr{Q}_k|_{\mathscr{I}'}\,, \quad \text{per } k = 1, \ldots, n, \tag{7.26}$$

$$\frac{dq^k}{dt} = \dot{q}^k\,, \quad \text{per } k = 1, \ldots, n, \tag{7.27}$$

le nuove componenti lagrangiane $\mathscr{Q}_k|_{\mathscr{I}'}$ sono relative alla forze inerziali presenti nel riferimento non inerziale $\mathscr{I}'$ e dipendono dal riferimento al contrario delle componenti lagrangiane $\mathscr{Q}_k$. Si ossertvi che se siamo nelle ipotesi di validità del teorema di esistenza ed unicità per entrambi i sistemi di equazioni di Eulero-Lagrange suddetti, ogni soluzione del primo deve anche essere soluzione del secondo e viceversa, dato che entrambi i sistemi equivalgono alle equazioni di Newton per i punti del sistema con le stesse forze vere (incluse quelle vincolari). Sussiste il seguente risultato.

Proposizione 7.1. *Consideriamo un sistema $\mathfrak{S}$ di N punti materiali descritto su $j^1(\mathbb{V}^{n+1})$. Assumiamo che i vettori posizione $\mathbf{x}_i := P_i - O$ di ciascun punto materiale di $\mathfrak{S}$ rispetto al riferimento $\mathscr{I}$ siano funzioni di classe C^3 nelle coordinate libere su $j^1(\mathbb{V}^{n+1})$ e nel tempo congiuntamente. e che il moto relativo del sistema di riferimento $\mathscr{I}'$ rispetto ad assi solidali con riferimento $\mathscr{I}$ sia assegnato da funzioni di classe C^3 nel tempo (in modo tale che $t \mapsto O'(t) - O$ sia C^3 e $t \mapsto \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)$ sia C^2).*
Allora le componenti lagrangiane $\mathscr{Q}_k|_{\mathscr{I}'}$ delle forze inerziali in $\mathscr{I}'$ sono individuate da un potenziale generalizzato $\mathscr{V}|_{\mathscr{I}'}$ e vale:

$$\mathscr{V}|_{\mathscr{I}'} = \mathscr{T}|_{\mathscr{I}} - \mathscr{T}|_{\mathscr{I}'}\,, \tag{7.28}$$

in ogni sistema di coordinate locali naturali $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ su $j^1(\mathbb{V}^{n+1})$. In particolare $\mathscr{T}|_{\mathscr{I}} - \mathscr{T}|_{\mathscr{I}'}$ è sempre una funzione lineare delle coordinate $\dot{q}^k$. $\diamondsuit$

Dimostrazione. L'energia cinetica $\mathscr{T}|_{\mathscr{I}'}$ sarà differente da $\mathscr{T}|_{\mathscr{I}}$. Indicheremo nel seguito con $\mathbf{x}_i(t, q^1, \ldots, q^n) = P_i - O$ il vettore posizione di P_i nel riferimento $\mathscr{I}$ e con $\mathbf{x}'_i(t, q^1, \ldots, q^n) := P_i - O'$ il vettore posizione dello stesso punto materiale rispetto al riferimento $\mathscr{I}'$. Queste funzioni sono C^2 per ipotesi. Indichiamo similmente con $\mathbf{v}_i(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ e $\mathbf{v}'_i(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ le corrispondenti velocità , che come noto sono funzioni *lineari* nelle $\dot{q}^k$ valendo:

$$\mathbf{v}_i(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \frac{\partial \mathbf{x}_i}{\partial t}(t, q^1, \ldots, q^n) + \sum_{k=1}^{n} \dot{q}^k \frac{\partial \mathbf{x}_i}{\partial q^k}(t, q^1, \ldots, q^n)\,.$$

La legge di trasformazione delle velocità si scrive allora:

$$\mathbf{v}_i = \mathbf{v}'_i + \mathbf{v}_{O'}|_{\mathscr{I}}(t) + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \mathbf{x}'_i$$

dove le funzioni $\mathbf{v}_{O'}|_{\mathscr{I}}(t)$ e $\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)$ sono assegnate per ipotesi e sono di classe C^2. Usando l'espressione per l'energia cinetica rispetto a $\mathscr{I}$ e $\mathscr{I}'$:

$$\mathscr{T}|_{\mathscr{I}} = \sum_{i=1} \frac{1}{2} m_i \mathbf{v}_i^2 \,, \quad \mathscr{T}|_{\mathscr{I}'} = \sum_{i=1} \frac{1}{2} m_i \mathbf{v}'^2_i$$

abbiamo che entrambe le funzioni sono di classe C^1 su $j^1(\mathbb{V}^{n+1})$. Nello stesso modo, le componenti lagrangiane delle forze inerziali

$$\mathbf{F}_i|_{\mathscr{I}'} = -m_i \mathbf{a}_{O'}|_{\mathscr{I}}(t) - m_i \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \mathbf{x}'_i) - 2m_i \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \mathbf{v}'_i$$
$$-m_P \dot{\boldsymbol{\omega}}_{\mathscr{I}'}|_{\mathscr{I}} \wedge \mathbf{x}'_i \,.$$

risultano essere di clsse C^1 su $j^1(\mathbb{V}^{n+1})$. Confrontando le due definizioni di energia cinetica, rispetto a $\mathscr{I}$ e rispetto a $\mathscr{I}'$ si trova infine che:

$$\mathscr{V}|_{\mathscr{I}'}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) - \mathscr{T}|_{\mathscr{I}'}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \,,$$

è di classe C^2 e vale:

$$\mathscr{V}|_{\mathscr{I}'}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \sum_{i=1}^{N} \frac{1}{2} m_i \left(\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \wedge \mathbf{x}'_i(t, q^1, \ldots, q^n) \right)^2$$
$$+ \frac{M}{2} \mathbf{v}_{O'}|_{\mathscr{I}}(t)^2 + \mathbf{v}_{O'}|_{\mathscr{I}}(t) \cdot \sum_{i=1}^{N} m_i \mathbf{v}'_i(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$$
$$+ \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t) \cdot \sum_{i=1}^{N} m_i \mathbf{x}'_i \wedge \left(\mathbf{v}_{O'}|_{\mathscr{I}}(t) + \mathbf{v}'_i(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \right) \,,$$

dove $M := \sum_{i=1}^{N} m_i$. Per costruzione, $\mathscr{V}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ è una funzione *lineare* nelle $\dot{q}^k$ e pertanto può essere interpretato come *potenziale generalizzato*. Per confronto tra (7.24) e (7.26) nel caso di totale assenza di forze attive vere e tenendo conto che, come osservato sopra, le soluzioni del primo sistema lo sono anche per il secondo e viceversa dato che siamo nelle ipotesi di validità del teorema di esistenza ed uncità , troviamo subito che vale

$$\frac{\partial \mathscr{V}|_{\mathscr{I}'}}{\partial q^k} - \frac{d}{dt} \frac{\partial \mathscr{V}|_{\mathscr{I}'}}{\partial \dot{q}^k} = \mathscr{Q}_k|_{\mathscr{I}'}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \,, \tag{7.29}$$

su ogni atto di moto $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ attraversato da una soluzione delle equazioni di Eulero-Lagrange (7.24)-(7.25) e (7.26)-(7.27). Dato che siamo nelle ipotesi di validità del teorema di esistenza, ogni atto di moto $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ è attraversato da una soluzione (quella che ha tale atto di moto come condizione iniziale). Quindi la (7.29) vale per ogni atto di moto. □

7.3 Configurazioni di equilibrio e stabilità .

In questa sezione specializzeremo alcune definizioni e risultati, ottenuti nel capitolo 5, a sistemi fisici di punti materiali che ammettono descrizione lagrangiana. La prima nozione di cui abbiamo bisogno è quella di *configurazione di equilibrio*. Si tratta di una ovvia specializzazione della definizione 5.2 al caso lagrangiano.

7.3.1 Configurazioni di equilibrio rispetto ad un riferimento.

Definizione 7.2. (**Configurazione di equilibrio rispetto ad un riferimento**.) Si consideri un sistema $\mathfrak{S}$ di N punti materiali sottoposti a $c = 3N - n \geq 0$ vincoli olonomi ideali e descritto su $j^1(\mathbb{V}^{n+1})$ e si fissi un sistema di riferimento $\mathscr{I}$. Consideriamo poi una carta locale naturale adattata alle fibre di $j^1(\mathbb{V}^{n+1})$ con coordinate $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ dove:

(i) $t \in \mathbb{R}$,

(ii) $(q^1, \ldots, q^n) \in V \subset \mathbb{R}^n$ aperto,

(iii) $(\ \dot{q}^1, \ldots,\ \dot{q}^n) \in \mathbb{R}^n$,

(iv) le coordinate $(t, q^1, \ldots, q^n)$ sono solidali (vedi definizione 6.1) con $\mathscr{I}$.

Si dice che la n-pla di coordinate $(q_0^1, \ldots, q_0^n) \in V$ individua una **configurazione di equilibrio nel riferimento** $\mathscr{I}$, se la curva:

$$(q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t)) = (q_0^1, \ldots, q_0^n, 0, \ldots, 0)\ , \quad \text{per ogni } t \in \mathbb{R}$$

è l'unica soluzione massimale in $\mathbb{R} \times V \times \mathbb{R}^n$ delle equazioni di Eulero-Lagrange (6.19) con condizioni iniziali

$$(q^1(0), \ldots, q^n(0),\ \dot{q}^1(0), \ldots,\ \dot{q}^n(0)) = (q_0^1, \ldots, q_0^n, 0, \ldots, 0)\ .$$

$\diamondsuit$

Osservazioni 7.5.

(1) È importante notare che la n-pla $(q_0^1, \ldots, q_0^n) \in V$ *non* individua una configurazione in $\mathbb{V}^{n+1}$, ma una *collezione di esse*, più precisamente *una soluzione delle equazioni di Eulero-Lagrange*. Tale curva appare descrivere un punto *in quiete nel riferimento* $\mathscr{I}$. La definizione di "configurazione di equilibrio" dipende quindi dal riferimento $\mathscr{I}$ scelto.

(2) La definizione tuttavia *non* dipende dal sistema di coordinate naturali usate purché esse siano solidali con $\mathscr{I}$. La dimostrazione è immediata conseguenza del fatto che, se due sistemi di coordinate naturali definiti sullo stesso aperto di $\mathbb{V}^{n+1}$: $(t, q^1, \ldots, q^n)$ e $(t', q'^1, \ldots, q'^n)$ adattati alle fibre di $\mathbb{V}^{n+1}$, sono entrambi solidali con $\mathscr{I}$, allora la trasformazione di coordinate $q^k = q^k(t', q'^1, \ldots, q'^n)$ per $k = 1, \ldots, n$ (e la sua inversa) non può dipendere esplicitamente dal tempo. Infatti, se $\mathbf{x}_i$ è il vettore posizione in $\mathscr{I}$ del punto materiale $P_i \in \mathfrak{S}$, $\mathbf{x}_i$ non può dipendere esplicitamente dal tempo quando parametrizzato tramite le coordinate primate oppure quelle non primate, per definizione di coordinate solidali definizione 6.1. D'altra parte la dipendenza

parametrica di $\mathbf{x}_i$ dalle coordinate primate si ottiene anche esprimendo (i) $\mathbf{x}_i$ in funzione delle coordinate non primate e poi (ii) esprimendo queste ultime in funzione di quelle primate:

$$\mathbf{x}_i = \mathbf{x}_i(q^1(t', q'^1, \ldots, q'^n) \ldots, q^n(t', q'^1, \ldots, q'^n))$$

Pertanto deve risultare che, per ogni scelta di $(t', q'^1, \ldots, q'^n)$ nell'intorno considerato:

$$\frac{\partial}{\partial t'}\mathbf{x}_i(q^1(t', q'^1, \ldots, q'^n) \ldots, q^n(t', q'^1, \ldots, q'^n)) = 0\,.$$

In altre parole:

$$\sum_{k=1}^{n} \frac{\partial \mathbf{x}_i}{\partial q^k}\frac{\partial q^k}{\partial t'} = \mathbf{0}\,.$$

Calcolando il quadrato di entrambi i membri dopo averli moltiplicati per le masse m_1 e sommato su i, si ricava

$$\sum_{k,h=1}^{n}\left(\sum_{i=1}^{N} m_i \frac{\partial \mathbf{x}_i}{\partial q^h}\cdot\frac{\partial \mathbf{x}_i}{\partial q^k}\right)\frac{\partial q^h}{\partial t'}\frac{\partial q^k}{\partial t'} = \sum_{k,h=1}^{n} 2a_{hk}\frac{\partial q^h}{\partial t'}\frac{\partial q^k}{\partial t'} = 0\,.$$

La matrice dei coefficienti a_{hk} è quella considerata in proposizione 6.2. In tale proposizione si è provato che tale matrice è strettamente definita positiva. Dobbiamo pertanto concludere che, per ogni $h = 1, \ldots, n$ e per tutte le configurazioni descritte da n-ple di coordinate $(q^1, \ldots, q^n)$ in V,

$$\frac{\partial q^h}{\partial t'} = 0\,.$$

Abbiamo la proposizione seguente, che fornisce condizioni necessarie e sufficienti affinché una configurazione sia di equilibrio.

Proposizione 7.2. *In relazione alla definizione 7.2 si considerino coordinate locali naturali su $j^1(\mathbb{V}^{n+1})$, $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ variabili nell'aperto $\mathbb{R}\times V\times\mathbb{R}^n$ e solidali con il riferimento $\mathscr{I}$. In tali coordinate si considerino le equazioni di Eulero-Lagrange:*

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial \dot{q}^k} - \dfrac{\partial \mathscr{T}|_{\mathscr{I}}}{\partial q^k} & = \mathscr{Q}_k|_{\mathscr{I}}\,, \\ \dfrac{dq^k}{dt} & = \dot{q}^k(t) \end{cases} \quad \textit{per } k = 1, \ldots, n,\,.$$

Si supponga infine che la funzione $\mathscr{T}|_{\mathscr{I}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ sia di classe C^2. Vale quanto segue.

(a) *Se le funzioni $\mathscr{Q}_k|_{\mathscr{I}} = \mathscr{Q}_k|_{\hat{\mathscr{I}}}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$, per $k = 1, \ldots, n$ sono di classe C^1*

e non dipendono esplicitamente dal tempo, $(q_0^1, \ldots, q_0^n) \in V$ *è una configurazione di equilibrio rispetto al riferimento* $\mathscr{I}$ *se e solo se*

$$\mathscr{Q}_k|_{\mathscr{I}}(q_0^1, \ldots, q_0^n, 0, \ldots, 0) = 0\,, \quad \textit{per } k = 1, \ldots, n. \tag{7.30}$$

(b) *Se il sistema* $\mathfrak{S}$ *ammette lagrangiana della forma* $\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$ *con* $\mathscr{U}|_{\hat{\mathscr{I}}} = \mathscr{U}|_{\hat{\mathscr{I}}}(q^1, \ldots, q^n)$ *di classe* C^2 *su* V, $(q_0^1, \ldots, q_0^n) \in V$ *individua una configurazione di equilibrio rispetto al riferimento* $\mathscr{I}$, *se e solo se* $(q_0^1, \ldots, q_0^n)$ *è un punto di stazionarietà di* $\mathscr{U}|_{\mathscr{I}}$*:*

$$\frac{\partial}{\partial q^k}\mathscr{U}|_{\mathscr{I}}(q^1, \ldots, q^n)\Big|_{(q_0^1, \ldots, q_0^n)} = 0\,, \quad \textit{per } k = 1, \ldots, n. \tag{7.31}$$

$\diamondsuit$

Dimostrazione. (a) Dato che le coordinate $t, q^1, \ldots, q^n$ sono solidali con $\mathscr{I}$, le funzioni che esplicitano i vettori posizione $\mathbf{x}_i$ dei punti $P_i \in \mathfrak{S}$ nel riferimento $\mathscr{I}$ non dipenderanno esplicitamente dal tempo. Di conseguenza l'espressione dell'energia cinetica $\mathscr{T}|_{\mathscr{I}}$ non dipenderà esplicitamente dal tempo ed avrà forma (vedi la sezione 6.1.3):

$$\mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n) = \sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k\,.$$

Le equazioni di Eulero-Lagrange saranno allora, tenendo conto che gli a_{hk} non dipendono esplicitamente dal tempo:

$$2\sum_k a_{hk}\frac{d^2q^k}{dt^2} + 2\sum_{k,r}\frac{\partial a_{hk}}{\partial q^r}\frac{dq^k}{dt}\frac{dq^r}{dt} - \sum_{r,k}\frac{\partial a_{rk}}{\partial q^h}\frac{dq^r}{dt}\frac{dq^k}{dt} = \mathscr{Q}_k|_{\mathscr{I}}(q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n)\,. \tag{7.32}$$

(Si osservi che, per costruzione, il primo membro si annulla se valutato su una curva costante $q^k(t) = q_0^k$ per ogni $k = 1, \ldots, n$ e ogni $t \in \mathbb{R}$; questo risultato avrebbe potuto essere falso se $\mathscr{T}|_{\mathscr{I}}$ avesse avuto una dipendenza temporale esplicita dal tempo (perché ?)). Tali equazioni si posso scrivere in forma normale semplicemente moltiplicando ambo i membri per l'inversa della matrice di coefficienti a_{hk} e con altre banali modifiche. Nelle ipotesi fatte, escludendo le funzioni $\mathscr{Q}^k$, le rimanenti funzioni delle coordinate $q^1, \ldots, q^n$ che appaiono nella forma esplicita delle equazioni di Eulero-Lagrange data sopra sono di classe C^1. I coefficienti della matrice inversa di $[a_{hk}]$ sono di classe C^1 nelle $q^1, \ldots, q^n$ dato che sono rapporti di polinomi (con denominatori mai nulli) dei coefficienti a_{hk}. Tenendo conto che, escludendo le funzioni $\mathscr{Q}^k$, la dipendenza dalle derivate delle coordinate q^k è polinomiale e che la funzione $\mathscr{Q}^k$ è di classe C^1, concludiamo che: *il secondo membro delle equazioni di Eulero Lagrange scritte in forma normale deve essere una funzione di classe (almeno)* C^1 *congiuntamente nelle variabili* q^k *e nelle derivate di esse.* Siamo allora nelle ipotesi del teorema di (esistenza e) unicità globale, teorema 3.5. Per provare la tesi è allora sufficiente verificare che se la soluzione costante:

$$(q^1(t), \ldots, q^n(t),\, \dot{q}^1(t), \ldots,\, \dot{q}^n(t)) = (q_0^1, \ldots, q_0^n, 0, \ldots, 0)\,, \quad \text{per ogni } t \in \mathbb{R} \tag{7.33}$$

risolve le equazioni di Eulero-Lagrange allora vale (7.30) e che se vale (7.30) allora

$$(q^1(t), \dots, q^n(t),\ \dot q^1(t), \dots,\ \dot q^n(t)) = (q^1_0, \dots, q^n_0, 0, \dots, 0)\ , \quad \text{per ogni } t \in \mathbb{R}$$

è una soluzione delle equazioni di Eulero-Lagrange.
Se la soluzione *costante* (7.33) risolve le equazioni di Eulero-Lagrange (7.32), il primo membro deve annullarsi su di essa per costruzione; ne consegue che vale la (7.30). Se d'altra parte vale (7.30), la soluzione costante (7.33) risolve le equazioni di Eulero-Lagrange (7.32) dato che ne annulla il primo membro ed il secondo è già nullo per (7.30)).
(b) La dimostrazione è immediata dalla precedente tenendo conto del fatto che:

$$\mathscr{Q}_k|_{\mathscr{I}} = -\frac{\partial}{\partial q^k}\mathscr{U}|_{\mathscr{I}}(q^1, \dots, q^n)\ .$$

□

Osservazioni 7.6. La dipendenza delle $\mathscr{Q}_k|_{\mathscr{I}}$ dal riferimento $\mathscr{I}$ sussiste solo quando si considerano riferimenti $\mathscr{I}$ che sono non inerziali: le forze inerziali dipendono dal riferimento. Le forze vere sono invece indipendenti dal riferimento.

La proposizione può essere modificata nella seguente utile estensione.

Proposizione 7.3. *In relazione alla definizione 7.2 si considerino coordinate locali naturali su $j^1(\mathbb{V}^{n+1})$, $(t, q^1, \dots, q^n, \dot q^1, \dots, \dot q^n)$ variabili nell'aperto $\mathbb{R}\times V\times\mathbb{R}^n$ e solidali con il riferimento $\mathscr{I}$. In tali coordinate si considerino le equazioni di Eulero-Lagrange:*

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial \mathscr{T}|_{\hat{\mathscr{I}}}}{\partial \dot q^k} - \dfrac{\partial \mathscr{T}|_{\hat{\mathscr{I}}}}{\partial q^k} & = \mathscr{Q}_k|_{\hat{\mathscr{I}}}\ , \\ \dfrac{dq^k}{dt} & = \dot q^k(t) \end{cases} \quad \textit{per } k = 1, \dots, n, .$$

dove $\hat{\mathscr{I}} \neq \mathscr{I}$. Si supponga infine che la funzione $\mathscr{T}|_{\hat{\mathscr{I}}}(t, q^1, \dots, q^n,\ \dot q^1, \dots,\ \dot q^n)$ sia di classe C^2 e non dipenda esplicitamente dal tempo. Vale quanto segue.
(a) *Si consideri la decomposizione standard di $\mathscr{T}|_{\hat{\mathscr{I}}}$ (vedi la sezione 6.1.3):*

$$\mathscr{T}|_{\hat{\mathscr{I}}} = 2\sum_{h,k} a_{hk}\,\dot q^h\,\dot q^k + \sum_k b_k\,\dot q^k + c\,. \tag{7.34}$$

Se le funzioni $\mathscr{Q}_k|_{\hat{\mathscr{I}}} = \mathscr{Q}_k|_{\hat{\mathscr{I}}}(t, q^1, \dots, q^n,\ \dot q^1, \dots,\ \dot q^n)$, per $k = 1, \dots, n$ sono di classe C^1 e non dipendono esplicitamente dal tempo, $(q^1_0, \dots, q^n_0) \in V$ è una configurazione di equilibrio rispetto al riferimento $\mathscr{I}$ se e solo se:

$$\mathscr{Q}_k|_{\hat{\mathscr{I}}}(q^1_0, \dots, q^n_0, 0, \dots, 0) = -\frac{\partial c(q^1_0, \dots, q^n_0, 0, \dots, 0)}{\partial q^k}\ , \quad \textit{per } k = 1, \dots, n. \tag{7.35}$$

(b) *Se il sistema* $\mathfrak{S}$ *ammette lagrangiana della forma* $\mathscr{L}|_{\hat{\mathscr{I}}} = \mathscr{T}|_{\hat{\mathscr{I}}} - \mathscr{U}|_{\hat{\mathscr{I}}}$ *con* $\mathscr{U}|_{\hat{\mathscr{I}}} = \mathscr{U}|_{\hat{\mathscr{I}}}(q^1, \ldots, q^n)$ *di classe* C^2 *su* V, $(q_0^1, \ldots, q_0^n) \in V$ *individua una configurazione di equilibrio rispetto al riferimento* $\mathscr{I}$, *se e solo se* $(q_0^1, \ldots, q_0^n)$ *è un punto di stazionarietà di* $\mathscr{U}' := \mathscr{U}|_{\hat{\mathscr{I}}} + c$:

$$\frac{\partial}{\partial q^k} \mathscr{U}'(q^1, \ldots, q^n)\Big|_{(q_0^1, \ldots, q_0^n)} = 0\,, \quad \textit{per } k = 1, \ldots, n. \tag{7.36}$$

$\diamondsuit$

Dimostrazione. Vale:

$$\mathscr{T}|_{\hat{\mathscr{I}}}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \sum_{h,k=1}^{n} a_{hk}(t, q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k + \sum_{k=1}^{n} b_k(t, q^1, \ldots, q^n)\, \dot{q}^k$$

$$+ c(t, q^1, \ldots, q^n)\,,$$

dove a primo membro non abbiamo scritto la variabile t dato che, per ipotesi $\mathscr{T}|_{\hat{\mathscr{I}}}$ non dipende esplicitamente da t. Derivando in t abbiamo dunque che:

$$0 = \sum_{h,k=1}^{n} \frac{\partial a_{hk}}{\partial t}\, \dot{q}^h\, \dot{q}^k + \sum_{k=1}^{n} \frac{\partial b_k}{\partial t}\, \dot{q}^k + \frac{\partial c}{\partial t}\,.$$

Per ogni scelta di $(t, q^1, \ldots, q^n)$ fissate, il polinomio di secondo grado nelle variabili $\dot{q}^k$ (che sono indipendenti dalle altre variabili) che appare a secondo membro è nullo. Un polinomio è identicamente nullo se e solo se ha tutti i coefficienti nulli. Ne consegue che le derivate parziali in t dei coefficienti a_{hk}, b_k e c sono identicamente nulle. Possiamo allora concludere che:

$$\mathscr{T}|_{\hat{\mathscr{I}}}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k + \sum_{k=1}^{n} b_k(q^1, \ldots, q^n)\, \dot{q}^k + c(q^1, \ldots, q^n)\,,$$

e la dimostrazione procede come nel caso precedente tenendo conto che ora le equazioni di Eulero-Lagrange risultano essere scritte:

$$2\sum_k a_{hk} \frac{d^2 q^k}{dt^2} + 2\sum_{k,r} \frac{\partial a_{hk}}{\partial q^r} \frac{dq^k}{dt} \frac{dq^r}{dt} + \sum_r \frac{\partial b_h}{\partial q^r} \frac{dq^r}{dt} - \sum_{r,k} \frac{\partial a_{rk}}{\partial q^h} \frac{dq^r}{dt} \frac{dq^k}{dt} + \sum_k^n \frac{\partial b_k}{\partial q^h} \frac{dq^k}{dt}$$

$$= \frac{\partial c(q_0^1, \ldots, q_0^n, 0, \ldots, 0)}{\partial q^k} + \mathscr{Q}_k|_{\hat{\mathscr{I}}}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)\,.$$

$\square$

Osservazioni 7.7.
(1) Come nel caso precedente, la dipendenza delle $\mathscr{Q}_k|_{\hat{\mathscr{I}}}$ dal riferimento $\hat{\mathscr{I}}$ sussiste solo quando si considerano riferimenti $\hat{\mathscr{I}}$ che sono non inerziali: le forze inerziali dipendono dal riferimento.

Le forze vere sono invece indipendenti dal riferimento.
(2) Ci si può chiedere se esistano davvero casi in cui $\mathscr{I} \neq \hat{\mathscr{I}}$, le coordinate libere sono solidali con $\mathscr{I}$, ma l'energia cinetica riferita a $\hat{\mathscr{I}}$ non dipende esplicitamente dal tempo. La risposta è positiva ed un esempio semplice si quando $\hat{\mathscr{I}}$ è un riferimento inerziale mentre $\mathscr{I}$ non lo è, ma ruota senza traslare rispetto a $\hat{\mathscr{I}}$ con vettore $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \neq \mathbf{0}$ costante.

Esempi 7.2. Si fissi un sistema destrorso di assi ortogonali $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ solidali con il riferimento $\mathscr{I}$ ed uscenti dall'origine O. Si consideri un punto materiale P di massa m vincolato sull'asse x, supposto liscio, e connesso ad O tramite una molla di massa nulla, costante elastica $k > 0$ e lunghezza nulla a riposo. Il riferimento $\mathscr{I}$ ruota rispetto al riferimento inerziale $\hat{\mathscr{I}}$ con $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = \Omega\, \mathbf{e}_z$ dove $\Omega > 0$ è costante. Sul punto P, oltre alla forza della molla ed alla reazione vincolare, agisce una forza viscosa $-\gamma \mathbf{v}_P|_{\mathscr{I}}$ con $\gamma \geq 0$ costante.
In questo caso, usando la coordinata x del punto P come coordinata libera, abbiamo definito un sistema di coordinate naturali (t, x) che soddisfa tutti i requisiti della definizione 7.2, rispetto al riferimento $\mathscr{I}$. Nel caso in esame abbiamo immediatamente che:

$$\mathscr{T}|_{\hat{\mathscr{I}}} = \frac{m}{2}\dot{x}^2 + \frac{m}{2}\Omega^2 x^2 \,.$$

La funzione c è quindi data da

$$c(x) = \frac{m}{2}\Omega^2 x^2 \,.$$

Riguardo alle forze abbiamo che (omettendo la specificazione $|_{\hat{\mathscr{I}}}$ essendo $\hat{\mathscr{I}}$ inerziale e le forze vere):

$$\mathscr{Q} = -kx - \gamma \dot{x} \,.$$

Le configurazioni di equilibrio x_0 nel riferimento non inerziale $\hat{\mathscr{I}}$ si ottengono, per la proposizione 7.3, imponendo che:

$$\left.\frac{dc(x)}{dx}\right|_{x_0} = \mathscr{Q}(x_0, 0) \,.$$

In altre parole deve essere $m\Omega^2 x_0 = k x_0$, da cui

$$x_0 = 0 \,,$$

dove abbiamo assunto che $m\Omega^2/k \neq 1$.
Possiamo studiare il problema delle configurazioni di equilibrio usando direttamente la proposizione 7.2. In questo caso si usa solo il riferimento $\mathscr{I}$. In esso

$$\mathscr{T}|_{\mathscr{I}} = \frac{m}{2}\dot{x}^2 \,.$$

Si osservi che nel riferimento $\mathscr{I}$ appaiono anche la forza inerziale di Coriolis, $-2m\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge \mathbf{v}_P|_{\mathscr{I}}$, e quella centrifuga $-m\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge (\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge \mathbf{x})$, con $\mathbf{x} := x\, \mathbf{e}_x$. La loro somma vale:

$$-2m\Omega\dot{x}\, \mathbf{e}_y + m\Omega^2 x\, \mathbf{e}_x \,.$$

Riguardo alle forze abbiamo che mettendo la specificazione $|_{\mathscr{I}}$ essendo $\mathscr{I}$ *non* inerziale ed essendo parte delle forze attive di natura inerziale):

$$\mathscr{Q}|_{\mathscr{I}} = -kx - \gamma\dot{x} + m\Omega^2 x \,.$$

Si osservi che la forza di Coriolis non produce componenti lagrangiane essendo:

$$-2m\Omega\dot{x}\,\mathbf{e}_y \cdot \frac{\partial \mathbf{x}}{\partial x} = -2m\Omega\dot{x}\,\mathbf{e}_y \cdot\ \mathbf{e}_x = 0 \,.$$

Le configurazioni di equilibrio x_0 nel riferimento non inerziale $\hat{\mathscr{I}}$ si ottengono, per la proposizione 7.2, imponendo che:

$$\mathscr{Q}|_{\mathscr{I}}(x_0, 0) = 0 \,.$$

In altre parole deve essere $-kx_0 + m\Omega^2 x = 0$, da cui esattamente come prima:

$$x_0 = 0 \,,$$

dove abbiamo assunto che $m\Omega^2/k \neq 1$.

7.3.2 Equilibrio stabile ed instabile, teorema di Lagrange-Dirichlet.

Ripetiamo per i sistemi lagrangiani alcune considerazioni già fatte nel caso generale nel capitolo 5. Dal punto di vista fisico, la nozione di configurazione di equilibrio data sopra non caratterizza in modo appropriato le configurazioni di equilibrio reali dei sistemi fisici (e degli altri sistemi della realtà descrivibili con equazioni differenziali). Una ragione, forse la più importante, è dovuta al fatto che le configurazioni dei sistemi fisici non corrispondono a configurazioni geometriche *ideali*, perché lo stato di ogni sistema fisico (la posizione e la forma stessa degli oggetti) è sottoposto a continue interazioni con l'ambiente circostante (per lo più interazioni di carattere termodinamico), che rendono lo stato fluttuante, sia pur con fluttuazioni molto piccole. Per questo motivo le configurazioni fisiche di equilibrio *osservabili nella realtà fisica* hanno anche una proprietà di *stabilità* sotto piccole perturbazioni dello stato iniziale. Queste proprietà di stabilità vengono definite a seconda del modello matematico adottato. Nel seguito ci riferiremo alle nozioni più elementari possibili specializzando al caso lagrangiano quelle illustrate nel capitolo 5. L'idea centrale della definizione di configurazione di equilibrio (rispetto al riferimento $\mathscr{I}$) stabile è quella che ogni soluzione rimanga confinata *vicino* alla configurazione (nel giudizio del riferimento $\mathscr{I}$) anche partendo da condizioni iniziali, che non coincidono con la configurazione di equilibri, ma se ne discostano un poco.

Definizione 7.3. (**Configurazione di equilibrio stabile rispetto ad un riferimento**) In riferimento alla definizione 7.2 ed alle notazioni usate in essa, se l'n-pla di coordinate $Q_0 := (q_0^1, \ldots, q_0^n) \in V$ individua una configurazione di equilibrio rispetto al riferimento $\mathscr{I}$, indichiamo con A_0 la $2n$-pla $(q_0^1, \ldots, q_0^n, 0, \ldots, 0) \in V \times \mathbb{R}^n$. Q_0 è detta:

(a) stabile nel futuro (rispetto al riferimento $\mathscr{I}$), se per ogni intorno aperto $B \ni A_0$ esiste un intorno aperto $E \ni A_0$ tale che ogni soluzione delle equazioni di Eulero-Lagrange (6.19) in $V \times \mathbb{R}^n$, $\gamma : I \to V \times \mathbb{R}^n$ con $\gamma(0) = \gamma_0 \in E$, soddisfa $\gamma(t) \in B$ se $I \ni t \geq 0$;

(b) stabile nel passato (rispetto al riferimento $\mathscr{I}$), se per ogni intorno aperto $B \ni A_0$ esiste un intorno aperto $E \ni A_0$ tale che ogni soluzione di (6.19) in $V \times \mathbb{R}^n$, $\gamma : I \to V \times \mathbb{R}^n$ con $\gamma(0) = \gamma_0 \in E$, soddisfa $\gamma(t) \in B$ se $I \ni t \leq 0$;

(c) stabile (rispetto al riferimento $\mathscr{I}$), se è stabile nel futuro e stabile nel passato;

(d) instabile (rispetto al riferimento $\mathscr{I}$) (rispettivamente **nel futuro** o **nel passato**), se non è stabile (rispettivamente nel futuro o nel passato);

(e) asintoticamente stabile nel futuro (rispetto al riferimento $\mathscr{I}$), se è stabile nel futuro ed esiste un intorno aperto $F \ni A_0$ tale che ogni soluzione massimale di (6.19) in $V \times \mathbb{R}^n$, $\gamma : I \to V \times \mathbb{R}^n$ con $\gamma(0) = \gamma_0 \in F$, è completa nel futuro e soddisfa $\gamma(t) \to A_0$ per $t \to +\infty$;

(f) asintoticamente stabile nel passato (rispetto al riferimento $\mathscr{I}$), se è stabile nel passato ed esiste un intorno aperto $P \ni A_0$ tale che ogni soluzione massimale di (6.19) in $V \times \mathbb{R}^n$, $\gamma : I \to V \times \mathbb{R}^n$ con $\gamma(0) = \gamma_0 \in P$, è completa nel passato e soddisfa $\gamma(t) \to A_0$ per $t \to -\infty$; $\diamondsuit$.

Osservazioni 7.8.
(1) Nel caso in cui le $\mathscr{Q}_k|_{\mathscr{I}}$ siano di classe C^1 e la $\mathscr{T}|_{\mathscr{I}}$ sia di classe C^2 se $Q_0 := (q_0^1, \ldots, q_0^n) \in V$ individua una configurazione di equilibrio rispetto al riferimento $\mathscr{I}$ che è stabile nel futuro, in virtù della proposizione 3.5, le soluzioni massimali delle equazioni di Eulero-Lagrange con condizione iniziale nell'intorno E sono sicuramente complete nel futuro (possiamo prendere B in modo che abbia chiusura compatta). Lo stesso risultato si ottiene per la stabilità nel passato.
(2) Bisogna avere ben chiaro il fatto che, trattando la stabilità di configurazioni di sistemi fisici, ed in particolare di sistemi lagrangiani, *la nozione di stabilità include informazioni e richieste sulle velocità e non solo sulle configurazioni.*

Possiamo ora enunciare l'analogo del teorema 5.5 in formulazione lagrangiana includendo l'analogo del teorema 5.7.

Teorema 7.3. (**Lagrange-Dirichlet**) *Si consideri un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$ sottoposti a $c = 3N - n \geq 0$ vincoli olonomi ideali e descritto su $j^1(\mathbb{V}^{n+1})$. Si consideri un sistema di coordinate naturali $(t, q^1, \ldots, q^n)$ variabili in $\mathbb{R} \times V$ e solidali con $\mathscr{I}$, e si assuma infine che le forze attive agenti su sistema si decompongano in una parte conservativa in $\mathscr{I}$ con energia potenziale totale $\mathscr{U}|_{\mathscr{I}}$ (di classe C^2 nelle variabili $(q^1, \ldots, q^n)$) ed una parte non conservativa data dalle forze non dipendenti esplicitamente dal tempo $\mathbf{F}_i|_{\mathscr{I}} = \mathbf{F}_i(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}})$, con $i = 1, \ldots, N$ (di classe C^1 quando espresse nelle variabili $(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$) eventualmente tutte nulle, dove $\mathbf{x}_i := P - O$ è il vettore posizione di P_i nel riferimento $\mathscr{I}$.*
Assumendo che l'energia cinetica $\mathscr{T}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ sia di classe C^2, una configurazione di equilibrio rispetto a $\mathscr{I}$ individuata da $Q_0 := (q_0^1, \ldots, q_0^n) \in V$ è :

(a) *stabile nel futuro se la restrizione di $\mathscr{U}|_{\mathscr{I}}$ ad un intorno aperto di di Q_0 ha un minimo*

stretto in tale punto in aggiunta alla condizione:

$\sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} \leq 0$ *per ogni scelta dei vettori velocità* $\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}$*;*

(b) *stabile nel passato e nel futuro se la restrizione di* $\mathscr{U}|_{\mathscr{I}}$ *ad un intorno aperto di di* Q_0 *ha un minimo stretto in* Q_0 *in aggiunta alla condizione:*

$\sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} \leq 0$ *per ogni scelta dei vettori velocità* $\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}$

(c) *stabile nel futuro e asintoticamente stabile nel futuro se il punto di minimo stretto* Q_0 *è anche un punto stazionario* isolato[2] *dell'energia potenziale* $\mathscr{U}_{\mathscr{I}}|$ *e in aggiunta alla la condizione:*

$\sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}} < 0$*, per* $(\mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \neq (\mathbf{0}, \ldots, \mathbf{0})$. $\diamondsuit$

Osservazioni 7.9. Prima di procedere con la dimostrazione osserviamo che se $\mathscr{Q}_k|_{\mathscr{I}}$ denota le componenti lagrangiane del sistema di forze non conservative $\mathbf{F}_i|_{\mathscr{I}}$, vale l'identità :

$$\sum_{k=1}^{n} \mathscr{Q}_k|_{\mathscr{I}}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots, dq^n)\, \dot{q}^k = \sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{v}_{P_1}|_{\mathscr{I}}, \ldots, \mathbf{v}_{P_N}|_{\mathscr{I}}) \cdot \mathbf{v}_{P_k}|_{\mathscr{I}}\,. \tag{7.37}$$

Questa identità , che esprime la *potenza complessiva* del sistema di forze $\{\mathbf{F}_i|_{\mathscr{I}}\}_{i=1,\ldots,N}$ rispetto al riferimento $\mathscr{I}$ usando coordinate naturali solidali con $\mathscr{I}$ consente di riformulare il teorema appena enunciato in modo alternativo, senza fare riferimento alle forze, ma solo alle componenti lagrangiane di esso.

Dimostrazione. (a) e (b). La dimostrazione è la stessa data per il teorema 5.5 con la differenza che ora la funzione di Liapunov (che coincide sempre con l'energia meccanica del sistema $\mathfrak{S}$ nel riferimento $\mathscr{I}$) è :

$$\mathscr{H}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) := \sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k + \mathscr{U}|_{\mathscr{I}}(q^1, \ldots, q^n)\,.$$

Si osservi prima di tutto che, per ogni scelta della n-pla $(q^1, \ldots, q^n)$ fissata, la funzione quadratica nelle $\dot{q}^k$: $\sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k$ ha un minimo stretto per $\dot{q}^k = 0$, $k = 1, \ldots, n$. Questo segue dal fatto che la matrice dei coefficienti a_{hk} è strettamente definita positiva come dimostrato nella proposizione 6.2.

Nel caso in esame, le equazioni di Eulero-Lagrange si scrivono, dove le $\mathscr{Q}_k$ sono le componenti lagrangiane del sistema di forze non conservative $\{\mathbf{F}_i|_{\mathscr{I}}\}_{i=1,\ldots,N}$:

$$2\sum_{k} a_{hk} \frac{d^2 q^k}{dt^2} + 2\sum_{k,r} \frac{\partial a_{hk}}{\partial q^r} \frac{dq^k}{dt} \frac{dq^r}{dt} - \sum_{r,k} \frac{\partial a_{rk}}{\partial q^h} \frac{dq^r}{dt} \frac{dq^k}{dt} + \frac{\partial \mathscr{U}|_{\mathscr{I}}}{\partial q^h} = \mathscr{Q}_h|_{\mathscr{I}}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$$

[2] Il fatto che $(q_0^1, \ldots, q_0^n)$ individui un minimo stretto dell'energia potenziale non è sufficiente ad assicurare che sia anche un punto di stazionarietà isolato dell'energia potenziale: si consideri, per $N = 1$, la funzione $\mathscr{U}(q) := q^4(2 + \sin(1/q))$ se $q \neq 0$ e $\mathscr{U}(0) := 0$. $\mathscr{U} \in C^2(\mathbb{R})$ ed ha un minimo stretto per $q = 0$ che non corrisponde ad un punto di stazionarietà isolato.

Moltiplicando per $\dot{q}^h$, sommando su h ed infine usando l'identità , valida sui moti del sistema, $\dot{q}^k = dq^k/dt$, si trova facilmente che, sulle soluzioni delle equazioni di Eulero-Lagrange vale:

$$\frac{d}{dt}\left(\sum_{h,k=1}^{n} a_{hk}(q^1(t),\ldots,q^n(t))\,\dot{q}^h(t)\,\dot{q}^k(t) + \mathscr{U}|_{\mathscr{I}}(q^1(t),\ldots,q^n(t))\right)$$

$$= \sum_{h=1}^{n} \dot{q}^h \mathscr{Q}_h|_{\mathscr{I}}(q^1(t),\ldots,q^n(t),\ \dot{q}^1(t),\ldots,\ \dot{q}^n(t))\,.$$

Tenendo conto di (7.37) si ha infine:

$$\frac{d}{dt}\left(\sum_{h,k=1}^{n} a_{hk}(q^1(t),\ldots,q^n(t))\,\dot{q}^h(t)\,\dot{q}^k(t) + \mathscr{U}|_{\mathscr{I}}(q^1(t),\ldots,q^n(t))\right)$$

$$= \sum_{k=1}^{N} \mathbf{F}_k(\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{v}_{P_1}|_{\mathscr{I}},\ldots,\mathbf{v}_{P_N}|_{\mathscr{I}})\cdot\mathbf{v}_{P_k}|_{\mathscr{I}}\,,$$

dove $\mathbf{x}_i = \mathbf{x}_i((q^1(t),\ldots,q^n(t)))$ e $\mathbf{v}_{P_k}|_{\mathscr{I}} = \mathbf{v}_{P_k}|_{\mathscr{I}}(q^1(t),\ldots,q^n(t),\ \dot{q}^1(t),\ldots,\ \dot{q}^n(t))$ sono valutati sul moto del sistema considerato. Il seguito della dimostrazione è identico a quello dato nella dimostrazione del teorema 7.3 per le asserzioni (a) e (b), usando $\mathscr{H}$ (ristretta ad un intorno di Q_0) come funzione di Liapunov al posto di W. (c) La dimostrazione è essenzialmente identica a quella del teorema 5.7 tenendo conto delle osservazioni fatte sopra. □

Per concludere enunciamo la proposizione analoga alla proposizione 5.8, una condizione sufficiente per l'instabilità basata sulle proprietà della matrice hessiana dell'energia potenziale, nel caso di sistemi con forze puramente conservative. La dimostrazione è lasciata per esercizio, ricalcando quella della proposizione 5.8.

Proposizione 7.4. *Nelle stesse ipotesi del teorema 7.3 si assuma che le forze siano completamente conservative (non siano cioé presenti le forze $\mathbf{F}_i$). Una configurazione di equilibrio rispetto a $\mathscr{I}$ individuata da $(q_0^1,\ldots,q_0^n) \in V$ è instabile nel futuro (rispettivamente nel passato) se la matrice hessiana dell'energia potenziale ha un autovalore negativo (rispettivamente positivo), in $(q_0^1,\ldots,q_0^n)$.* ♢

Capitolo 8

Simmetrie e leggi di conservazione: teoremi di Nöther e di Jacobi.

In questo capitolo consideriamo altri argomenti più avanzati di Meccanica Lagrangiana che hanno avuto delle fondamentali conseguenze nei successivi sviluppi della fisica matematica e teorica, anche in ambiti molto lontani dalla meccanica classica.
Introdurremo i teoremi che mostrano la relazione tra simmetrie della lagrangiana e le varie quantità conservate (integrali primi): il teorema di Nöther e quello di Jacobi. Nell'ultima sezione riformuleremo il teorema di Nöther in una versione molto generale, usando il linguaggio appropriato della geometria differenziale, che include il teorema di Jacobi come sottocaso.

8.1 Il legame tra simmetria e leggi di conservazione: coordinate cicliche.

In questa sezione cominceremo a studiare il legame tra simmetrie di un sistema che ammette lagrangiana ed alcuni dei suoi integrali primi, cioé quantità conservate sui moti del sistema. Ci concentreremo sulla nozione di *coordinata ciclica* e di momento coniugato associato. Esemplificheremo i risultati mostrando come, sotto le debite ipotesi sul sistema fisico in esame, l'invarianza per traslazioni sia connessa alla conservazione dell'impulso, mentre l'invarianza per rotazioni sia connessa alla conservazione del momento angolare.

8.1.1 Coordinate cicliche e conservazione dei momenti coniugati.

Consideriamo un sistema fisico di punti materiali $\mathfrak{S}$ descritto nello spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$ in termini di una lagrangiana $\mathscr{L}$. Questa lagrangiana potrà essere associata ad un riferimento (inerziale o meno), ma per il momento rimaniamo su un piano generale. Supponiamo di avere scelto coordinate locali naturali $t, q^1, \ldots, q^n$, $\dot{q}^1, \ldots,$ $\dot{q}^n$ e che, nello spaziotempo degli

atti di moto nelle dette coordinate risulti, in riferimento ad una fissata coordinata q^j:

$$\frac{\partial \mathscr{L}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)}{\partial q^j} = 0\,, \quad \text{per ogni scelta di } (t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n).$$

Diremo in tal caso che la coordinata q^j è **ciclica** o **ignorabile**. Si osservi che la condizione di ciclicità rispetto alla coordinata q^j è del tutto *equivalente* alla *condizione di invarianza* della lagrangiana sotto trasformazioni che alterino solo la coordinata q^j lasciando immutate le altre:

$$(q^1, \ldots, q^{j-1}, q^j, q^{j+1}, \ldots, q^n) \mapsto (q^1, \ldots, q^{j-1}, q^j + \Delta q^j, q^{j+1}, \ldots, q^n)\,.$$

Dire che vale (8.2) è infatti evidentemente equivalente a dire che:

$$\begin{cases} \mathscr{L}(t, q^1, \ldots, q^{j-1}, q^j + \Delta q^j, q^{j+1}, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \mathscr{L}(t, q^1, q^{j-1}, q^j, q^{j+1}, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n), \\ \qquad\qquad \text{per ogni scelta di } (t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n). \end{cases}$$

Ora le equazioni di Eulero-Lagrange (6.44) forniscono immediatamente l'identità :

$$\frac{d}{dt}\frac{\partial \mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots, \dot{q}^j(t))}{\partial\, \dot{q}^j} = \frac{\mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots, \dot{q}^j(t))}{\partial q^j}$$

dove $q^k = q^k(t)$ per $k = 1, \ldots, n$, è una generica soluzione delle equazioni di Eulero-Lagrange rappresentata nelle coordinate locali dette (e pertanto, su di essa $\dot{q}^k(t) = dq^k/dt$ per $k = 1, \ldots, n$). Nelle nostre ipotesi di ciclicità della coordinata q^j, la derivata parziale a secondo membro è nulla e pertanto:

$$\frac{d}{dt}\frac{\partial \mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots, \dot{q}^j(t))}{\partial\, \dot{q}^j} = 0\,.$$

Questa equazione dice che il **momento coniugato** alla coordinata q^j, cioè la funzione definita nelle coordinate locali su $j^1(\mathbb{V}^{n+1})$ considerate:

$$p_j(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) := \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)}{\partial\, \dot{q}^j} \tag{8.1}$$

si conserva nel tempo *su ogni moto del sistema*. Abbiamo quindi provato la seguente utile proposizione.

Proposizione 8.1. *Consideriamo un sistema fisico di punti materiali $\mathfrak{S}$ descritto nello spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$ in termini di una lagrangiana $\mathscr{L}$. Supponiamo di avere scelto coordinate locali naturali $t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n$ e che, nelle dette coordinate risulti,*

$$\frac{\partial \mathscr{L}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)}{\partial q^j} = 0\,, \quad \textit{per ogni scelta di } (t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n). \tag{8.2}$$

ovvero, equivalentemente, la lagrangiana $\mathscr{L}$ sia invariante sotto trasformazioni

$$(q^1, \ldots, q^{j-1}, q^j, q^{j+1}, \ldots, q^n) \mapsto (q^1, \ldots, q^{j-1}, q^j + \Delta q^j, q^{j+1}, \ldots, q^n)\,, \tag{8.3}$$

per ogni scelta di $(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ e per Δq^j. In queste ipotesi il momento coniugato a q^j, p_j definito in (8.1) si conserva nel tempo su ogni soluzione delle equazioni di Eulero-Lagrange (cioé su ogni moto del sistema). $\diamondsuit$

Il significato fisico di p_j dipende dal caso considerato. Vedremo tra poco due esempi importanti riferiti al caso in cui la lagrangiana ha struttura standard "energia cinetica più potenziale" rispetto ad un fissato sistema di riferimento inerziale $\mathscr{I}$ e le coordinate libere $q^1, \ldots, q^n$ sono scelte in modo da essere solidali con $\mathscr{I}$.
Consideriamo un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$ dotati di massa $m_1, \ldots, m_N$ rispettivamente e sottoposto a vincoli olonomi ideali oltre che a forze date da un potenziale $\mathscr{V}|_{\mathscr{I}}$ rispetto al riferimento $\mathscr{I}$. Ammettiamo che le funzioni di vincolo *non dipendano dal tempo* quando espresse in coordinate cartesiane solidali con $\mathscr{I}$. Scegliendo coordinate lagrangiane $q^1, \ldots, q^n$ *solidali* con $\mathscr{I}$ si avrà allora una lagrangiana

$$\mathscr{L}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) + \mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n)\,.$$

Notare l'assenza della coordinata t nell'energia cinetica. Dato che le coordinate sono solidali con $\mathscr{I}$, se $\mathbf{x}_i := P_i - O$ è il vettore posizione del punto i-esimo rispetto all'origine O di un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$, avremo che

$$\mathbf{x}_i = \mathbf{x}_i(q^1, \ldots, q^n)\,. \tag{8.4}$$

Questo spiega in particolare l'assenza della variabile t nell'espressione dell'energia cinetica. La coordinata t sarà assente anche da $\mathscr{V}$ nel caso in cui le forze attive siano tutte conservative, in modo tale che $\mathscr{V}|_{\mathscr{I}} = -\mathscr{U}|_{\mathscr{I}}(q^1, \ldots, q^n)$, essendo quest'ultima l'energia potenziale del sistema $\mathfrak{S}$. Nel caso in esame l'energia cinetica si potrà esplicitare come:

$$\mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \sum_{i=1}^{N} \frac{m_i}{2} \sum_{h,k=1}^{n} \frac{\partial \mathbf{x}_i}{\partial q^k} \cdot \frac{\partial \mathbf{x}_i}{\partial q^h}\, \dot{q}^k\, \dot{q}^h\,.$$

Pertanto, con semplici calcoli otteniamo:

$$p_j = \sum_{i=1}^{N} \sum_{k=1}^{n} m_i \frac{\partial \mathbf{x}_i}{\partial q^k}\, \dot{q}^k \cdot \frac{\partial \mathbf{x}_i}{\partial q^j} \tag{8.5}$$

ovvero, tenendo conto di (8.4), sui moti del sistema possiamo scrivere che:

$$p_j = \sum_{i=1}^{N} \sum_{k=1}^{n} m_i \mathbf{v}_i|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^j} \tag{8.6}$$

8.1.2 Invarianza traslazionale e conservazione dell'impulso.

Consideriamo ancora un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$ dotati di massa $m_1, \ldots, m_N$ rispettivamente e sottoposto a vincoli olonomi ideali oltre che a forze date da un potenziale $\mathscr{V}|_{\mathscr{I}}$ rispetto al riferimento $\mathscr{I}$. Ammettiamo ancora che le funzioni di vincolo *non dipendano dal tempo* quando espresse in coordinate cartesiane solidali con $\mathscr{I}$, per cui scegliendo coordinate lagrangiane $q^1, \ldots, q^n$ *solidali* con $\mathscr{I}$ si ha la lagrangiana

$$\mathscr{L}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) + \mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n) \,.$$

Possiamo allora applicare le formule (8.5) e (8.6) per il calcolo dei momenti coniugati.
Supponiamo ora in particolare che la coordinata q^j descriva *traslazioni rigide del sistema nella direzione* $\mathbf{n}$. In altre parole, per ogni $\Delta q^j \in \mathbb{R}$ deve valere:

$$\begin{cases} \mathbf{x}_i(q^1, \ldots, q^{j-1}, q^j + \Delta q^j, q^{j+1}, \ldots, q^n) = \mathbf{x}_i(q^1, q^{j-1}, q^j, q^{j+1}, \ldots, q^n) + \Delta q^1 \mathbf{n} \,, \\ \qquad\qquad \text{per ogni } i = 1, \ldots, N. \end{cases} \tag{8.7}$$

Notare che *tutti* i punti sono traslati, al variare di q^j, della *stessa* traslazione $\Delta q^j \mathbf{n}$. In questo senso la traslazione è *rigida*. Assumiamo anche che la coordinata q^j sia ciclica per la lagrangiana $\mathscr{L}|_{\mathscr{I}}$. In questo caso il momento coniugato p_j assume forma sul moto per (8.6)

$$p_j = \sum_{i=1}^{N} \sum_{k=1}^{n} m_i \mathbf{v}_i|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^j} = \sum_{k=1}^{n} m_i \mathbf{v}_i|_{\mathscr{I}} \cdot \mathbf{n} \,,$$

ovvero

$$p_j = \mathbf{P}|_{\mathscr{I}} \cdot \mathbf{n} \,. \tag{8.8}$$

Abbiamo di conseguenza la seguente proposizione.

Proposizione 8.2. *Si consideri un sistema di punti materiali $\mathfrak{S}$ soggetto a vincoli olonomi ideali, indipendenti dal tempo nel riferimento $\mathscr{I}$ nel quale le forze attive sono individuate da una funzione potenziale $\mathscr{V}|_{\mathscr{I}}$, e si usino coordinate libere $q^1, \ldots, q^n$ solidali con $\mathscr{I}$. Vale quanto segue.*
(a) *Se la coordinata q^j descrive traslazioni rigide in $\mathscr{I}$ di $\mathfrak{S}$ nella direzione $\mathbf{n}$, il momento coniugato p_j coincide con la componente lungo $\mathbf{n}$ dell'impulso totale (rispetto a $\mathscr{I}$) su ogni moto di $\mathfrak{S}$ del sistema.*
(b) *Se la lagrangiana $\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}} + \mathscr{V}|_{\mathscr{I}}$ è invariante sotto traslazioni rigide del sistema nella direzione $\mathbf{n}$, cioé q^j è ciclica ,allora l'impulso totale di $\mathfrak{S}$ (rispetto a $\mathscr{I}$) lungo la direzione $\mathbf{n}$ si conserva su ogni moto.* $\diamondsuit$

Esempi 8.1.
1. L'esempio più semplice è quello di un sistema $\mathfrak{S}$ di N punti materiali $P_1, \cdots, P_N$, con masse $m_1, \ldots, m_n$, assumendo che $\mathfrak{S}$ *non* sia sottoposto a vincoli e che le forze (che sono tutte attive)

siano tutte interne e conservative. In tal modo l'energia potenziale dipende solo dalle differenze di posizione tra i punti $P_i - P_j$ con $i, j = 1, \ldots, N$ e $i \neq j$. In questo caso possiamo descrivere $\mathfrak{S}$ in un riferimento inerziale $\mathscr{I}$, usando come coordinate libere le $3N$ coordinate cartesiane dei punti di $\mathfrak{S}$ rispetto ad un sistema di coordinate cartesiane ortogonali solidali con $\mathscr{I}$ di assi $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ e origine O. Successivamente cambiamo coordinate libere tramite una trasformazione lineare, possiamo fare in modo che le prime 3 coordinate libere q^1, q^2, q^3 descrivano traslazioni rigide del sistema $\mathfrak{S}$ rispettivamente lungo gli assi delle coordinate $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ solidali con $\mathscr{I}$. Le 3 coordinate q^1, q^2, q^3 suddette possono essere fissate come le componenti del vettore centro di massa G del sistema, le rimanenti coordinate libere possono per esempio essere fissate come le $n-3$ componenti degli $N-1$ vettori $P_i - P_1$ con $1 < i \leq N$. Un sistema di coordinate alternativo è quello delle cosiddette *coordinate di Jacobi* sulle quali non ci soffermiamo.
Dato che sia l'energia cinetica non è funzione delle coordinate non puntate e l'energia potenziale dipende dalle differenze $P_i - P_j$, che sono invarianti sotto traslazioni rigide dell'intero sistema, concludiamo che la lagrangiana è invariante per traslazioni lungo $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Dalla proposizione precedente abbiamo immediatamente che si conservano le tre componenti dell'impulso totale del sistema.
2. Consideriamo un sistema $\mathfrak{S}$ costituito da due punti materiali P_1 e P_2 di masse m_1 e m_2, vincolati a stare su una retta liscia r, in quiete in un riferimento inerziale $\mathscr{I}$, con r passante per l'origine O di un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ e con versore tangente $\mathbf{n}$. I punti sono interagenti l'uno con l'altro attraverso una coppia di forze conservative di energia potenziale $\mathscr{U} = \mathscr{U}(||P_1 - P_2||)$. In questo caso, a differenza dell'esempio precedente, sono presenti le reazioni vincolari (ideali) dovute al vincolo (4 equazioni funzionalmente indipendenti) ed il numero dei gradi di libertà del sistema si riduce a 2. Usando come coordinate libere $q^1 := (s_1 + s_2)/2$ e $q^2 := (s_1 - s_2)/2$, dove s_i è la coordinata di P_i su r presa a partire da O con segno positivo nella direzione $\mathbf{n}$, si vede subito che la coordinata q^1 genera traslazioni rigide del sistema. Ciò segue immediatamente dalle relazioni: $s_1 = q^1 + q^2$ e $s_2 = q^1 - q^2$. La lagrangiana $\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$ non dipenderà, per costruzione, dalla coordinata q^1 (lo si dimostri!). Concludiamo pertanto che, in base alla proposizione precedente, si conserva l'impulso totale del sistema nella direzione $\mathbf{n}$.

8.1.3 Invarianza rotazionale e conservazione del momento angolare.

Consideriamo ancora un sistema $\mathfrak{S}$ di N punti materiali $P_1, \ldots, P_N$ dotati di massa $m_1, \ldots, m_N$ rispettivamente e sottoposto a vincoli olonomi ideali oltre che a forze date da un potenziale $\mathscr{V}|_{\mathscr{I}}$ rispetto al riferimento $\mathscr{I}$. Ammettiamo ancora che le funzioni di vincolo *non dipendano dal tempo* quando espresse in coordinate cartesiane solidali con $\mathscr{I}$, per cui scegliendo coordinate lagrangiane $q^1, \ldots, q^n$ *solidali* con $\mathscr{I}$ si ha la lagrangiana

$$\mathscr{L}|_{\mathscr{I}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) + \mathscr{V}|_{\mathscr{I}}(t, q^1, \ldots, q^n) .$$

Possiamo allora applicare le formule (8.5) e (8.6) per il calcolo dei momenti coniugati.
Supponiamo ora in particolare che la coordinata q^j descriva *rotazioni rigide del sistema attorno*

all'asse $\mathbf{n}$ *rispetto all'origine* O. In altre parole, per ogni $\Delta q^j \in \mathbb{R}$ deve valere:

$$\begin{cases} \mathbf{x}_i(q^1, \ldots, q^{j-1}, q^j + \Delta q^j, q^{j+1}, \ldots, q^n) = R_{\mathbf{n},\Delta q^1}\mathbf{x}_i(q^1, q^{j-1}, q^j, q^{j+1}, \ldots, q^n)\,, \\ \qquad\qquad\qquad\qquad \text{per ogni } i = 1, \ldots, N. \end{cases} \tag{8.9}$$

$R_{\mathbf{n},\Delta q^1} : V^3 \to V^3$ è l'operatore che ruota il vettore su cui è applicato di un angolo Δq^1 positivamente attorno al versore $\mathbf{n}$. Notare che *tutti* i vettori posizione $\mathbf{x}_i = P_i - O$ sono ruotati, al variare di q^j, della *stessa* rotazione $R_{\mathbf{n},\Delta q^1}$. In questo senso la rotazione è *rigida*. Assumiamo anche che la coordinata q^j sia ciclica per la lagrangiana $\mathscr{L}|_{\mathscr{I}}$. In questo caso il momento coniugato p_j assume forma sul moto per (8.6)

$$p_j = \sum_{i=1}^{N}\sum_{k=1}^{n} m_i \mathbf{v}_i|_{\mathscr{I}} \cdot \frac{\partial \mathbf{x}_i}{\partial q^j} = \sum_{k=1}^{n} m_i \mathbf{v}_i|_{\mathscr{I}} \cdot \mathbf{n} \wedge \mathbf{x}_i\,,$$

dove abbiamo usato il fatto che:

$$\frac{d}{d\theta} R_{\mathbf{n},\theta}\mathbf{x} = \mathbf{n} \wedge (R_{\mathbf{n},\theta}\mathbf{x})\,.$$

Tenendo conto del fatto che $\mathbf{v}_i|_{\mathscr{I}} \cdot \mathbf{n} \wedge \mathbf{x}_i = \mathbf{n} \cdot \mathbf{x}_i \wedge m_i \mathbf{v}_i$, possiamo riscrivere p_j ottenuto sopra come:

$$p_j = \mathbf{n} \cdot \sum_{k=1}^{n} \mathbf{x}_i \wedge m_i \mathbf{v}_i\,,$$

ovvero

$$p_j = \Gamma_O|_{\mathscr{I}} \cdot \mathbf{n}\,. \tag{8.10}$$

dove $\Gamma_O|_{\mathscr{I}}$ è il momento angolare totale del sistema $\mathfrak{S}$ rispetto al polo O valutato nel riferimento $\mathscr{I}$. Abbiamo di conseguenza la seguente proposizione.

Proposizione 8.3. *Si consideri un sistema di punti materiali* $\mathfrak{S}$ *soggetto a vincoli olonomi ideali, indipendenti dal tempo nel riferimento* $\mathscr{I}$ *nel quale le forze attive sono individuate da una funzione potenziale* $\mathscr{V}|_{\mathscr{I}}$*, e si usino coordinate libere* $q^1, \ldots, q^n$ *solidali con* $\mathscr{I}$*. Vale quanto segue.*
(a) *Se la coordinata* q^j *descrive rotazioni rigide in* $\mathscr{I}$ *di* $\mathfrak{S}$ *attorno all'asse* $\mathbf{n}$ *e rispetto al punto* O*, il momento coniugato* p_j *coincide con la componente lungo* $\mathbf{n}$ *del momento angolare totale del sistema* $\mathfrak{S}$ *rispetto al polo* O *valutato nel riferimento* $\mathscr{I}$*.*
(b) *Se la lagrangiana* $\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}} + \mathscr{V}|_{\mathscr{I}}$ *è invariante sotto rotazioni rigide del sistema attorno all'asse* $\mathbf{n}$ *e rispetto al punto* O*, cioé la coordinata* q^j *è ciclica allora la componente lungo* $\mathbf{n}$ *del momento angolare totale del sistema* $\mathfrak{S}$ *rispetto al polo* O *valutato nel riferimento* $\mathscr{I}$ *si conserva su ogni moto.* $\diamondsuit$

Esempi 8.2.
1. L'esempio più semplice è quello di un sistema $\mathfrak{S}$ di N punti materiali $P_1, \cdots, P_N$, con masse

$m_1, \ldots, m_n$, assumendo che $\mathfrak{S}$ *non* sia sottoposto a vincoli e che le forze (che sono tutte attive) siano tutte interne e conservative. In tal modo l'energia potenziale dipende solo dalle differenze di posizione tra i punti $P_i - P_j$ con $i, j = 1, \ldots, N$ e $i \neq j$. In questo caso possiamo descrivere $\mathfrak{S}$ in un riferimento inerziale $\mathscr{I}$, usando come coordinate libere le $3N$ coordinate cilindriche φ_i, r_i, z_i di ciascuno degli N punti di $\mathfrak{S}$, rispetto ad un comune sistema di assi ortogonali di origine O solidali con $\mathscr{I}$. Successivamente cambiamo coordinate libere tramite una trasformazione lineare degli angoli dei singoli punti, usando come nuove coordinate, q^1 data dalla coordinata angolare del centro di massa del sistema, le rimanenti coordinate di tipo angolare sono invece le differenze $\varphi_i - \varphi_1$ per $1 < \varphi_i \leq N$. Le rimanenti coordinate libere rimangono le stesse di prima (quelle di tipo r_i e z_i). In questo modo la prima coordinata q^1 descrive rotazioni rigide del sistema $\mathfrak{S}$ attorno all'asse z. Dato che sia l'energia cinetica non è funzione delle coordinate non puntate e l'energia potenziale dipende dalle norme delle differenze $P_i - P_j$, che sono invarianti sotto rotazioni rigide dell'intero sistema, concludiamo che la lagrangiana è invariante per rotazioni attorno all'asse z e rispetto all'origine O. Dalla proposizione precedente abbiamo immediatamente che si conserva la componente lungo l'asse z del momento angolare totale del sistema riferito al polo O e valutato nel riferimento $\mathscr{I}$. Si osservi che in realtà ciò vale rispetto *ad ogni asse ed ogni origine* O (entrambi in quiete in $\mathscr{I}$), in quanto la scelta iniziale di dove posizionare l'asse z è del tutto arbitratia.

2. Consideriamo un sistema $\mathfrak{S}$ costituito da due punti materiali P_1 e P_2 di masse m_1 e m_2, vincolati a stare su una circonferenza liscia C, in quiete in un riferimento inerziale $\mathscr{I}$, con C centrata nell'origine O di un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$ e normale all'asse z. I punti sono interagenti l'uno con l'altro attraverso una coppia di forze conservative di energia potenziale $\mathscr{U} = \mathscr{U}(||P_1 - P_2||)$. In questo caso, a differenza dell'esempio precedente, sono presenti le reazioni vincolari (ideali) dovute al vincolo (4 equazioni funzionalmente indipendenti) ed il numero dei gradi di libertà del sistema si riduce a 2. Usando come coordinate libere $q^1 := (\varphi_1 + \varphi_2)/2$ e $q^2 := (\varphi_1 - \varphi_2)/2$, dove φ_i è la coordinata angolare di P_i riferita a coordinate polari piane nel piano $z = 0$, si vede subito che la coordinata q^1 genera traslazioni rigide del sistema. Ciò segue immediatamente dalle relazioni: $\varphi_1 = q^1 + q^2$ e $\varphi_2 = q^1 - q^2$. La lagrangiana $\mathscr{L}|_{\mathscr{I}} = \mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$ non dipenderà , per costruzione, dalla coordinata q^1 (lo si dimostri!). Concludiamo pertanto che, in base alla proposizione precedente, si conserva su ogni moto di $\mathfrak{S}$ la componente lungo $\mathbf{e}_z$ del momento angolare totale del sistema $\mathfrak{S}$ rispetto al polo O valutato nel riferimento $\mathscr{I}$.

8.2 Il legame tra simmetrie e leggi di conservazione: il teorema di Emmy Nöther.

In questa sezione formuliamo e dimostriamo il teorema di Emmy Nöther (nel caso di un sistema meccanico) che si può dire essere uno dei più importanti teoremi della fisica matematica e teorica. Il teorema suddetto[1] illustra lo strettissimo legame che esiste tra simmetria di un sistema

[1] La formulazione generale del teorema, valida per sistemi continui inclusi i campi, apparve in "Invariante Variationsprobleme" Nachr.d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse, 235-257 (1918).

lagrangiano sotto un gruppo di trasformazioni e esistenza di grandezze conservate sui moti del sistema. Il teorema può essere formulato nel caso di sistemi fisici del tutto generali, anche sistemi continui e campi (anche in versioni relativistiche), purché descritti in formulazione lagrangiana.

8.2.1 Trasformazioni su $j^1(\mathbb{V}^{n+1})$.

Consideriamo un sistema meccanico $\mathfrak{S}$ descritto sullo spaziotempo delle configurazioni $\mathbb{V}^{n+1}$, fissiamo un sistema di coordinate locali naturali $t, q^1, \ldots, q^n \in (a,b) \times \mathcal{V}$ definite in un aperto di $\mathbb{V}^{n+1}$. Ricordiamo che le coordinate $q^1, \ldots, q^n$ sono coordinate locali su ciascuno spazio delle configurazioni $\mathbb{Q}_t$ ad ogni tempo t. Consideriamo infine una trasformazione attiva in $(t_1, t_2) \times \mathcal{U} \subset (a,b) \times \mathcal{V}$, che definisca un diffeomorfismo tra i punti di coordinate $(t, q^1, \ldots, q^n)$ nei punti di coordinate $(t', q'^1, \ldots, q'^n) \in (t'_1, t'_1) \times \mathcal{U}' \subset (a,b) \times \mathcal{V}$ e che ha la struttura:

$$\begin{cases} t' = t\,, \\ q'^k = q'^k(t, q^1, \ldots, q^n) \quad \text{per } k = 1, \ldots, n. \end{cases} \tag{8.11}$$

In altre parole, per ogni istante t, *i punti di ogni $\mathbb{Q}_t$ vengono spostati, ma sono mandati biettivamente e bidifferenziabilmente in punti dello stesso spazio delle configurazioni $\mathbb{Q}_t$ al tempo t.* Dato che lo spazio delle configurazioni al tempo t è un sottoinsieme, individuato dai vincoli, dello spazio assoluto al tempo t (più precisamente del prodotto cartesiano degli spazi assoluti Σ_t al tempo t dei punti materiali di cui è costituito $\mathfrak{S}$), queste trasformazioni non sono altro che trasformazioni locali differenziabili con inversa differenziabile delle configurazioni di $\mathfrak{S}$ che: (1) rispettano i vincoli a cui è sottoposto il sistema e che (2) agiscono nello spazio assoluto ad ogni istante, trasformando punti dello spazio assoluto all'istante t in punti dello spazio assoluto allo stesso istante t. Quello che queste trasformazioni *non fanno* è mandare punti che esistono nello spazio assoluto Σ_t all'istante t in un altro spazio assoluto $\Sigma_{t'}$ con $t \neq t'$.
Una trasformazioni locale in $\mathbb{V}^{n+1}$ di questo tipo è detta **preservare le fibre** di $\mathbb{V}^{n+1}$. Ricordiamo che le fibre di $\mathbb{V}^{n+1}$ sono proprio gli spazi $\mathbb{Q}_t$ per ogni t.

Esempi 8.3.
(**1**). Se $\mathfrak{S}$, è un singolo punto materiale e q^1, q^2, q^3 sono le coordinate cartesiane del punto riferite ad un riferimento $\mathscr{I}$, una trasformazione del tipo descritto sopra è quella che, ad ogni istante fissato, ruota la posizione del punto attorno ad un fissato asse passante per un punto fissato (entrambi indipendenti dal tempo) secondo un angolo ϵ fissato.
(**2**). L'esempio precedente si generalizza immediatamente al caso di due punti P, Q connessi da un'asta rigida di lunghezza fissata descritti con coordinate libere $q^1, \ldots, q^5$, dove q^1, q^2, q^3 sono le coordinate di P in un sistema di coordinate cartesiane ortonormali di assi $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ e origine O solidali con un fissato riferimento $\mathscr{I}$, mentre q^4, q^5 sono angoli polari che individuano Q rispetto ad un sistema di assi ortonormali uscenti da P paralleli agli assi $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. Una trasformazione

Una traduzione in inglese si può trovare su http://arxiv.org/pdf/physics/0503066.

che preserva le fibre di $\mathbb{V}^{5+1}$ è ora quella che ruota P e Q attorno ad un fissato asse $\mathbf{n}$ passante per O (indipendenti dal tempo) secondo un angolo ϵ, *preservando la distanza reciproca dei due punti.*

Una trasformazione che preserva le fibre si estende immediatamente alle coordinate puntate, cioé dallo spaziotempo delle configurazioni $\mathbb{V}^{n+1}$ allo spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$, nel modo che segue:

$$\begin{cases} t' = t\,, \\ q'^k = q'^k(t, q^1, \ldots, q^n) \quad \text{per } k = 1, \ldots, n. \\ \dot{q}'^k = \dfrac{\partial q'^k}{\partial t}(t, q^1, \ldots, q^n) + \displaystyle\sum_{j=1}^{n} \dfrac{\partial q'^k}{\partial q^j}(t, q^1, \ldots, q^n)\, \dot{q}^j \quad \text{per } k = 1, \ldots, n, \end{cases} \tag{8.12}$$

dove $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) \in (t_1, t_2) \times \mathcal{U} \times \mathbb{R}^n$.
Nello scrivere l'ultima riga abbiamo supposto euristicamente che le $\dot{q}^k$ fossero le derivate temporali delle q^k secondo una curva fissata in $\mathbb{V}^{n+1}$ e parametrizzata nel tempo. In realtà tale curva non esiste dato che le $\dot{q}^k$ sono coordinate *indipendenti* dalle q^k. Tuttavia l'interpretazione data è corretta per ogni una curva fissata in $\mathbb{V}^{n+1}$ e parametrizzata nel tempo, pur di interpretare le $\dot{q}^k$ come le derivate temporali delle q^k. Questo accade in particolare sulle soluzioni delle equazioni di Eulero-Lagrange.
La trasformazione (8.12) si dice **indotta** dalla trasformazione che preserva le fibre (8.11).

8.2.2 Il teorema di Nöther in forma locale elementare.

Consideriamo un sistema fisico di punti materiali $\mathfrak{S}$ descritto in $j^1(\mathbb{V}^{n+1})$ e, in riferimento a coordinate locali naturali $(t, q^1, \ldots, q^n) \in (a, b) \times \mathcal{V}$ definite su un aperto di $\mathbb{V}^{n+1}$, supponiamo di avere una *classe* di diffeomorfismi locali $\mathcal{T}$ che preservano le fibre di $\mathbb{V}^{n+1}$ dipendenti dal parametro $\epsilon \in (-\alpha, \alpha)$ con $\alpha > 0$ fissato:

$$\begin{cases} t' = t\,, \\ q'^k = q'^k(\epsilon, t, q^1, \ldots, q^n) \quad \text{per } k = 1, \ldots, n. \end{cases} \tag{8.13}$$

Assumiamo che le funzioni a secondo membro siano differenziabili congiuntamente in tutte le variabili (inclusa ϵ) e che la trasformazione si riduca all'identità per $\epsilon = 0$:

$$q'^k(0, t, q^1, \ldots, q^n) = q^k \quad \text{per } k = 1, \ldots, n, \tag{8.14}$$

ed infine che valga la proprietà , per $k = 1, \ldots, n$:

$$q'^k\left(\epsilon', t, q'^1(\epsilon, t, q^1, \ldots, q^n), \ldots, q'^n(\epsilon, t, q^1, \ldots, q^n)\right) = q'^k(\epsilon' + \epsilon, t, q^1, \ldots, q^n) \tag{8.15}$$

quando entrambi i membri siano definiti. In questo caso diremo che $\mathcal{T}$ è un **gruppo ad un parametro di diffeomorfismi locali** su $\mathbb{V}^{n+1}$ che preservano le fibre (vedi la Sezione 3.5.3).

Esempi 8.4. In riferimento all'esempio (2) in esempi 8.3, un esempio si ottiene considerando la classe di trasformazioni che si ottiene al variare dell'angolo ϵ in un $(-\pi, \pi)$. Notare che per $\epsilon = 0$ si ha la rotazione nulla e pertanto (8.14) è verificata. Inoltre la l'azione successiva di due rotazioni, la prima di un angolo ϵ e la seconda di un angolo ϵ', coincide con l'unica rotazione di un angolo $\epsilon + \epsilon'$. Quindi anche (8.15) è valida.

Definizione 8.1. (**Sistema lagrangiano invariante sotto un gruppo di trasformazioni**.) Diremo che un sistema di punti materiali $\mathfrak{S}$ descritto su $\mathbb{V}^{n+1}$ è **invariante** sotto il gruppo ad un parametro di diffeomorfismi locali che preservano le fibre di $\mathbb{V}^{n+1}$, $\mathcal{T}$, dato da (8.13), se c'è una lagrangiana $\mathscr{L} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ che descrive la dinamica del sistema $\mathfrak{S}$ tale che, nel sistema di coordinate naturali usate per definire $\mathcal{T}$ vale:

$$\left.\frac{\partial \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)}{\partial \epsilon}\right|_{\epsilon=0} = 0\,, \quad \text{per ogni } t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n \tag{8.16}$$

dove:

$$\dot{q}'^k = \frac{\partial q'^k}{\partial t}(\epsilon, t, q^1, \ldots, q^n) + \sum_{j=1}^{n} \frac{\partial q'^k}{\partial q^j}(\epsilon, t, q^1, \ldots, q^n)\, \dot{q}^j \quad \text{per } k = 1, \ldots, n. \tag{8.17}$$

In questo caso $\mathcal{T}$ è detto **gruppo (di trasformazioni di) simmetria** del sistema $\mathfrak{S}$.
Diremo invece che $\mathfrak{S}$ descritto su $\mathbb{V}^{n+1}$ è **debolmente invariante** sotto $\mathcal{T}$ se c'è una lagrangiana $\mathscr{L} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ che descrive la dinamica del sistema $\mathfrak{S}$ per la quale vale

$$\left.\frac{\partial \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)}{\partial \epsilon}\right|_{\epsilon=0} = G(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n), \quad \text{per ogni } t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n \tag{8.18}$$

essendo $t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n$ il sistema di coordinate locali su $j^1(\mathbb{V}^{n+1})$ usato per definire $\mathcal{T}$, dove vale (8.17) e dove infine, la funzione G ha la struttura di "derivata totale":

$$G(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \frac{\partial g}{\partial t}(t, q^1, \ldots, q^n) + \sum_{k=1}^{n} \frac{\partial g}{\partial q^k}(t, q^1, \ldots, q^n)\, \dot{q}^k \tag{8.19}$$

per $g = g(t, q^1, \ldots, q^n)$ differenziabile definita localmente su $\mathbb{V}^{n+1}$.
In questo caso $\mathcal{T}$ è detto **gruppo (di trasformazioni di) simmetria debole** del sistema $\mathfrak{S}$. $\diamondsuit$

Possiamo ora enunciare e provare il *teorema di Nöther*.

Teorema 8.1. (**Teorema di Nöther**.) *Si consideri un sistema di punti materiali $\mathfrak{S}$ descritto su $\mathbb{V}^{n+1}$ ed invariante sotto il gruppo ad un parametro di diffeomorfismi locali che preservano le*

fibre di $\mathbb{V}^{n+1}$, $\mathcal{T}$, dato da (8.13). Se $\mathscr{L}$ è la lagrangiana di $\mathfrak{S}$ che soddisfa (8.16), sui moti del sistema (fino a quando sono confinati nel domino delle coordinate usate per descrivere l'azione locale di $\mathcal{T}$ su $j^1(\mathbb{V}^{n+1})$) si conserva nel tempo la funzione:

$$I_{\mathcal{T}}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} (t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \,. \tag{8.20}$$

Se invece $\mathfrak{S}$ è debolmente invariante sotto $\mathcal{T}$ e $\mathscr{L}$ e g sono le funzioni che soddisfano (8.18) e (8.19), allora sui moti del sistema (fino a quando sono conciliati nel domino delle coordinate usate per descrivere l'azione locale di $\mathcal{T}$ su $j^1(\mathbb{V}^{n+1})$) si conserva nel tempo la funzione:

$$I_{\mathcal{T},g}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} (t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) - g(t, q^1, \ldots, q^n) \,. \tag{8.21}$$

$\diamondsuit$

Dimostrazione. Se esplicitiamo il primo membro di (8.16) abbiamo che esso vale:

$$\sum_k \frac{\partial \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)}{\partial q'^k} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} + \sum_k \frac{\partial \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)}{\partial \dot{q}'^k} \frac{\partial \dot{q}'^k}{\partial \epsilon}\bigg|_{\epsilon=0}$$

che, tenendo conto di (8.14), può riscriversi come:

$$\sum_k \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial q^k} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} + \sum_k \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial \dot{q}^k} \frac{\partial \dot{q}'^k}{\partial \epsilon}\bigg|_{\epsilon=0} \,.$$

Ora consideriamo un moto del sistema $q^k = q^k(t)$, $\dot{q}^k(t) = dq^k/dt(t)$. Da (8.17) si ha facilmente che, sviluppando separatamente i du membri e confrontandoli:

$$\frac{\partial \dot{q}'^k}{\partial \epsilon}(t) = \frac{d}{dt} \frac{\partial q'^k(\epsilon, t, q^1(t), \ldots, q^n(t))}{\partial \epsilon}\bigg|_{\epsilon=0} \,.$$

Inserendo il risultato nell'espressione trovata sopra per $\frac{\partial}{\partial \epsilon}|_{\epsilon=0} \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)$, dove ora *assumiamo esplicitamente che le variabili della lagrangiana siano valutate sul moto considerato*, abbiamo che:

$$0 = \frac{\partial}{\partial \epsilon}|_{\epsilon=0} \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n) =$$

$$\sum_k \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial q^k} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} + \sum_k \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial \dot{q}^k} \frac{d}{dt} \frac{\partial q'^k(\epsilon, t, q^1(t), \ldots, q^n(t))}{\partial \epsilon}\bigg|_{\epsilon=0} \,.$$

Abbiamo in definitiva trovato che:

$$\sum_k \frac{\partial \mathscr{L}}{\partial q^k} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} + \sum_k \frac{\partial \mathscr{L}}{\partial \dot{q}^k} \frac{d}{dt} \frac{\partial q'^k}{\partial \epsilon}\bigg|_{\epsilon=0} = 0 \,.$$

Dato che tutto è valutato su una soluzione delle equazioni di Eulero-Lagrange, possiamo riscrivere l'identità trovata come:

$$\sum_k \left(\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial\, \dot{q}^k}\right) \left.\frac{\partial q'^k}{\partial \epsilon}\right|_{\epsilon=0} + \sum_k \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k}\frac{d}{dt} \left.\frac{\partial q'^k}{\partial \epsilon}\right|_{\epsilon=0} = 0\,,$$

che può essere infine riscritta come:

$$\frac{d}{dt}\left(\sum_{k=1}^n \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k} \left.\frac{\partial q'^k}{\partial \epsilon}\right|_{\epsilon=0}\right) = 0\,,$$

dove il primo membro, prima di eseguire la derivata totale nel tempo, *è valutato su una qualsiasi soluzione delle equazioni di Eulero-Lagrange*. Questa è la tesi nel primo caso. Per quanto riguarda il secondo caso la dimostrazione è essenzialmente la stessa. Lavorando come nel caso precedente, la condizione di invarianza debole (8.18) si riscrive, su ogni fissata soluzione delle equazioni di Eulero-Lagrange:

$$\frac{d}{dt}\left(\sum_{k=1}^n \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k} \left.\frac{\partial q'^k}{\partial \epsilon}\right|_{\epsilon=0}\right) = \frac{d}{dt} g(t, q^1(t), \ldots, q^n(t))\,,$$

da cui segue immediatamente la tesi. □

Esempi 8.5.
1. Il caso di lagrangiana con coordinate cicliche ricade come sottocaso particolare del teorema di Nëther: l'integrale primo di Nöther coincide con il momento coniugato conservato. Infatti se $\mathscr{L} = \mathscr{L}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ è tale che $\partial\mathscr{L}/\partial q^j = 0$, si consideri il gruppo il gruppo locale ad un parametro di trasformazioni che preservano le fibre di $\mathbb{V}^{n+1}$, $\mathcal{T}$ dato dalle trasformazioni: $t' = t$, $q'^j = q^j + \epsilon$ e $q'^k = q^k$ se $k \neq j$. Quindi (8.17) fornisce immediatamente:

$$\dot{q}'^k = \dot{q}^k \quad \text{per } k = 1, \ldots, n. \tag{8.22}$$

Banalmente la lagrangiana soddisfa (8.40) sotto l'azione di $\mathcal{T}$ non dipendendo esplicitamente da q^j. Di conseguenza soddisfa anche (8.16). L'integrale primo ottenuto dal teorema di Nöther è quindi:

$$I_{\mathcal{T}} = \sum_k \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k} \left.\frac{\partial q'^k}{\partial \epsilon}\right|_{\epsilon=0} = \sum_k \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k}\delta_{kj} = \frac{\partial \mathscr{L}}{\partial\, \dot{q}^j} = p_j\,.$$

Conseguentemente, per sistemi invarianti sotto traslazioni rigide ovvero sotto rotazioni rigide con lagrangiana del tipo $\mathscr{T} + \mathscr{V}$, gli integrali primi conservati in virtù del teorema di Nöther sono le corrispondenti componenti dell'impulso totale e del momento angolare totale come visto nella sezione precedente negli esempi 8.3 e 8.4.
2. Un esempio molto più interessante è il seguente. Consideriamo un sistema $\mathfrak{S}$ di N punti materiali $P_1, \cdots, P_N$, con messe $m_1, \ldots, m_n$, assumendo che $\mathfrak{S}$ *non* sia sottoposto a vincoli

e che le forze (che sono tutte attive) siano tutte interne e conservative. In tal modo l'energia potenziale dipende solo dalle differenze di posizione tra i punti $P_i - P_j$ con $i, j = 1, \ldots, N$ e $i \neq j$. In questo caso possiamo descrivere $\mathfrak{S}$ in un riferimento inerziale $\mathscr{I}$, usando come coordinate libere le $3N$ coordinate dei punti P_i rispetto ad un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$, di origine O e assi $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. Consideriamo l'azione di un sottogruppo ad un parametro $\mathcal{T}$ del *sottogruppo delle trasformazioni pure di Galileo* sui punti del sistema: per $i = 1, \ldots, N$, se $\mathbf{x}_i := P_i - O$,

$$\mathbf{x}_i \to \mathbf{x}'_i = \mathbf{x}_i + \epsilon t \mathbf{n} \,, \quad \text{con } \epsilon \in \mathbb{R} \,. \tag{8.23}$$

Il versore $\mathbf{n}$ è fissato una volta per tutte. Si osservi che la trasformazione dipende parametricamente dal tempo. L'azione sulle coordinate puntate è quindi:

$$\dot{\mathbf{x}}_i \to \dot{\mathbf{x}}'_i = \dot{\mathbf{x}}_i + \epsilon \mathbf{n} \,, \quad \text{con } \epsilon \in \mathbb{R} \,. \tag{8.24}$$

Sotto le trasformazioni dette l'energia potenziale rimane invariata dato che dipende dalle differenze di posizione dei punti. L'energia cinetica invece si trasforma come:

$$\mathscr{T} \to \mathscr{T}' = \mathscr{T} + \epsilon \sum_{i=1}^{N} m_i \dot{\mathbf{x}}_i \cdot \mathbf{n} + \frac{\epsilon^2}{2} \sum_{i=1}^{N} m_i \,.$$

Pertanto:

$$\mathscr{L}' = \mathscr{L} + \epsilon \sum_{i=1}^{N} m_i \dot{\mathbf{x}}_i \cdot \mathbf{n} + \frac{\epsilon^2}{2} \sum_{i=1}^{N} m_i$$

e quindi:

$$\left. \frac{\partial \mathscr{L}'}{\partial \epsilon} \right|_{\epsilon=0} = \mathbf{n} \cdot \sum_{i=1}^{N} m_i \dot{\mathbf{x}}_i \,.$$

Il secondo membro è una derivata totale della funzione:

$$g(t, \mathbf{x}) := \mathbf{n} \cdot \sum_{i=1}^{N} m_i \mathbf{x}_i \,.$$

Abbiamo provato che il sistema è debolmente invariante sotto l'azione di $\mathcal{T}$. Per il teorema di Nöther si ha l'integrale primo dipendente esplicitamente dal tempo:

$$I_{\mathcal{T}}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1 \ldots, \dot{\mathbf{x}}_N) := \mathbf{n} \cdot t \sum_{i=1}^{N} m_i \dot{\mathbf{x}}_i - \mathbf{n} \cdot \sum_{i=1}^{N} m_i \mathbf{x}_i \,. \tag{8.25}$$

Su ogni moto, l'integrale primo sopra ottenuto può essere trascritto come:

$$I_{\mathcal{T}}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1 \ldots, \dot{\mathbf{x}}_N) := \mathbf{n} \cdot \left(t \mathbf{P}_{\mathscr{I}} - \sum_{i=1}^{N} m_i \mathbf{x}_i \right) \,.$$

L'equazione di conservazione $dI_{\mathcal{T}}/dt = 0$, tenendo conto che l'impulso totale $\mathbf{P}_{\mathscr{I}}$ si conserva nelle ipotesi fatte, fornisce:

$$\mathbf{n} \cdot \left(\mathbf{P}|_{\mathscr{I}} - \frac{d}{dt} \sum_{i=1}^{N} m_i \mathbf{x}_i(t) \right) = 0 \, ,$$

che deve valere per ogni direzione $\mathbf{n}$, dato che il versore $\mathbf{n}$ iniziale era del tutto arbitrario. In definitva abbiamo ottenuto, usando le notazioni della Sezione 4.1:

$$\mathbf{P}|_{\mathscr{I}} = \frac{d}{dt} M \mathbf{v}_G|_{\mathscr{I}} \, ,$$

Questo è il teorema del centro di massa già trovato nella Sezione 4.1 espresso dall'equazione (4.1).

8.2.3 Invarianza dell'integrale primo di Nöther per trasformazione di coordinate.

Vogliamo qui sottolineare come l'integrale primo associato ad un gruppo ad un parametro di diffeomorfismi locali $\mathcal{T}$ *non* dipenda dalle coordinate naturali usate per rappresentare $\mathcal{T}$. Consideriamo il gruppo ad un parametro $\mathcal{T}$. Per descriverlo consideriamo una carta locale (U, ψ) su $\mathbb{V}^{n+1}$ con $\psi : U \ni p \mapsto (t, q^1, \ldots, q^n)$, dove $U \subset \mathbb{V}^{n+1}$ è un aperto. In queste coordinate il gruppo $\mathcal{T}$ è descritto dalle trasformazioni attive, dipendenti dal parametro $\epsilon \in (-\delta, \delta)$: $T_\epsilon : (t, q^1, \ldots, q^n) \mapsto (t', q'^1, \ldots, q'^n)$ espresse in coordinate di (U, ψ) come:

$$\begin{cases} t' = t \, , \\ \\ q'^k = q'^k(\epsilon, t, q^1, \ldots, q^n) \quad \text{per } k = 1, \ldots, n. \end{cases} \tag{8.26}$$

Facciamo un cambiamento di coordinate passando alla carta locale (V, ψ) con $\phi : V \ni q \mapsto (\hat{t}, \hat{q^1}, \ldots, \hat{q^n})$ con $V \subset \mathbb{V}^{n+1}$ un altro aperto tale che $V \cap U \neq \emptyset$. La legge di trasformazione di coordinate, essendo i due sistemi di coordinate adattati alle fibre di $\mathbb{V}^{n+1}$, avrà al solito la forma: $\hat{t} = t + c$ e $\hat{q}^k = \hat{q}^k(t, q^1, \ldots, q^n)$ per $k = 1, \ldots, n$. Il gruppo $\mathcal{T}$ può essere espresso su $V \cap U$ nelle nuove coordinate, attraverso la classe di trasformazioni date in coordinate $\hat{t}, \hat{q^1}, \ldots, \hat{q^n}$ da:

$$\hat{T}_\epsilon := \psi \circ (\phi^{-1} \circ T_\epsilon \circ \phi) \circ \psi^{-1}$$

corrispondenti, in coordinate, a

$$\begin{cases} \hat{t}' = \hat{t} \, , \\ \\ \hat{q}'^k = \hat{q}'^k \left(\epsilon, \hat{t}, q^1 \left(\epsilon, t(\hat{t}, \hat{q^1}, \ldots, \hat{q^n}), q^1(\hat{t}, \hat{q^1}, \ldots, \hat{q^n}), \ldots q^n(\hat{t}, \hat{q^1}, \ldots, \hat{q^n}) \right), \ldots, \right. \\ \quad \left. \ldots q^n \left(\epsilon, t(\hat{t}, \hat{q^1}, \ldots, \hat{q^n}), q^1(\hat{t}, \hat{q^1}, \ldots, \hat{q^n}), \ldots q^n(\hat{t}, \hat{q^1}, \ldots, \hat{q^n}) \right) \right) \\ \text{per } k = 1, \ldots, n. \end{cases} \tag{8.27}$$

Assumiamo ora che la il sistema fisico descritto dalla lagrangiana $\mathscr{L}$ sia debolmente invariante sotto $\mathcal{T}$:

$$\left.\frac{\mathscr{L}(t', q'^1, \ldots, q'^n,\ \dot{q}'^1, \ldots,\ \dot{q}'^n)}{\partial \epsilon}\right|_{\epsilon=0} = G(t, q'^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)\ , \tag{8.28}$$

dove G è la solita "derivata totale" di g. G e $\mathscr{L}$ sono *pensate come funzione scalari* su $j^1(\mathbb{V}^{n+1})$. È chiaro che la condizione (8.28) non dipende dalle coordinate usate: passando alle coordinate $\hat{t}, \hat{q^1}, \ldots, \hat{q^n}$, essa continua ad essere soddisfatta come si verifica immediatamente per computo diretto. Nello stesso modo si trova subito da (8.27) che (lasciamo il calcolo per esercizio):

$$I_{\mathcal{T},g} = \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial\, \dot{q}^k}\, \left.\frac{\partial q'^k}{\partial \epsilon}\right|_{\epsilon=0} - g(t, q^1, \ldots, q^n) = \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{\hat{q}}^k}\, \left.\frac{\partial \hat{q}'^k}{\partial \epsilon}\right|_{\epsilon=0} - g(\hat{t}, \hat{q}^1, \ldots, \hat{q}^n)\ . \tag{8.29}$$

8.2.4 *Le trasformazioni di simmetria (debole) di un sistema lagrangiano trasformano soluzioni delle equazioni di E.-L. in soluzioni delle stesse.

In questa sezione vogliamo mostrare l'aspetto duale della nozione di invarianza di un sistema fisico sotto un gruppo di trasformazioni. Mostreremo che se un sistema fisico è *debolmente* invariante (e quindi in particolare se è invariante) sotto un gruppo ad un parametro $\mathcal{T}$, allora l'azione del gruppo $\mathcal{T}$ trasforma soluzioni delle equazioni di Eulero-Lagrange in soluzioni delle equazioni di Eulero-Lagrange riferite alla *stessa* lagrangiana.

Teorema 8.2. *Si consideri un sistema di punti materiali $\mathfrak{S}$ descritto su $\mathbb{V}^{n+1}$ e debolmente invariante sotto il gruppo ad un parametro di diffeomorfismi locali $\mathcal{T}$, dato da (8.13), che preservano le fibre di $\mathbb{V}^{n+1}$, e siano $\mathscr{L}$ e g le funzioni che soddisfano (8.18) e (8.19).*
Si consideri una soluzione delle equazioni di Eulero-Lagrange riferite a $\mathscr{L}$:

$$q^k = q^k(t),\ \dot{q}^k =\ \dot{q}^k(t)\ , \quad \textit{per } t \in I.$$

Si consideri infine, la curva che si ottiene facendo agire $\mathcal{T}$ su $q^k = q^k(t),\ \dot{q}^k =\ \dot{q}^k(t)$ per $t \in I$:

$$\begin{cases} q'^k(t) = q'^k(\epsilon_1, t, q^1(t), \ldots, q^n(t)) \quad \textit{per } k = 1, \ldots, n. \\ \\ \dot{q}'^k(t) = \dfrac{\partial q'^k}{\partial t}(\epsilon_1, t, q^1(t), \ldots, q^n(t)) + \displaystyle\sum_{j=1}^{n} \frac{\partial q'^k}{\partial q^j}(\epsilon, t, q^1, \ldots, q^n)\, \dot{q}^j \quad \textit{per } k = 1, \ldots, n. \end{cases} \tag{8.30}$$

avendo scelto il valore $\epsilon = \epsilon_1$ per il parametro di $\mathcal{T}$. La curva, definita localmente su $j^1(\mathbb{V}^{n+1})$, $q'^k = q'^k(t),\ \dot{q}'^k =\ \dot{q}'^k(t)$ per $t \in I$ è ancora una soluzione delle equazioni di Eulero-Lagrange riferite alla stessa $\mathscr{L}$, purché $\epsilon_1 \in \mathbb{R}$ sia sufficientemente piccolo (in modo tale che le curve $q'^k = q'^k(t),\ \dot{q}'^k =\ \dot{q}'^k(t)$, per $t \in I$, per $t \in I$ siano incluse nel dominio della carta locale usata per definire $\mathcal{T}$ per ogni $\epsilon \in [0, \epsilon_1]$ se $\epsilon_1 > 0$ oppure $\epsilon \in [\epsilon_1, 0]$ se $\epsilon_1 < 0$). $\diamondsuit$

Dimostrazione.

Definiamo per $|\epsilon| \leq |\epsilon_1|$:

$$\begin{cases} q_\epsilon^k(t) = q'^k(\epsilon, t, q^1(t), \ldots, q^n(t)) \quad \text{per } k = 1, \ldots, n. \\ \\ \dot{q}_\epsilon^k(t) = \dfrac{\partial q'^k}{\partial t}(\epsilon, t, q^1(t), \ldots, q^n(t)) + \displaystyle\sum_{j=1}^n \dfrac{\partial q'^k}{\partial q^j}(\epsilon, t, q^1, \ldots, q^n)\, \dot{q}^j \quad \text{per } k = 1, \ldots, n. \end{cases} \tag{8.31}$$

Per comodità di lettura ammettiamo che il numero dei gradi di libertà sia $n = 1$, la generalizzazione al caso $n > 1$ è immediata. Con le definizioni poste abbiamo che:

$$\mathscr{L}(t, q'(t),\ \dot{q}'(t)) = \mathscr{L}(t, q(t),\ \dot{q}(t)) + \int_0^{\epsilon_1} g(t, q_\epsilon(t),\ \dot{q}_\epsilon(t)) d\epsilon\ . \tag{8.32}$$

Applichiamo ora ad ambo i memebri l'operatore differenziale:

$$\left[\frac{d}{dt}\left(\frac{\partial}{\partial\, \dot{q}}\right) - \frac{\partial}{\partial q}\right]\Bigg|_{(t, q(t),\ \dot{q}(t))}\ . \tag{8.33}$$

Il primo addendo a secondo membro di (8.32) produrrà un risultato nullo in virtù del fatto che la curva $q = q(t)$, $\dot{q} = \dot{q}(t)$ risolve le equazioni di Eulero-Lagrange. Il secondo addendo, può essere studiato passando le derivate sotto il simbolo di integrale (possiamo procedere in tal modo perché le funzioni sono differenziabili con continuità in tutti gli argomenti e $[-\epsilon_1, \epsilon_1]$ è compatto). Il risultato che si ottiene da esso è :

$$\int_0^{\epsilon_1} \left[\frac{d}{dt}\left(\frac{\partial}{\partial\, \dot{q}}\right) - \frac{\partial}{\partial q}\right]\Bigg|_{(t, q(t),\ \dot{q}(t))} g(t, q_\epsilon(t),\ \dot{q}_\epsilon(t)) d\epsilon\ .$$

L'applicazione dell'operatore (8.33) al primo membro di (8.32) fornisce invece:

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial\, \dot{q}'}\frac{\partial\, \dot{q}'}{\partial\, \dot{q}}\right)\Bigg|_{(t, q(t),\ \dot{q}(t))} - \frac{\partial \mathscr{L}}{\partial q'}\frac{\partial q'}{\partial q}\Bigg|_{(t, q(t),\ \dot{q}(t))} - \frac{\partial \mathscr{L}}{\partial\, \dot{q}'}\frac{\partial\, \dot{q}'}{\partial q}\Bigg|_{(t, q(t),\ \dot{q}(t))}\ .$$

Sviluppando le varie derivate tenendo conto di (8.31) si trova facilmente il risultato finale :

$$\left[\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial\, \dot{q}}\right) - \frac{\partial \mathscr{L}}{\partial q}\right]\Bigg|_{(t, q'(t),\ \dot{q}'(t))} \frac{\partial q'}{\partial q}\Bigg|_{(t, q(t),\ \dot{q}(t))} = \int_0^{\epsilon_1} \left[\frac{d}{dt}\left(\frac{\partial g}{\partial\, \dot{q}}\right) - \frac{\partial g}{\partial q}\right]\Bigg|_{(t, q_\epsilon(t),\ \dot{q}_\epsilon(t))} \frac{\partial q_\epsilon}{\partial q}\Bigg|_{(t, q(t),\ \dot{q}(t))} d\epsilon.$$

La funzione di t che appare nell'integrando è identicamente nulla in virtù del fatto che $g(t, q,\ \dot{q})$ ha una struttura di "derivata totale". Concludiamo che, ad ogni istante $t \in I$:

$$\left[\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial\, \dot{q}}\right) - \frac{\partial \mathscr{L}}{\partial q}\right]\Bigg|_{(t, q'(t),\ \dot{q}'(t))} \frac{\partial q'}{\partial q}\Bigg|_{(t, q(t),\ \dot{q}(t))} = 0\ .$$

Dato che la trasformazione che manda le q nelle q' è differenziabile invertibile con inversa differenziabile, la matrice jacobiana deve essere ovunque invertibile. In definitiva possiamo omettere

il secondo fattore a primo membro moltiplicando entrambi i membri per la matrice jacobiana inversa, ottenendo alla fine:

$$\left[\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}}\right) - \frac{\partial \mathscr{L}}{\partial q}\right]\Bigg|_{(t, q'(t),\, \dot{q}'(t))} = 0\,,$$

per ogni valore di $t \in I$. Questo è quanto volevamo provare. $\square$

8.3 L'integrale primo di Jacobi, invarianza sotto "traslazioni temporali" e conservazione dell'energia meccanica.

In questa sezione ci occupiamo di un integrale primo ben noto molto tempo prima del teorema di Nöther. Si tratta dell'*integrale primo di Jacobi*. Enunciamo subito il teorema corrispondente.

Teorema 8.3. **(Teorema di Jacobi.)** *Sia $\mathfrak{S}$ un sistema di N punti materiali sottoposto a $c = 3N - n \geq 0$ vincoli olonomi ideali con lagrangiana $\mathscr{L}$. Se in riferimento ad un sistema di coordinate naturali $(t, q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n)$ definite sull'aperto U nello spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$ vale l'identità :*

$$\frac{\partial \mathscr{L}}{\partial t} = 0\,, \quad \textit{ovunque su } U, \tag{8.34}$$

allora valgono i seguenti fatti.
(a) *La funzione definita su U, detta* **hamiltoniana del sistema**,

$$\mathscr{H}(t, q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n) := \sum_{k=1}^{n} \dot{q}^k \frac{\partial \mathscr{L}}{\partial \dot{q}^k} - \mathscr{L}\,, \tag{8.35}$$

si conserva sui moti del sistema (nella porzione di essi contenuta in U) e si chiama in tal caso **integrale primo di Jacobi**.
(b) *Se (i) le forze attive su $\mathfrak{S}$ sono conservative nel riferimento $\mathscr{I}$, (ii) nello stesso riferimento $\mathscr{I}$ i vincoli non dipendono esplicitamente dal tempo, (iii) le coordinate libere $(t, q^1, \ldots, q^n)$ sono solidali con $\mathscr{I}$ ed infine: $\mathscr{L} = \mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$, allora $\mathscr{H}$ coincide con l'energia meccanica totale di $\mathfrak{S}$ valutata in $\mathscr{I}$:*

$$\mathscr{H}(q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n) := \mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n) + \mathscr{U}|_{\mathscr{I}}(q^1, \ldots, q^n,\, \dot{q}^1, \ldots,\, \dot{q}^n)\,. \tag{8.36}$$

$\diamondsuit$

Dimostrazione. (a) Sia $t \mapsto (t, q^1(t), \ldots, q^n(t),\, \dot{q}^1(t), \ldots,\, \dot{q}^n(t))$ una soluzione delle equazioni di Eulero-Lagrange. Direttamente dalla definizione di $\mathscr{H}$ abbiamo che:

$$\frac{d}{dt}\mathscr{H}(t, q^1(t), \ldots, q^n(t),\, \dot{q}^1(t), \ldots,\, \dot{q}^n(t))$$

$$= \sum_k \frac{d\dot{q}^k}{dt}\frac{\partial \mathscr{L}}{\partial \dot{q}^k} + \sum_k \dot{q}^k \frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{q}^k} - \frac{d}{dt}\mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))\,.$$

Usando le equazioni di Eulero-Lagrange possiamo riscrivere l'identità di sopra come:

$$\frac{d}{dt}\mathscr{H}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))$$

$$= \sum_k \frac{d\dot{q}^k}{dt}\frac{\partial \mathscr{L}}{\partial \dot{q}^k} + \sum_k \dot{q}^k \frac{\partial \mathscr{L}}{\partial q^k} - \frac{d}{dt}\mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))\,.$$

Aggiungendo e togliendo $\partial \mathscr{L}/\partial t$ a secondo membro si ottiene infine:

$$\frac{d}{dt}\mathscr{H} = \sum_k \frac{d\dot{q}^k}{dt}\frac{\partial \mathscr{L}}{\partial \dot{q}^k} + \sum_k \dot{q}^k \frac{\partial \mathscr{L}}{\partial q^k} + \frac{\partial \mathscr{L}}{\partial t} - \frac{d}{dt}\mathscr{L} - \frac{\partial \mathscr{L}}{\partial t}\,.$$

D'altra parte, dato che $\mathscr{L} = \mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))$ deve anche essere:

$$\frac{d}{dt}\mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t)) = \sum_k \frac{d\dot{q}^k}{dt}\frac{\partial \mathscr{L}}{\partial \dot{q}^k} + \sum_k \dot{q}^k \frac{\partial \mathscr{L}}{\partial q^k} + \frac{\partial \mathscr{L}}{\partial t}\,.$$

Inserendo questa identità nell'espressione trovata sopra otteniamo infine che, su ogni soluzione delle equazioni di Eulero-Lagrange vale l'identità notevole:

$$\frac{d}{dt}\mathscr{H}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t)) = -\frac{\partial}{\partial t}\mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t))\,.$$

Dato che vale (8.34) si ha infine:

$$\frac{d}{dt}\mathscr{H}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t)) = -\frac{\partial}{\partial t}\mathscr{L}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t)) = 0\,,$$

e pertanto su ogni soluzione delle equazioni di Eulero-Lagrange $\mathscr{H}(t, q^1(t), \ldots, q^n(t), \dot{q}^1(t), \ldots, \dot{q}^n(t))$ è una costante al variare del tempo.
(b) Nelle ipotesi fatte, usando il contenuto della sezione 6.1.3, abbiamo che:

$$\mathscr{T}|_{\mathscr{I}}(q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k$$

dove

$$a_{hk}(q^1, \ldots, q^n) = \sum_{i=1}^{N} \frac{m_i}{2}\frac{\partial \mathbf{x}_i}{\partial q^h} \cdot \frac{\partial \mathbf{x}_i}{\partial q^k}\,,$$

dove infine $\mathbf{x}_i = P_i(q^1, \ldots, q^n) - O$ è il vettore posizione in $\mathscr{I}$ del punto materiale i-esimo (O è l'origine di un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$). In questo caso troviamo:

$$\mathscr{L}|_{\mathscr{I}}(t, q^1(t), \ldots, q^n(t),\ \dot{q}^1(t), \ldots,\ \dot{q}^n(t)) = \sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k - \mathscr{U}(q^1, \ldots, q^n)\,.$$

Il calcolo diretto usando (8.35) produce:

$$\mathscr{H}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \sum_{h,k=1}^{n} a_{hk}(q^1, \ldots, q^n)\, \dot{q}^h\, \dot{q}^k + \mathscr{U}|_{\mathscr{I}}(q^1, \ldots, q^n)\,.$$

ossia

$$\mathscr{H}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \mathscr{T}|_{\mathscr{I}} + \mathscr{U}|_{\mathscr{I}}\,.$$

Questo conclude la dimostrazione. □

Osservazioni 8.1.
(1) Sembrerebbe che il teorema di Jacobi sia logicamente slegato dal teorema di Nöther, in particolare perché le trasformazioni coinvolte nel teorema di Jacobi sono trasformazioni che muovono le coordinate temporali (vedi l'osservazione seguente) e pertanto non preservano le fibre di $j^1(\mathbb{V}^{n+1})$. In realtà ciò è dovuto al fatto che abbiamo enunciato il teorema di Nöther in una versione troppo elementare per includere il teorema di Jacobi. Nella prossima sezione mostreremo che il teorema di Jacobi è un sottocaso di una forma più generale del teorema di Nöther.
(2) La condizione (8.34):

$$\frac{\partial \mathscr{L}}{\partial t} = 0\,, \quad \text{ovunque su } U,$$

si può interpretare come *invarianza del sistema fisico sotto traslazioni temporali.* In altre parole (8.34) sancisce che il sistema fisico ammette una lagrangiana che verifica:

$$\mathscr{L}(t+\tau, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n) = \mathscr{L}(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)\,, \quad \text{per ogni } \tau\,.$$

In questo senso, se siamo nelle ipotesi dell'enunciato (b), il teorema di Jacobi dice che l'energia meccanica si conserva quando la lagrangiana è invariante sotto traslazioni temporali. Si deve avere ben chiaro però che la condizione (8.34) dipende *strettamente* dalla scelta delle coordinate libere usate $(t, q^1, \ldots, q^n)$. La condizione (8.34)) riferita alle nuove coordinate libere $(t', q'^1, \ldots, q'^n)$, cioé : (supponendo per semplicità $n = 1$)

$$\frac{\partial \mathscr{L}(t'+c, q(t', q'),\ \dot{q}(t', q',\ \dot{q}'))}{\partial t'} = 0\,,$$

può essere falsa a causa della possibile dipendenza esplicita da t' nella relazione che lega le nuove coordinate alle vecchie:

$$t = t' + c\,, \quad q^k = q^k(t', q'^1, \ldots, q'^n)\,.$$

Nello stesso modo, l'hamiltoniana $\mathscr{H}$ dipende strettamente dalle coordinate libere usate: al contrario di $\mathscr{L}$, *non* è un campo scalare su $j^1(\mathbb{V}^{n+1})$. Inoltre $\mathscr{H}$ in generale *non* è definito globalmente su + tutto lo spaziotempo degli atti di moto, ma solo su ogni aperto U dotato di coordinate naturali adattate alle fibre.
(3) Supponiamo che $\mathscr{L}$, data da $\mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$, non dipenda esplicitamente dal tempo come

richiesto in (8.34), ma le coordinate libere usate non siano adattate a $\mathscr{I}$. In tal caso l'integrale primo di Jacobi esiste comunque, ma non ha il significato di energia meccanica totale in $\mathscr{I}$. Può in effetti coincidere con l'energia meccanica totale in un altro sistema di riferimento. Un esempio concreto di questa situazione è descritto nell'esercizio 6.1.**3**.

8.4 Commenti finali sul teorema di Nöther.

8.4.1 Invarianza sotto il gruppo di Galileo in meccanica lagrangiana.

Possiamo ora enunciare il postulato di invarianza della teoria rispetto al gruppo di Galileo riferendoci al contenuto della sezione 2.4.1 e riadattandolo alla descrizione fisica in termini di Lagrangiane. La richiesta è semplice: *in un fissato sistema di riferimento inerziale, ogni sistema di punti materiali interagenti, non sottoposto a vincoli, isolato all'esterno e che ammette descrizione in termini di lagrangiana, deve essere tale che la sua lagrangiana oppure una equivalente ad essa (nel senso che produca le stesse equazioni di Eulero-Lagrange) sia invariante o debolmente invariante sotto l'azione dei sottogruppi ad un parametro del gruppo di Galileo: traslazioni rispetto ad una fissata (arbitrariamente) direzione, rotazioni attorno ad un fissato (arbitrariamente) asse di rotazione, traslazioni temporali, trasformazioni pure di Galileo in una fissata (arbitrariamente) direzione.*
In altre parole abbiamo imposto che nella classe dei riferimenti inerziali la descrizione della meccanica sia tale da verificare, nel linguaggio della meccanica lagrangiana: l'*omogeneità dello spazio*, l'*isotropia dello spazio*, l'*omogeneità del tempo* e l'*invarianza galileiana pura*. In base al teorema di Nöther (che *include* quello di Jacobi nella formulazione avanzata che daremo nella prossima sezione), ad ognuna di queste proprietà di simmetria corrisponde una grandezza fisica conservata. Riassumendo tutti i risultati ottenuti in questo capitolo per sistemi isolati descritti nel riferimento inerziale $\mathscr{I}$ e che ammettono lagrangiana della forma $\mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$, con $\mathscr{U}$ energia potenziale, abbiamo che:

(i) *l'omogeneità spaziale implica la conservazione dell'impulso totale del sistema*,

(ii) *l'isotropia spaziale implica la conservazione del momento angolare totale del sistema*,

(iii) *l'omogeneità temporale implica la conservazione dell'energia meccanica totale del sistema.*

(iv) *l'invarianza galileiana pura fornisce una grandezza conservata tale che la relazione di conservazione si interpreta come il teorema del centro di massa (vedi l'esempio 2 in esempi 8.5).*

8.4.2 Formulazione lagrangiana e teorema di Nöther oltre la meccanica classica.

Per concludere la discussione riassuntiva della sezione precedente, ci sarebbe da discutere cosa accade per sistemi che ammettono lagrangiana di forma diversa da $\mathscr{T}|_{\mathscr{I}} - \mathscr{U}|_{\mathscr{I}}$. Ma simili discussioni sarebbero accademiche dato che il mondo fisico *non* è invariante sotto il gruppo di Galileo,

ma lo è invece (almeno lontano da forti campi gravitazionali) rispetto al gruppo di Poincarè. La discussione richiederebbe un'estensione del formalismo lagrangiano e del teorema di Nöther, alle teorie relativistiche includendo la teoria dei campi. Diciamo solo che tale estensione esiste, è molto naturale ed ha provato essere di grandissima potenza nelle formulazioni della fisica moderna (specialmente nella teoria delle interazioni fondamentali e delle particelle elementari), dove il legame simmetrie-leggi di conservazione ha giocato e gioca un ruolo fondamentale.
A titolo di esempio si pensi già in elettrodinamica quanto sia difficile definire l'impulso del campo eletromagnetico (oppure il suo momento angolare). L'estensione del formalismo lagrangiano e del teorema di Nöther alla teoria dei campi consente di *definire* l'impulso del campo elettromagnetico come quella grandezza che si conserva come conseguenza dell'omogeneità spaziale (invarianza sotto traslazioni spaziali della lagrangiana del campo elettromagnetico). Alla prova dei fatti si vede che questa è la definizione *fisicamente* corretta anche perchè, in particolare, consente di implementare il principio di conservazione dell'impulso: con la definizione data si dimostra che per un sistema isolato, composto di particelle e campi elettromagnetici che interagiscono tra di essi, l'impulso totale di si conserva nel tempo.

8.5 *Formulazione generale e globale del Teorema di Nöther.

Il teorema 3.11 permette di associare univocamente un campo vettoriale ad ogni gruppo ad un parametro di diffeomorfismi locali che preservano le fibre di $\mathbb{V}^{n+1}$, $\mathcal{T}$ dato in (8.13). Infatti, se consideriamo, al tempo fissato t, il campo vettoriale definito localmente su $\mathbb{Q}_t$ definito nelle coordinate $q^1, \ldots, q^n$ da

$$X_0(t, q^1, \ldots, q^n) := \sum_{k=1}^{n} \left.\frac{\partial q'^k(\epsilon, t, q^1, \ldots, q^n)}{\partial \epsilon}\right|_{\epsilon=0} \left.\frac{\partial}{\partial q^k}\right|_{(t,q^1,\ldots,q^n)} + 0 \left.\frac{\partial}{\partial t}\right|_{(t,q^1,\ldots,q^n)}$$

risulta facilmente usando (8.14) e (8.15) che, per un fissato punto $q^1, \ldots, q^n$, se consideriamo la curva $\epsilon \mapsto (t'(\epsilon, t, q^1, \ldots, q^n), q'^1(\epsilon, t, q^1, \ldots, q^n), \cdots, q'^n(\epsilon, t, q^1, \ldots, q^n)) \in \mathbb{V}^{n+1}$ vale

$$\frac{dt'}{dt} = 0\,, \quad \frac{dq'^k}{d\epsilon} = X^k(q'^1(\epsilon, q^1, \ldots, q^n), \cdots, q'^n(\epsilon, q^1, \ldots, q^n))\,.$$

In altre parole il gruppo ad un parametro $\mathcal{T}$ è dato dalle curve integrali del campo X_0 definito sopra: il campo X_0 è il generatore del gruppo. Il punto individuato dalle coordinate

$$(t'(\epsilon, t, q^1, \ldots, q^n), q'^1(\epsilon, t, q^1, \ldots, q^n), \cdots, q'^n(\epsilon, t, q^1, \ldots, q^n))$$

è il punto che si trova, per il valore del parametro ϵ, sull'unica curva integrale in $\mathbb{V}^{n+1}$ del campo X che parte da $(t, q^1, \ldots, q^n)$ per $\epsilon = 0$.
Con un pò di pazienza si può verificare che l'insieme delle trasformazioni su $j^1(\mathbb{V}^{n+1})$, indotte

da $\mathcal{T}$,

$$\begin{cases} t' = t\,, \\ q'^k = q'^k(\epsilon, t, q^1, \dots, q^n) \quad \text{per } k = 1, \dots, n. \\ \dot{q}'^k = \dfrac{\partial q'^k}{\partial t}(\epsilon, t, q^1, \dots, q^n) + \displaystyle\sum_{j=1}^{n} \dfrac{\partial q'^k}{\partial q^j}(\epsilon, t, q^1, \dots, q^n)\, \dot{q}^j \quad \text{per } k = 1, \dots, n. \end{cases} \tag{8.37}$$

soddisfano ancora le analoghe delle proprietà (8.14) e (8.15) per $\epsilon \in (-a, a)$. Pertanto l'azione del gruppo $\mathcal{T}$ estesa localmente ad $\mathbb{A}_t$ tramite le (8.37) si ottiene ancora dalle curve integrali del campo, definito localmente su $\mathbb{A}_t$:

$$X(t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n) := \sum_{k=1}^{n} \left.\frac{\partial q'^k(\epsilon, t, q^1, \dots, q^n)}{\partial \epsilon}\right|_{\epsilon=0} \left.\frac{\partial}{\partial q^k}\right|_{(t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n)}$$

$$+\, 0 \left.\frac{\partial}{\partial t}\right|_{(t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n)} + \sum_{k=1}^{n} \left.\frac{\partial\, \dot{q}'^k(\epsilon, t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n)}{\partial \epsilon}\right|_{\epsilon=0} \left.\frac{\partial}{\partial\, \dot{q}^k}\right|_{(t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n)}. \tag{8.38}$$

Le equazioni (8.37) definiscono ancora un gruppo ad un parametro di diffeomorfismi locali agente su un aperto di $j^1(\mathbb{V}^{n+1})$, il cui generatore è il campo X definito in (8.38). La condizione di invarianza (8.16) può allora essere scritta equivalentemente, pensando il campo X come operatore differenziale sui campi scalari definiti su $\mathbb{A}_t$:

$$X\,(\mathscr{L}) = 0\,. \tag{8.39}$$

Osservazioni 8.2. La (8.39) significa che su ogni $\mathbb{A}_t$ e lungo ogni linea integrale di $\widetilde{X}$, $(c, d) \ni \epsilon \mapsto \gamma(\epsilon) \in \mathbb{A}_t$, vale

$$\frac{d}{d\epsilon}\mathscr{L}(t, \gamma(\epsilon)) = 0\,, \quad \text{per ogni } \epsilon \in (c, d)$$

e pertanto la lagrangiana assume valore costante (che dipenderà dalla linea integrale). Integrando la (8.39) lungo la linea integrale che parte da $q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n$ e termina in $q'^1, \dots, q'^n,\, \dot{q}'^1, \dots,\, \dot{q}'^n$ si ottiene la condizione apparentemente molto più forte di (8.16) (che equivale alla (8.39)):

$$\mathscr{L}(t', q'^1, \dots, q'^n,\, \dot{q}'^1, \dots,\, \dot{q}'^n) = \mathscr{L}(t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n) \tag{8.40}$$

per ogni scelta di $\epsilon \in (-a, a)$ e $(t, q^1, \dots, q^n,\, \dot{q}^1, \dots,\, \dot{q}^n)$ nel sistema di coordinate locali su $j^1(\mathbb{V}^{n+1})$ usato per definire $\mathcal{T}$. In realtà, come appena verificato, le due condizioni (8.16) e (8.40) sono *del tutto equivalenti*.

Nel seguito enunceremo una forma generalizzata del teorema di Nöther usando questa identificazione dei gruppi ad un parametro di diffeomorfismi locali e campi vettoriali: invece di parlare di gruppi ad un parametro di trasformazioni parleremo semplicemente di campi vettoriali. L'ipotesi principale del teorema di Nöther, cioé la condizione di invarianza della lagrangiana sotto un gruppo di trasformazioni sarà scritta in termini di campi vettoriali come in (8.39). Quello che ci si guadagna in questo modo, dato che i campi vettoriali sono definibili globalmente su $j^1(\mathbb{V}^{n+1})$ senza troppe complicazioni tecniche, è che l'enunciato e la validità del teorema risultano completamente indipendenti dalla scelta delle coordinate e risultano essere validi globalmente.
Passiamo a trascrivere, con gli stessi benefici, anche la tesi del teorema di Nöther nel linguaggio dei campi vettoriali sullo spaziotempo degli atti di moto. Ricordiamo a tal fine che le stesse soluzioni massimali delle equazioni di Eulero-Lagrange sono definite globalmente su $j^1(\mathbb{V}^{n+1})$ e non sono altro che le linee integrali del campo (6.47), definito globalmente, e in coordinate locali naturali dato da:

$$Z(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \frac{\partial}{\partial t} + \sum_{k=1}^{n} \dot{q}^k \frac{\partial}{\partial q^k} + \sum_{k=1}^{n} z^k(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) \frac{\partial}{\partial \dot{q}^k} .$$

Utilizzando il vettore Z abbiamo la seguente proposizione:

Proposizione 8.4. *Si consideri un sistema fisico descritto su $j^1(\mathbb{V}^{n+1})$ da equazioni di Eulero-Lagrange associate al campo vettoriale dinamico Z su $j^1(\mathbb{V}^{n+1})$. Assumiamo che Z sia di classe C^1. La condizione che una funzione $f : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$, di classe C^1, sia un integrale primo è equivalente alla condizione:*

$$Z(f) = 0 , \quad \textit{ovunque su } j^1(\mathbb{V}^{n+1}) . \tag{8.41}$$

$\diamondsuit$

Dimostrazione. Le equazioni di Eulero-Lagrange si scrivono in forma globale, tenendo conto della definizione di Z:

$$\dot{\gamma}(t) = Z(\gamma(t)) ,$$

dove $I \ni t \mapsto \gamma(t) \in j^1(\mathbb{V}^{n+1})$ è una curva sullo spaziotempo degli atti di moto. Di conseguenza (8.41) implica che:

$$\frac{d}{dt} f(\gamma(t)) = 0 , \quad \text{per ogni valore di } t \in I$$

per ogni soluzione γ delle equazioni di Eulero-Lagrange e quindi che f è un integrale primo. Viceversa, se f è un integrale primo vale

$$\frac{d}{dt} f(\gamma(t)) = 0 , \quad \text{per ogni valore di } t \in I$$

per ogni soluzione $\gamma : I \to j^1(\mathbb{V}^{n+1})$ delle equazioni di Eulero-Lagrange. Ma allora, per definizione di Z abbiamo in particolare che, se $t_0 \in I$:

$$Z_{\gamma(t_0)}(f) = \frac{d}{dt}|_{t=t_0} f(\gamma(t)) = 0 .$$

In altri termini vale:

$$Z_p(f) = 0$$

per ogni punto $p \in j^1(\mathbb{V}^{n+1})$ che ammette una soluzione γ delle equazioni di Eulero-Lagrange definita nell'intorno di t_0 e con $\gamma(t_0) = p$. Per il teorema di esistenza, essendo il campo Z di classe C^1 almeno, ogni punto $p \in j^1(\mathbb{V}^{n+1})$ soddisfa tale richiesta. $\square$

Nella tesi del teorema di Nöther la conservazione dell'integrale primo I sulle soluzioni delle equazioni di Eulero-Lagrange, sarà dunque enunciata semplicemente con $Z(I) = 0$ ovunque sullo spaziotempo degli atti di moto.

8.5.1 Il teorema di Nöther nella forma generale.

Vogliamo ora enunciare il teorema di Nöther nel linguaggio dei campi vettoriali. Il campo X sarà globalmente definito su $j^1(\mathbb{V}^{n+1})$ ed è pensato come generatore di un gruppo ad un parametro di diffeomorfismi locali che lasciano invariata la lagrangiana di un sistema fisico (almeno in senso debole). In coordinate naturali adattate alle fibre di $j^1(\mathbb{V}^{n+1})$ ogni campo X si può decomporre come:

$$X = X^t \frac{\partial}{\partial t} + \sum_{k=1}^{n} X_0^k \frac{\partial}{\partial q^k} + \sum_{k=1}^{n} X^k \frac{\partial}{\partial \dot{q}^k} , \tag{8.42}$$

dove X^t, X_0^k e X^k sono funzioni differenziabili di $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$. Nelle ipotesi del teorema di Nöther in forma elementare si assumeva che, pensando Z come operatore differenziale sulle funzioni differenziabili definite su $j^1(\mathbb{V}^{n+1})$,

$$X^k = Z\left(X_0^k\right) , \quad k = 1, \ldots, n .$$

Infatti, questa condizione è sicuramente verificata dal generatore di ogni gruppo locale della forma (8.38):

$$Z\left(\frac{\partial q'^k(\epsilon, t, q^1, \ldots, q^k)}{\partial \epsilon}|_{\epsilon=0}\right) = \frac{\partial}{\partial t}\frac{\partial q'^k(\epsilon, t, q^1, \ldots, q^k)}{\partial \epsilon}|_{\epsilon=0} + \sum_k \dot{q}^k \frac{\partial q'^k(\epsilon, t, q^1, \ldots, q^k)}{\partial \epsilon}|_{\epsilon=0}$$

$$= \left.\frac{\partial\, \dot{q}'^k(\epsilon, t, q^1, \ldots, q^k)}{\partial \epsilon}\right|_{\epsilon=0} ,$$

dove nell'ultimo passaggio abbiamo tenuto conto dell'ultima equazione in (8.37). Questa ipotesi sembra essenziale dal punto di vista fisico: ricorda in qualche modo al gruppo di trasformazioni come il mondo delle velocità (delle coordinate $\dot{q}^k$) e quello delle posizioni (le coordinate q^k) sono imparentati. Nel seguito assumeremo ancora questa ipotesi anche se assumeremo, molto più generalmente di quanto fatto nel teorema di Nöther in forma elementare, che le funzioni X_0^k possano dipendere anche dalle coordinate $\dot{q}^k$. I gruppi ad un parametro che ne conseguono

sono tali che la loro azione nello spaziotempo delle configurazioni dipende non solo dalle configurazioni stesse, ma anche dalle velocità che il sistema assume nelle varie configurazioni. Si dice in questo caso che si considerano **simmetrie dinamiche** (quelle cioé in cui X_0^k è funzione di tutte le coordinate $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$) piuttosto che **simmetrie geometriche** (quelle cioé in cui X_0^k è funzione delle sole coordinate $(t, q^1, \ldots, q^n)$).
La richiesta $X^k = Z\left(X_0^k\right)$, se la si assume valida, come è naturale, in *ogni sistema di coordinate naturali su* $j^1(\mathbb{V}^{n+1})$, comporta dei vincoli non banali nel caso generale in cui ogni X_0^k è funzione di tutte le coordinate $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$. Vale infatti la seguente proposizione.

Proposizione 8.5. *Se un campo X su $j^1(\mathbb{V}^{n+1})$ di classe C^1 soddisfa la condizione*

$$X^k = Z\left(X_0^k\right), \quad \text{per } k = 1, \ldots, n \text{ in ogni sistema di coordinate naturali su } j^1(\mathbb{V}^{n+1}), \tag{8.43}$$

rispetto ad un campo vettoriale dinamico $Z : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ (indotto da qualche lagrangiana), allora è verificata la condizione

$$Z(\langle X, dT\rangle) = 0\,. \tag{8.44}$$

Se vale (8.44) e vale (8.43) in un atlante di coordinate naturali locali di $j^1(\mathbb{V}^{n+1})$, allora (8.43) vale in ogni sistema di coordinate locali naturali di $j^1(\mathbb{V}^{n+1})$. $\diamondsuit$

Dimostrazione. (8.44) si scrive localmente $Z(X^t) = 0$ in ogni sistema di coordinate naturali. Dato che le componenti X^t si trasformano scalarmente cambiando carta locale naturale:

$$X'^{t'} = \frac{\partial t'}{\partial t} X^t = X^t\,,$$

per provare che vale la prima parte della tesi è sufficiente mostrare che c'é un sistema di coordinate naturali nell'intorno di ogni punto in cui $Z(X^t) = 0$. Supponiamo che (8.43) valga in ogni sistema di coordinate locali naturali adattati alle fibre di $j^1(\mathbb{V}^{n+1})$. Se $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ e $(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)$ sono siffatti sistemi di coordinate definiti su uno stesso aperto $U \subset j^1(\mathbb{V}^{n+1})$, tenendo conto delle (6.31),(6.32) e (6.33), avremo le relazioni per $j = 1, \ldots, n$:

$$X'^j = \sum_{k=1}^n \frac{\partial\, \dot{q}'^j}{\partial\, \dot{q}^k} X^k + \sum_{k=1}^n \frac{\partial\, \dot{q}'^j}{\partial q^k} X_0^k + \frac{\partial\, \dot{q}'^k}{\partial t} X^t\,, \tag{8.45}$$

$$X_0'^j = \sum_{k=1}^n \frac{\partial q'^j}{\partial q^k} X_0^k + \frac{\partial q'^j}{\partial t} X^t\,, \tag{8.46}$$

$$X'^{t'} = X^t\,.$$

Dato che (8.43) vale in entrambi i sistemi di coordinate, dobbiamo concludere che:

$$Z\left(\sum_{k=1}^n \frac{\partial q'^j}{\partial q^k} X_0^k + \frac{\partial q'^j}{\partial t} X^t\right) = \sum_{k=1}^n \frac{\partial\, \dot{q}'^j}{\partial\, \dot{q}^k} Z(X_0^k) + \sum_{k=1}^n \frac{\partial\, \dot{q}'^j}{\partial q^k} X_0^k + \frac{\partial\, \dot{q}'^j}{\partial t} X^t\,. \tag{8.47}$$

Tenendo conto della relazione (6.33) otteniamo:

$$\frac{\partial\, \dot{q}'^j}{\partial\, \dot{q}^k} = \frac{\partial q'^j}{\partial q^k}\ ,$$

per cui possiamo riscrivere l'identità di sopra come:

$$Z\left(\sum_{k=1}^{n} \frac{\partial q'^j}{\partial q^k} X_0^k + \frac{\partial q'^j}{\partial t} X^t\right) = \sum_{k=1}^{n} \frac{\partial q'^j}{\partial q^k} Z(X_0^k) + \sum_{k=1}^{n} \frac{\partial\, \dot{q}'^j}{\partial q^k} X_0^k + \frac{\partial\, \dot{q}'^j}{\partial t} X^t\ ,$$

che si esplicita in:

$$\sum_{k=1}^{n} Z\left(\frac{\partial q'^j}{\partial q^k}\right) X_0^k + \sum_{k=1}^{n} \frac{\partial q'^j}{\partial q^k} Z\left(X_0^k\right) + \sum_{k=1}^{n} Z\left(\frac{\partial q'^j}{\partial t}\right) X^t + \frac{\partial q'^j}{\partial t} Z\left(X^t\right)$$

$$= \sum_{k=1}^{n} \frac{\partial q'^j}{\partial q^k} Z(X_0^k) + \sum_{k=1}^{n} \frac{\partial\, \dot{q}'^j}{\partial q^k} X_0^k + \frac{\partial\, \dot{q}'^j}{\partial t} X^t\ . \tag{8.48}$$

Dato che:

$$Z\left(\frac{\partial q'^j}{\partial q^k}\right) = \frac{\partial\, \dot{q}'^j}{\partial q^k} \quad \text{e} \quad Z\left(\frac{\partial q'^j}{\partial t}\right) = \frac{\partial\, \dot{q}'^j}{\partial t}\ ,$$

(8.48) si riduce a:

$$0 = \frac{\partial q'^j}{\partial t} Z\left(X^t\right) \quad \text{per ogni coppia di sistemi di coordinate locali naturali su } j^1(\mathbb{V}^{n+1})\ . \tag{8.49}$$

Mostriamo che la validità di questa identità implica la tesi. Nell'intorno di ogni punto $p \in j^1(\mathbb{V}^{n+1})$, fissato il sistema di coordinate naturali $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ con origine corrispondente a p, possiamo fissare le nuove coordinate $(t', q'^1, \ldots, q'^n,\ \dot{q}'^1, \ldots,\ \dot{q}'^n)$ in modo tale che $q'^j = q^j$ se $j > 1$ e $q'^1 = q^1 + ct$, dove $c \neq 0$ (queste nuove coordinate sono ammissibili nell'intorno di p in quanto la matrice jacobiana della trasformazione di coordinate ha determinante non nullo in p). L'identità (8.49) imposta in p implica allora che:

$$0 = cZ\left(X^t\right)|_p\ ,$$

per cui $Z\left(X^t\right)|_p = 0$. Questo risultato vale per ogni $p \in j^1(\mathbb{V}^{n+1})$ in un sistema di coordinate naturali definite nell'intorno di p come volevamo provare.
La seconda parte della tesi è ora ovvia: se vale vale (8.44) e in un atlante di coordinate naturali (non primate) valgono le relazioni (8.43), in ogni sistema locale di coordinate naturali (primate) deve valere (8.47). Usando (8.43) per le coordinate (non primate) dell'atlante, concludiamo da (8.45) e (8.46) che

$$Z(X_0'^j) = X'^j\ ,$$

cioé la validità di (8.43) in ogni sistema di coordinate locali naturali. □

Osservazioni 8.3.
(1) Nelle applicazioni è bene tener presente che la condizione (8.44), si scrive in coordinate locali naturali ed in relazione alla decomposizione (8.42):

$$Z(X^t) = 0\,. \tag{8.50}$$

(2) Campi vettoriali X su $j^1(\mathbb{V}^{n+1})$ che soddisfi la condizioni (8.43) si ottengono *rialzando* a $j^1(\mathbb{V}^{n+1})$ *per mezzo del campo* Z ogni campo vettoriale $\tilde{X}$ su $\mathbb{V}^{n+1}$ che soddisfi (8.43) purché siano soddisfatte opportune ipotesi. In coordinate locali naturali adattate alle fibre, se

$$\tilde{X} = \tilde{X}^t \frac{\partial}{\partial t} + \tilde{X}^k \frac{\partial}{\partial q^k}\,,$$

dove $X^k = X^k(t, q^1, \ldots, q^n)$, il suo rialzamento a $j^1(\mathbb{V}^{n+1})$, X, è definito come:

$$X = \tilde{X}^t \frac{\partial}{\partial t} + \tilde{X}^k \frac{\partial}{\partial q^k} + Z(\tilde{X}^k) \frac{\partial}{\partial\, \dot{q}^k}\,.$$

Questa definizione non dipende dalle coordinate locali usate e quindi definisce un campo globalmente su $j^1(\mathbb{V}^{n+1})$ quando, come sappiamo dalla proposizione precedente, $Z(\tilde{X}^t) = 0$. Quest'ultima condizione si riduce in questo caso, dato che $\tilde{X}^t$ può dipendere solo da t e dalle q^k, ma non dalle coordinate puntate, alla richiesta:

$$\frac{\partial \tilde{X}^t}{\partial t} + \sum_{k=1}^{n} \frac{\partial \tilde{X}^t}{\partial q^k}\, \dot{q}^k = 0\,,$$

che deve valere ovunque. Da questa condizione si ricava subito che le derivate parziali di $\tilde{X}^t$ rispetto a t ed alle q^k devono essere ovunque nulle e pertanto l'unica possibilità è quella di $\tilde{X}^t =$ costante.
Concludiamo che, *per un campo X su $j^1(\mathbb{V}^{n+1})$ che soddisfi (8.43), $\langle X, dT\rangle$ può non essere costante solo quando X non è il rialzamento di un campo preesistente su $\mathbb{V}^{n+1}$*. In particolare le simmetrie geometriche sono tali che $\langle X, dT\rangle$ sia costante.
(3) Le due condizioni (8.43) e (8.44) sono equivalenti alle seguenti, lasciamo la prova per esercizio al lettore:

$$Z(\langle X, dT\rangle) = 0\,, \tag{8.51}$$

$$\langle \mathcal{L}_Z(X), \omega_k\rangle = 0\,, \quad \text{per ogni forma di contatto } \omega_k. \tag{8.52}$$

Dove abbiamo usato la derivata di Lie $\mathcal{L}_Z$ rispetto a Z (vedi l'appendice A). Le forme di contatto, definite in (8.57), sono riferite ad un sistema di coordinate locali naturali, *per ogni forma di contatto* ω_k significa che bisogna variare anche i sistemi di coordinate locali.

la condizione $\langle X, dT\rangle = 0$ in (8.38) (dove $T : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ è il tempo assoluto) dice che il gruppo ad un parametro generato da X *preserva le fibre di* $j^1(\mathbb{V}^{n+1})$. Si osservi infatti che tale

condizione si esplicita semplicemente in $X^t = 0$ ovunque in ogni sistema di coordinate locali naturali adattate alle fibre per cui il significato di preservazione delle fibre da parte del gruppo ad un parametro di diffeomorfismi locali generati da X è del tutto evidente.
Questa condizione è lasciata cadere nell'enunciato del teorema di Nöther in forma globale e generale che diamo di seguito.

Teorema 8.4. (**Teorema di Nöther in forma generale e globale**.) *Si consideri un sistema fisico descritto su $j^1(\mathbb{V}^{n+1})$ dalla lagrangiana $\mathscr{L}$. Sia X dato è un campo vettoriale differenziabile su $j^1(\mathbb{V}^{n+1})$ che soddisfa la proprietà (8.43):*

$$X^k = Z\left(X_0^k\right) \ , \quad \textit{per } k = 1, \ldots, n \textit{ in ogni sistema di coordinate naturali su } j^1(\mathbb{V}^{n+1}), \tag{8.53}$$

essendo Z il campo vettoriale dinamico su $j^1(\mathbb{V}^{n+1})$ associato a $\mathscr{L}$. Se vale la condizione di **invarianza debole** *di $\mathscr{L}$ sotto l'azione del campo X:*

$$X(\mathscr{L}) = Z(f) \ , \tag{8.54}$$

dove $f : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ è una funzione differenziabile, allora si ha l'integrale primo di Nöther $I_{X,f} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ definito globalmente da:

$$I_{X,f} := \langle X, \Omega_{\mathscr{L}} \rangle - f \ , \tag{8.55}$$

dove abbiamo introdotto la **forma di Poincaré -Cartan**, *$\Omega_{\mathscr{L}}$, individuata in coordinate locali naturali adattate alle fibre di di $j^1(\mathbb{V}^{n+1})$, da:*

$$\Omega_{\mathscr{L}} := \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k} \omega_k + \mathscr{L} dT \ , \tag{8.56}$$

essendo ω_k le **forme di contatto** *associate alle coordinate locali naturali usate:*

$$\omega_k := dq^k - \dot{q}^k dt \ . \tag{8.57}$$

$\diamondsuit$

Dimostrazione. Il fatto che $I_{X,f}$ (definito localmente dal secondo membro di (8.55)) risulti essere un campo scalare ben definito su tutto $j^1(\mathbb{V}^{n+1})$ segue subito dal fatto che la 1-forma di Poincaré -Cartan è globalmente definita e che f è un campo scalare. Quest'ultimo fatto si prova per via diretta verificando che cambiando coordinate nell'atlante delle carte locali naturali adattate alle fibre, le componenti del secondo membro di (8.56) si trasformano come le componenti di vettore covariante. Lasciamo la verifica di ciò al lettore. Nella verifica si deve tenere conto del fatto che $\mathscr{L}$ è un campo scalare su $j^1(\mathbb{V}^{n+1})$.
Lavoriamo in una carta locale $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ e verifichiamo che $Z(I_{X,f}) = 0$ in tale

carta. Questo è sufficiente a provare la tesi dato che i sistemi di coordinate naturali formano un atlante di $j^1(\mathbb{V}^{n+1})$. Abbiamo che:

$$Z(I_{X,f}) = Z\left(\sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k}(X_0^k - X^t\,\dot{q}^k) + X^t\mathscr{L} - f\right)$$

$$= \sum_{k=1}^{n} Z\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k} X_0^k\right) - Z(f) - \sum_{k=1}^{n} X^t Z\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\,\dot{q}^k\right) + X^t Z(\mathscr{L}) =$$

$$\sum_{k=1}^{n} Z\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right) X_0^k + \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k} Z\left(X_0^k\right) - Z(f) - \sum_{k=1}^{n} X^t Z\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right)\dot{q}^k - \sum_{k=1}^{n} X^t \frac{\partial \mathscr{L}}{\partial \dot{q}^k} Z\left(\dot{q}^k\right) + X^t Z(\mathscr{L})$$

dove, in base alla proposizione 8.5, abbiamo usato il fatto che $Z(X^t) = 0$ e anche $Z(h \cdot g) = h \cdot Z(g) + g \cdot Z(h)$ per ogni coppia di funzioni differenziabili $f, g : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$, se $\cdot$ indica il prodotto punto per punto. Usando le equazioni di Eulero-Lagrange:

$$Z\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right) = \frac{\partial \mathscr{L}}{\partial q^k}$$

e la condizione (8.44) troviamo allora che:

$$Z(I_{X,f}) = \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial q^k} X_0^k + \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k} X^k - Z(f) - X^t\left(\sum_{k=1}^{n} \dot{q}^k \frac{\partial \mathscr{L}}{\partial q^k} + \sum_{k=1}^{n} z^k \frac{\partial \mathscr{L}}{\partial \dot{q}^k} - Z(\mathscr{L})\right),$$

dove:

$$Z = \frac{\partial}{\partial t} + \sum_k \dot{q}^k \frac{\partial}{\partial q^k} + \sum_k z^k \frac{\partial}{\partial \dot{q}^k}\,.$$

Dato che:

$$\sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial q^k} X_0^k + \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k} X^k = X(\mathscr{L}) - X^t \frac{\partial \mathscr{L}}{\partial t}$$

e

$$\sum_{k=1}^{n} \dot{q}^k \frac{\partial \mathscr{L}}{\partial q^k} + \sum_{k=1}^{n} z^k \frac{\partial \mathscr{L}}{\partial \dot{q}^k} = Z(\mathscr{L}) - \frac{\partial \mathscr{L}}{\partial t}\,,$$

l'espressione trovata sopra per $Z(I_{X,f})$ si può riscrivere come:

$$Z(I_{X,f}) = X(\mathscr{L}) - Z(\mathscr{L}) - X^t \frac{\partial \mathscr{L}}{\partial t} + X^t \frac{\partial \mathscr{L}}{\partial t}\,,$$

che è nulla per ipotesi. □

Osservazioni 8.4.
(1) Questa versione del teorema di Nöther è più sofisticata della versione data nel teorema 8.1

perché qui l'integrale primo è definito globalmente, ma è anche più generale dato che la funzione f che appare nella condizione di simmetria debole (8.54) può , localmente, dipendere dalle coordinate $\dot{q}^k$ a differenza del caso elementare trattato nel teorema 8.1. Infine la condizione (8.43), automaticamente soddisfatta nelle nostre ipotesi, generalizza generalizza ed include la richiesta di preservazione delle fibre del gruppo ad un parametro generato da X. Questa richiesta si scrive esplicitamente $\langle X, dT\rangle = 0$ che è un sottocaso banale della (8.43). Nel caso generale quindi, il gruppo ad un parametro di diffeomorfismi locali generati dal campo vettoriale X non preserva la fibra, ma muove anche nel tempo i punti dello spaziotempo degli atti di moto.
(2) Abbandonando la richiesta $\langle X, dT\rangle = 0$, cioé che il gruppo ad un parametro generato da X preservi le fibre di $j^1(\mathbb{V}^{n+1})$, non c'è più la garanzia che l'azione di tale gruppo su una sezione di $j^1(\mathbb{V}^{n+1})$ (cioè su una curva in tale spazio che può essere rappresentata in funzione del tempo e quindi pensata come un moto del sistema per qualche vettore dinamico Z, non necessariamente quello indotto da $\mathscr{L}$) venga trasformata in una sezione (e quindi in un possibile moto del sistema). Dal punto di vista fisico alcune simmetrie della lagrangiana $\mathscr{L}$ potrebbero non avere quindi significato fisico, anche se sarebbero comunque associate, tramite il teorema di Nöther, ad un integrale primo. In realtà , per ogni simmetria di $\mathscr{L}$ (anche debole) indotta da un campo vettoriale X su $j^1(\mathbb{V}^{n+1})$ che rispetta le ipotesi del teorema 8.4, c'è sempre una seconda simmetria $\tilde{X}$ che rispetta le ipotesi del teorema 8.4 inducendo lo stesso integrale primo di X, ma che in più soddisfa anche $\langle \tilde{X}, dT\rangle = 0$ e pertanto il gruppo ad un parametro di diffeomorfismi locali generato da $\tilde{X}$ trasforma sezioni di $j^1(\mathbb{V}^{n+1})$ in sezioni di $j^1(\mathbb{V}^{n+1})$. La dimostrazione è data nella seguente proposizione.

Proposizione 8.6. *In riferimento al teorema di Nöther 8.4, sia X un campo vettoriale su $j^1(\mathbb{V}^{n+1})$ che soddisfa (8.53) e (8.54). Il campo vettoriale su $j^1(\mathbb{V}^{n+1})$:*

$$\tilde{X} := X - \langle X, dT\rangle Z \ ,$$

soddisfa ancora (8.53) e (8.54), con f rimpiazzata da $\tilde{f} := f - \langle X, dT\rangle \mathscr{L}$, soddisfa la condizione di preservazione delle fibre: $\langle \tilde{X}, dT\rangle = 0$ ovunque su $j^1(\mathbb{V}^{n+1})$ e produce lo stesso integrale primo prodotto da X. $\diamondsuit$

Dimostrazione. La proprietà $\langle \tilde{X}, dT\rangle = 0$ è ovvia per costruzione, mentre vale

$$\tilde{X}(\mathscr{L}) = X(\mathscr{L}) - \langle X, dT\rangle Z(\mathscr{L}) = Z(f) - Z(\langle X, dT\rangle \mathscr{L}) = Z(f - \langle X, dT\rangle \mathscr{L}) \ ,$$

dove abbiamo usato il fatto che $Z(\langle X, dT\rangle) = 0$ come dimostrato nella proposizione 8.5. Tenendo conto dello stesso fatto ed anche della struttura del vettore Z segue che $\langle X, dT\rangle Z$ soddisfa l'ipotesi (8.43). Per linearità , tale ipotesi sarà soddisfatta anche da $\tilde{X}$. Infine, tenendo conto del fatto che $\langle Z, \omega_k\rangle = 0$ si ha:

$$\langle X, \Omega_{\mathscr{L}}\rangle - f = \langle \tilde{X}, \Omega_{\mathscr{L}}\rangle - f + \langle X, dT\rangle \mathscr{L} = \langle \tilde{X}, \Omega_{\mathscr{L}}\rangle - \tilde{f} \ .$$

Questo conclude la dimostrazione. $\square$

8.5.2 Il vettore di Runge-Lenz dal teorema di Nöther.

Illustriamo qui un caso interessante di *simmetria dinamica* e di corrispondente integrale primo dovuto al teorema di Nöther nelle forma generale data sopra. Consideriamo un sistema inerziale $\mathscr{I}$ e in esso un punto materiale P di massa m e vettore posizione $\mathbf{x} = P - O$ riferito a $\mathscr{I}$, sottoposto alla forza centrale coulombiana con centro dato dall'origine O di coordinate cartesiane ortonormali solidali con $\mathscr{I}$. Avremo dunque un energia potenziale:

$$\mathscr{U}(\mathbf{x}) = -\frac{mg}{||\mathbf{x}||}\,,$$

dove $g \neq 0$ è una costante. La lagrangiana del sistema, usando le componenti di $\mathbf{x}$ come coordinate libere, sarà allora data da:

$$\mathscr{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} m \dot{\mathbf{x}}^2 + \frac{mg}{||\mathbf{x}||}\,.$$

Le equazioni del moto risultano essere banalmente:

$$m\frac{d^2 x^i}{dt^2} = \frac{mg}{||\mathbf{x}||^3} x^i\,, \quad \text{per i=1,2,3.}$$

Dato che questa lagrangiana è evidentemente invariante per rotazioni attorno ad ogni asse $\mathbf{n}$ uscente da O, in base al teorema di Nöther (vedi esempio 8.5.1) si conserverà nel tempo il momento angolare $\mathbf{\Gamma}_O|_{\mathscr{I}} = \mathbf{x}(t) \wedge m\mathbf{v}(t)$. Di conseguenza, se $\mathbf{\Gamma}_O|_{\mathscr{I}} \neq \mathbf{0}$, il moto avverrà sul piano passante per O e perpendicolare a $\mathbf{\Gamma}_O|_{\mathscr{I}}$. Assumendo $\mathbf{\Gamma}_O|_{\mathscr{I}} \neq \mathbf{0}$, ruotiamo gli assi in modo che l'asse $\mathbf{e}_3$ sia diretto lungo $\mathbf{\Gamma}_O|_{\mathscr{I}}$. La lagrangiana sopra scritta ristretta al piano $z = 0$ diventa:

$$\mathscr{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} m \dot{\mathbf{x}}^2 + \frac{mg}{||\mathbf{x}||}\,.$$

dove però ora $\mathbf{x} = x^1\,\mathbf{e}_1 + x^2\,\mathbf{e}_2$ e le equazioni del moto sono di conseguenza:

$$m\frac{d^2 x^i}{dt^2} = \frac{mg}{||\mathbf{x}||^3} x^i\,, \quad \text{per i=1,2.}$$

In altre parole la lagrangiana ristretta al piano $x^3 = 0$ produce i moti della lagrangiana iniziale che sono confinati al piano $x^3 = 0$. In coordinate polari piane, sul piano $x^3 = 0$ la nuova lagrangiana può essere riscritta equivalentemente come:

$$\mathscr{L} = \frac{m}{2}\left(r^2\dot{\varphi}^2 + \dot{r}^2\right) + \frac{mg}{r}\,. \tag{8.58}$$

Useremo d'ora in poi la lagrangiana ridotta (8.58) limitandoci a studiare i moti del punto P confinati al piano $z = 0$ (con momento angolare iniziale non nullo). Questa lagrangiana non sembra avere simmetrie evidenti che possano produrre ulteriori integrali primi oltre a quello già ottenuto (usando la lagrangiana completa) ed oltre l'integrale primo di Jacobi (l'energia meccanica del

sistema). Vogliamo provare che in realtà esiste una *simmetria dinamica* che produce un nuovo integrale primo detto *vettore di Runge-Lenz*. Tale vettore è ben noto nella dinamica classica ma la sua esistenza non può essere facilmente spiegata in base a principio generali nelle formulazioni elementari della meccanica.
Consideriamo il campo vettoriale $\mathbf{X}$ su $j^1(\mathbb{V}^{3+1})$ che ha la struttura, per un fissato $\alpha \in \mathbb{R}$:

$$\mathbf{X} = r^2\dot{\varphi}\sin(\varphi+\alpha)\frac{\partial}{\partial r} + r\dot{\varphi}\cos(\varphi+\alpha)\frac{\partial}{\partial \varphi} + Z(X_0^r)\frac{\partial}{\partial \dot{r}} + Z(X_0^\varphi)\frac{\partial}{\partial \dot{\varphi}}\,, \tag{8.59}$$

dove abbiamo usato le notazioni della sezione precedente:

$$X_0^r := r^2\dot{\varphi}\sin(\varphi+\alpha)\,, \quad \text{e} \quad X_0^\varphi := r\dot{\varphi}\cos(\varphi+\alpha)\,,$$

e dove il vettore dinamico Z associato alla lagrangiana risulta essere, dopo un facile calcolo:

$$Z = \frac{\partial}{\partial t} + \dot{r}\frac{\partial}{\partial r} + \dot{\varphi}\frac{\partial}{\partial \varphi} + \left(r\dot{\varphi}^2 - \frac{mg}{r}\right)\frac{\partial}{\partial \dot{r}} - \frac{2\dot{r}\dot{\varphi}}{r}\frac{\partial}{\partial \dot{\varphi}} \tag{8.60}$$

Il calcolo esplicito mostra che:

$$\mathbf{X}(\mathscr{L}) = Z(mg\cos(\varphi+\alpha))\,. \tag{8.61}$$

Siamo dunque nelle ipotesi del teorema 8.4, e più precisamente nel caso con $\langle \mathbf{X}, dT\rangle = 0$, per cui ci sarà un integrale primo $I_{\mathbf{X}}$ associato alla simmetria debole indotta da $\mathbf{X}$. Si osservi che il gruppo ad un parametro di diffeomorfismi locali generato da $\mathbf{X}$ non è del tipo trattato nelle versioni elementari del teorema di Nöther (teorema 8.1), dato che le componenti X_0^r e X_0^φ dipendono *anche* dalle coordinate puntate: siamo dunque alla prese con una genuina simmetria dinamica. Il calcolo esplicito fornisce:

$$I = (\sin\alpha)\left[m\dot{r}\dot{\varphi}r^2\cos\varphi - (mr^3\dot{\varphi}^2 - mg)\sin\varphi\right] + (\cos\alpha)\left[m\dot{r}\dot{\varphi}r^2\sin\varphi + (mr^3\dot{\varphi}^2 - mg)\cos\varphi\right]\,.$$

Si osservi che abbiamo un integrale primo per ogni $\alpha \in [0, 2\pi]$. La classe completa degli integrali primi si ottiene dalle combinazioni lineari dei due integrali primi per $\alpha = 0$ e $\alpha = \pi/2$ rispettivamente. In particolare il seguente vettore, detto *vettore di Runge-Lenz* porta le stesse informazioni della classe degli integrali primi suddetti:

$$\mathbf{K} := \left[m\dot{r}\dot{\varphi}r^2\sin\varphi + (mr^3\dot{\varphi}^2 - mg)\cos\varphi\right]\,\mathbf{e}_1 + \left[-m\dot{r}\dot{\varphi}r^2\cos\varphi + (mr^3\dot{\varphi}^2 - mg)\sin\varphi\right]\,\mathbf{e}_2\,.$$

Se ora teniamo conto del fatto che nelle nostre ipotesi: $m\dot{\varphi}r^2\,\mathbf{e}_3 = \mathbf{\Gamma}_O|_{\mathscr{I}}$, si verifica facilmente che:

$$\mathbf{K} = \frac{\mathbf{p}|_{\mathscr{I}} \wedge \mathbf{\Gamma}_O|_{\mathscr{I}}}{m} - mg\frac{\mathbf{x}}{||\mathbf{x}||}\,,$$

dove $\mathbf{p}|_{\mathscr{I}} = m\mathbf{v}|_{\mathscr{I}}$ è l'impulso (diverso istante per istante) del punto materiale. Il vettore $\mathbf{K}$ è una costante del moto in quanto le sue componenti sono tali (sono integrali primi). Nel caso in cui i moti della particella siano chiusi e quindi ellissi (in questo caso la forza coulombiana deve essere attrattiva e quindi $g > 0$), quando il punto materiale attraversa l'asse maggiore dell'ellisse accade che $\mathbf{p}|_{\mathscr{I}} \wedge \mathbf{\Gamma}_O|_{\mathscr{I}}$ e $\mathbf{x}$ sono entrambi diretti lungo tale asse. Dato che $\mathbf{K}$ è costante nel tempo, esso sarà dunque diretto lungo l'asse maggiore dell'ellisse ad ogni istante. Nel caso il moto avvenga su una circonferenza, si verifica facilmente che $\mathbf{K} = 0$.

8.5.3 L'integrale primo di Jacobi come conseguenza del teorema di Nöther.

L'esistenza dell'integrale primo di Jacobi non è sicuramente conseguenza del teorema di Nöther nella forma elementare data nel teorema 8.1, in quanto non è essere associato ad un gruppo ad un parametro di diffeomorfismi locali che preservano le fibre di $j^1(\mathbb{V}^{n+1})$. Mostriamo ora è possibile dedurre l'esistenza dell'integrale primo di Jacobi dal teorema di Nöther nella versione generale data nel teorema 8.4. Mettiamoci nelle stesse ipotesi del teorema di Jacobi 8.3 e sia Z il vettore dinamico associato alla lagrangiana $\mathscr{L}$. Consideriamo il sistema di coordinate naturali $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ definite sull'aperto U nello spaziotempo degli atti di moto $j^1(\mathbb{V}^{n+1})$ su cui valgono le ipotesi del teorema di Jacobi rispetto al il campo vettoriale su $j^1(\mathbb{V}^{n+1})$:

$$X := \frac{\partial}{\partial t} \, .$$

Questo campo vettoriale soddisfa per costruzione le ipotesi del teorema 8.4 (in particolare con $\langle X, dT \rangle = 1$ costantemente). Inoltre nelle ipotesi del teorema di Jacobi:

$$X(\mathscr{L}) = 0 \, ,$$

per cui la funzione f che appare nella tesi del teorema 8.4 è nulla. Sappiamo allora che deve esserci un integrale primo $\mathscr{H}_X := -I_{X,0}$ che, *nelle coordinate considerate* ha la forma:

$$\mathscr{H}_X := \sum_k \frac{\partial \mathscr{L}}{\partial \, \dot{q}^k} (\, \dot{q}^k - X_0^k) - \mathscr{L} = \sum_k \frac{\partial \mathscr{L}}{\partial \, \dot{q}^k} \, \dot{q}^k - \mathscr{L} \, ,$$

dato che nelle coordinate considerate vale $X_0^k = 0$ identicamente. Abbiamo ritrovato l'integrale primo di Jacobi $\mathscr{H} = \mathscr{H}_X$. Si deve osservare che il campo $\frac{\partial}{\partial t}$ è definito solo in un aperto di $j^1(\mathbb{V}^{n+1})$, per cui l'integrale primo trovato ha validità solo locale. Tuttavia la procedura si può generalizzare globalmente se il campo X è definito globalmente.

Capitolo 9

Fondamenti di Meccanica Hamiltoniana.

In questo capitolo introduciamo la *formulazione di Hamilton* della meccanica classica. Si tratta di una rielaborazione della meccanica lagrangiana che ha avuto una notevole influenza nello sviluppo teorico di vari campi della fisica (moderna). Nella formulazione di Hamilton, l'oggetto matematico fondamentale, che gioca un ruolo simile a quello giocato dalla lagrangiana, $\mathscr{L}$ nella formulazione di Lagrange portando tutte le informazioni sulla dinamica del sistema fisico in esame, è la funzione hamiltoniana, $\mathscr{H}$, che abbiamo introdotto con il teorema di Jacobi. Come sappiamo essa è legata all'energia meccanica e coincide con essa in determinate situazioni. Quindi, al contrario della lagrangiana, l'hamiltoniana $\mathscr{H}$ è , sotto certe condizioni una grandezza direttamente legata all'esperienza e che entra in gioco *direttamente* in alcuni fenomeni fisici. Questo fatto ha avuto una certa rilevanza nello sviluppo della meccanica statistica classica in cui le grandezze macroscopiche all'equilibrio termodinamico possono essere solo funzione degli integrali primi del sistema. In particolare l'integrale primo che riveste maggiore importanza è l'energia meccanica. Le distribuzioni della meccanica statistica sono infatti descritte da funzioni dell'hamiltoniana del sistema. Il risultato vale anche in meccanica statistica quantistica.
Nella formulazione hamiltoniana le equazioni del moto, equivalenti a quelle di Eulero-Lagrange espresse nelle variabili lagrangiane q^k, $\dot{q}^k$, sono invece espresse in termini di *variabili hamiltoniane* anche dette *variabili canoniche*: q^k, p_k. Le equazioni del moto, ma anche tutta la struttura della teoria, ha una forma molto più simmetrica delle equazioni di Eulero-Lagrange e della teoria collegata, quando si scambia il ruolo dei due tipi di variabili. Questo punto di vista si è rivelato decisivo nel passaggio dalla meccanica classica alla meccanica quantistica, che ha ereditato dalla meccanica hamiltoniana gran parte del linguaggio e dell'impostazione.

9.1 Lo spaziotempo delle fasi e le equazioni di Hamilton.

L'ambiente nel quale si imposta la formulazione di Hamilton è lo *spaziotempo delle fasi.* Vediamo euristicamente come si arriva alla definizione di esso. Nella sezione 8.1.1 abbiamo introdotto,

per un sistema fisico descritto su $j^1(\mathbb{V}^{n+1})$ dalla lagrangiana $\mathscr{L}$, il *momento coniugato* alla coordinata q^k:

$$p_k(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) := \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial \dot{q}^k} .$$

L'idea centrale della formulazione di Hamilton è di usare, per formulare il problema del moto, le coordinate $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ invece che quelle lagrangiane $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$. In questo senso le coordinate $(p_1, \ldots, p_n)$ vengono pensate come *indipendenti* dalle q^k e vengono utilizzate *al posto* delle coordinate $\dot{q}^k$. Lo spazio (varietà $2n+1$ dimensionale) nel quale le coordinate $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ sono coordinate locali naturali è appunto lo *spaziotempo delle fasi*, che gioca un ruolo simmetrico rispetto allo spaziotempo degli atti di moto.
Se cambiamo sistemi di coordinate in $j^1(\mathbb{V}^{n+1})$ in modo tale che le trasformazioni di coordinate siano quelle che ben conosciamo:

$$t' = t + c , \tag{9.1}$$

$$q'^k = q'^k(t, q^1, \ldots, q^n) , \tag{9.2}$$

$$\dot{q}'^k = \sum_{h=1}^{n} \frac{\partial q'^k}{\partial q^h} \dot{q}^h + \frac{\partial q'^k}{\partial t} , \tag{9.3}$$

la lagrangiana soddisferà la seguente proprietà , essendo un campo scalare:

$$\frac{\partial \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)}{\partial \dot{q}'^h} = \sum_{r=1}^{n} \frac{\partial \dot{q}^r}{\partial \dot{q}'^h} \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial \dot{q}^r}$$

che può essere riscritta come:

$$\frac{\partial \mathscr{L}(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)}{\partial \dot{q}'^h} = \sum_{r=1}^{n} \frac{\partial q^r}{\partial q'^h} \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial \dot{q}^r} .$$

Se teniamo conto dell'identificazione della derivata parziale di $\mathscr{L}$ rispetto a $\dot{q}^k$ con p_k, dobbiamo concludere che eseguire il cambio di coordinate naturali su $j^1(\mathbb{V}^{n+1})$ tramite le equazioni (9.1), (9.2) e (9.3) è equivalente a cambiare coordinate $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ secondo le leggi di trasformazione seguenti:

$$\begin{aligned} t' &= t + c , \\ q'^h &= q'^h(t, q^1, \ldots, q^n) , \\ p'_h &= \sum_{r=1}^{n} \frac{\partial q^r}{\partial q'^h} p_r . \end{aligned}$$

Questo risultato giustifica euristicamente le leggi di trasformazione tra le coordinate naturali sullo spazio delle fasi che assumeremo d'ora in poi. Tuttavia deve essere *ben chiaro* che le coordinate p_k si identificano con i momenti coniugati solo *dopo* che è stata fissata una lagrangiana. Lo spaziotempo delle fasi è invece pensato come ente astratto indipendente dalla scelta della lagrangiana del sistema.

9.1.1 Lo spaziotempo delle Fasi $F(\mathbb{V}^{n+1})$.

Possiamo dare la seguente definizione.

Definizione 9.1. **(Spaziotempo delle Fasi, variabili coniugate).** Lo **spaziotempo delle fasi** $F(\mathbb{V}^{n+1})$, sul quale si descrive l'evoluzione del sistema fisico $\mathfrak{S}$, è una varietà differenziabile di dimensione $2n+1$ che, dal punto di vista insiemistico, è l'unione sottovarietà (embedded) $2n$-dimensionali, $\mathbb{F}_t$, dove t corre sull'asse del tempo assoluto. Ogni spazio $\mathbb{F}_t$ è detto **spazio delle fasi al tempo** t ed i suoi punti sono detti **punti rappresentativi del sistema $\mathfrak{S}$ (al tempo t)**.
$F(\mathbb{V}^{n+1})$ è dotato di un atlante le cui carte locali sono dette sistemi di **coordinate locali naturali (adattate alle fibre) di** $F(\mathbb{V}^{n+1})$, sono individuati da coordinate locali che della forma $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ in cui t è il tempo assoluto a meno di una costante additiva, $(t, q^1, \ldots, q^n)$, variabile in un aperto di $\mathbb{R}^n$, è associato a un sistema di coordinate locali naturali su $\mathbb{V}^{n+1}$ e $(p_1, \ldots, p_n)$ varia in tutto $\mathbb{R}^n$. La coordinata p_j viene detta **variabile coniugata della coordinata** q^j e viceversa. All'interno di tale atlante le leggi di trasformazione tra due sistemi di coordinate sono trasformazioni differenziabili, invertibili e con inversa differenziabile, della forma:

$$t' = t + c\,, \tag{9.4}$$

$$q'^k = q'^k(t, q^1, \ldots, q^n)\,, \tag{9.5}$$

$$p'_k = \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} p_h\,. \tag{9.6}$$

$\diamondsuit$

Osservazioni 9.1.
(1) Un *punto rappresentativo del sistema* $\mathfrak{S}$ *all'istante* t è quindi determinato, in un sistema di coordinate naturali del tipo detto, dall'assegnazione di una precisa $2n+1$-pla di numeri $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$. Si osservi che tale $2n+1$-pla individua anche la configurazione del sistema al tempo t, in $\mathbb{Q}_t$, attraverso il vettore riga $(q^1, \ldots, q^n)$
(2) Quando è assegnata una lagrangiana $\mathscr{L}$ del sistema $\mathfrak{S}$, le coordinate p_j si identificano con i *momento coniugati* come detto sopra. Con questa identificazione, come proveremo tra poco, ogni $2n+1$-pla $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ individua biunivocamente una $2n+1$-pla $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ e pertanto determina posizione e velocità di ogni punto materiale di $\mathfrak{S}$ all'istante t in ogni sistema di riferimento $\mathscr{I}$. In questo senso fissare un punto rappresentativo del sistema all'istante t è equivalente posizione e velocità di ogni punto materiale di $\mathfrak{S}$ all'istante t.

Affrontiamo più rigorosamente la questione della corrispondenza, mediata da una fissata lagrangiana $\mathscr{L}$, tra le coordinate locali naturali sullo spaziotempo delle fasi e quelle sullo spaziotempo degli atti di moto. Fissiamo una carta locale naturale su $\mathbb{V}^{n+1}$, con coordinate $(t, q^1, \ldots, q^n)$. Questa carta si può sempre completare in due modo distinti: come una carta locale naturale

di $j^1(\mathbb{V}^{n+1})$ passando alle coordinate $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ dove $(t, q^1, \ldots, q^n) \in U$ aperto di $\mathbb{R}^{n+1}$ e $(\dot{q}^1, \ldots, \dot{q}^n) \in \mathbb{R}^n$, oppure come una carta locale naturale su $F(\mathbb{V}^{n+1})$ passando alle coordinate $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ dove $(t, q^1, \ldots, q^n) \in U$ aperto di $\mathbb{R}^{n+1}$ (lo stesso di prima) e $(p_1, \ldots, p_n) \in \mathbb{R}^n$. Si consideri infine la trasformazione di coordinate sui domini detti:

$$
\begin{aligned}
t &= t\,, && (9.7)\\
q^k &= q^k\,, \quad \text{per } k = 1, \ldots, n\,, && (9.8)\\
p_k &= \frac{\partial \mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)}{\partial \dot{q}^k}\,, \quad \text{per } k = 1, \ldots, n\,. && (9.9)
\end{aligned}
$$

L'ultima trasformazione, la (9.9), *quando considerata a* $(t, q^1, \ldots, q^n)$ *fissati* e quindi pensata come una trasformazione da $\mathbb{R}^n \ni (\dot{q}^1, \ldots, \dot{q}^n)$ in $\mathbb{R}^n \ni (p_1, \ldots, p_n)$, è detta **trasformazione di Legendre**. Similmente si dice che l'hamiltoniana, *espressa in funzione delle variabili di Hamilton connessa a quelle di Lagrange tramite la trasformazione di Legendre*, è la funzione **trasformata di Legendre** della Lagrangiana.
Vogliamo ora provare che la trasformazione (9.9) è *biettiva* sui domini detti, oltre ad essere differenziabile con inversa differenziabile, per tutte le lagrangiane che abbiamo incontrato fino ad ora.
Tutte le lagrangiane che abbiamo incontrato nella formulazione lagrangiana della meccanica, includendo lagrangiane contenenti potenziali generalizzati (vedi la sezione 7.2.2) sono quadratiche nelle $\dot{q}^k$:

$$
\mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \sum_{k,h=1}^{n} a_{kh}(t, q^1, \ldots, q^n)\dot{q}^k\dot{q}^h + \sum_{k=1}^{n} \beta_k(t, q^1, \ldots, q^n)\dot{q}^k + \gamma(t, q^1, \ldots, q^n) \tag{9.10}
$$

dove la matrice simmetrica di coefficienti $a_{hk}(t, q^1, \ldots, q^n)$ è *strettamente definita positiva* e tutte le funzioni delle variabili $(t, q^1, \ldots, q^n)$ sono di classe C^1 almeno. In questo caso la trasformazione di Legendre si riduce a:

$$
p_k = \sum_{h=1}^{n} 2a_{kh}(t, q^1, \ldots, q^n)\,\dot{q}^h + \beta_k(t, q^1, \ldots, q^n)\,. \tag{9.11}
$$

Dato che la matrice simmetrica dei coefficienti a_{kh} è strettamente definita positiva deve anche essere invertibile. Pertanto la trasformazione di Legendre scritta sopra, per ogni scelta di $t, q^1, \ldots, q^n$ fissati, si riduce ad una semplice trasformazione biettiva da $\mathbb{R}^n$ a $\mathbb{R}^n$. Di conseguenza la trasformazione di coordinate (9.7)-(9.9) è una trasformazione biettiva (sui domini delle coordinate precedentemente indicati) di classe C^1 almeno.

Esempi 9.1.
1. Consideriamo un sistema di N punti materiali, $P_1, \ldots, P_N$ di masse $m_1, \ldots, m_N$ rispettivamente, non sottoposti a vincoli ma interagenti tramite forze dedotte da un potenziale totale

$\mathscr{V}|_{\mathscr{I}}$. $\mathscr{I}$ è un sistema di riferimento inerziale nel quale i punti dono individuati da vettori posizione $\mathbf{x}_1, \ldots, \mathbf{x}_n$. La lagrangiana del sistema sarà allora data da, usando le componenti di tutti i vettori $\mathbf{x}_i$ come coordinate libere:

$$\mathscr{L}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_n) = \sum_{i=1}^{N} \frac{m_i}{2} \dot{\mathbf{x}}_i^2 + \mathscr{V}|_{\mathscr{I}}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N) \, .$$

In questo caso, riorganizzando le coordinate p_k a tre a tre come per i vettori posizione, abbiamo che:

$$\mathbf{p}_i = \nabla_{\dot{\mathbf{x}}_i} \mathscr{L}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_N) = m_i \dot{\mathbf{x}}_i \, .$$

Possiamo concludere che per sistemi fisici di punti materiali non sottoposti a vincoli, in cui le forze sono conservative o dedotte da un potenziale *non generalizzato*, se si usano come coordinate libere le componenti dei vettori posizione delle particelle, le coordinate p_k associate si identificano con le componenti degli impulsi delle particelle su ogni moto del sistema.

2. Un caso meno banale del precedente si ha quando il sistema di forze che agisce sui punti materiali è indotto da un potenziale generalizzato. Consideriamo il caso di una particella carica di massa m e carica e immersa in un campo elettromagnetico assegnato tramite i potenziali elettromagnetici $\mathbf{A}$ e φ. Sappiamo dalla sezione 7.2.1 che la sua lagrangiana è , in un riferimento inerziale ed usando le componenti del vettore posizione della particella come coordinate libere:

$$\mathscr{L}(t, \mathbf{x}, \dot{\mathbf{x}}) = \frac{m}{2} \dot{\mathbf{x}}^2 - e\varphi(t, \mathbf{x}) + \frac{e}{c} \mathbf{A}(t, \mathbf{x}) \cdot \dot{\mathbf{x}} \, .$$

In questo caso si ha subito:

$$\mathbf{p} = m\dot{\mathbf{x}} + \frac{e}{c} \mathbf{A}(t, \mathbf{x}) \, ,$$

e quindi i momenti coniugati associati alle tre coordinate spaziali non coincidono più (sul moto) con le componenti dell'impulso della particella.

3. Consideriamo una particella di massa m, vincolata a stare su una sfera liscia di raggio $R > 0$ e centro O in quiete in un riferimento inerziale $\mathscr{I}$, e sottoposta ad una forza dovuta ad un potenziale $\mathscr{V}|_{\mathscr{I}}$. Usiamo coordinate angolari sferiche θ e ϕ (solidali con $\mathscr{I}$) per individuare la particella P sulla sfera. Sappiamo che (vedi capitolo 1), se $\mathbf{v}$ è la velocità della particella in $\mathscr{I}$, in coordinate polari sferiche vale:

$$\mathbf{v} = \dot{r} \, \mathbf{e}_r + r\dot{\theta} \, \mathbf{e}_\theta + r\dot{\phi} \sin\theta \, \mathbf{e}_\varphi \, .$$

Adattando al nostro caso ($r = R$ costantemente) la formula di sopra, possiamo arguire che la lagrangiana della particella è :

$$\mathscr{L}(t, \theta, \phi, \dot{\theta}, \dot{\phi}) = \frac{mR^2}{2} \left(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta \right) + \mathscr{V}|_{\mathscr{I}}(t, \theta, \phi) \, .$$

Segue immediatamente che i due momenti coniugati p_ϕ e ϕ_θ hanno la struttura:

$$p_\phi = mR^2 \dot{\phi} \sin^2\theta \, , \quad p_\theta = mR^2 \dot{\theta} \, .$$

È interessante notare che il significato fisico, di p_ϕ , sui moti della particella, è quello di *componente lungo l'asse z del momento angolare della particella* (valutato in $\mathscr{I}$ rispetto al polo O). Infatti è

$$\Gamma_O|_{\mathscr{I}} = m(P-O)\wedge \mathbf{v} = mR\,\mathbf{e}_r \wedge \left(R\dot\theta\,\mathbf{e}_\theta + R\dot\phi\sin\theta\,\mathbf{e}_\varphi\right) = mR^2\left(\dot\theta\,\mathbf{e}_\varphi - \dot\phi\sin\theta\,\mathbf{e}_\theta\right)\ ,$$

e quindi:

$$\Gamma_O|_{\mathscr{I}}\cdot\ \mathbf{e}_z = mR^2\left(\dot\theta\,\mathbf{e}_\varphi\cdot\ \mathbf{e}_z - \dot\phi\sin\theta\,\mathbf{e}_\theta\cdot\ \mathbf{e}_z\right) = 0 + mR^2\dot\phi\sin^2\theta = p_\phi\ .$$

Si osservi che lo stesso risultato riguardante il significato fisico di p_ϕ si ha anche studiando una particella non sottoposta a vincoli (ma sottoposta ad un potenziale), usando coordinate libere date dalle tre coordinate polari sferiche della particella in un riferimento inerziale.

9.1.2 Le equazioni di Hamilton.

Vogliamo ora determinare quali sono le equazioni che corridpondono, nello spaziotempo delle fasi, alle equazioni di Eulero-Lagrange. Abbiamo la seguente fondamentale proposizione che permette di giungere a tali equazioni.

D'ora in poi, per evitare di appesantire inutilmente la notazione, useremo anche la seguente notazione contratta: $(t, q,\ \dot q)$ significherà $(t, q^1, \ldots, q^n,\ \dot q^1, \ldots,\ \dot q^n)$ e similmente (t, q, p) indicherà $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$.

Proposizione 9.1. *Si consideri un sistema fisico $\mathfrak{S}$ descritto dalla lagrangiana di classe C^1 $\mathscr{L} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ con struttura (9.10) in ogni sistema di coordinate naturali di $j^1(\mathbb{V}^{n+1})$ e dove la matrice dei coefficienti a_{hk} è ovunque strettamente definita positiva.*
Riferendoci a due sistemi di coordinate locali naturali: $(t, q^1, \ldots, q^n,\ \dot q^1, \ldots,\ \dot q^n)$ su $j^1(\mathbb{V}^{n+1})$ e $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ su $F(\mathbb{V}^{n+1})$, connessi dalle trasformazioni

$$t = t\,, \tag{9.12}$$

$$q^k = q^k\,, \quad \textit{per } k = 1, \ldots, n\,, \tag{9.13}$$

$$p_k = \frac{\partial \mathscr{L}(t, q,\ \dot q)}{\partial\,\dot q^k}\,, \quad \textit{per } k = 1, \ldots, n\,, \tag{9.14}$$

valgono le seguenti identità :

$$\frac{\partial \mathscr{H}}{\partial t} = -\left.\frac{\partial \mathscr{L}}{\partial t}\right|_{(t, q,\ \dot q(t,q,p))}\,, \tag{9.15}$$

$$\frac{\partial \mathscr{H}}{\partial p_k} = \dot q^k(t, q, p)\,, \quad \textit{per } k = 1, \ldots, n \tag{9.16}$$

$$\frac{\partial \mathscr{H}}{\partial q^k} = -\left.\frac{\partial \mathscr{L}}{\partial q^k}\right|_{(t, q,\ \dot q(t,q,p))}\,, \quad \textit{per } k = 1, \ldots, n, \tag{9.17}$$

dove l'hamiltoniana $\mathscr{H}$ associata a $\mathscr{L}$ nelle coordinate dette è stata espressa nelle variabili di Hamilton:

$$\mathscr{H}(t,q,p) = \sum_{k=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}^k}\bigg|_{(t,q,\,\dot{q}(t,q,p))} \dot{q}^k\,(t,q,\,\dot{q}(t,q,p)) - \mathscr{L}(t,q,\,\dot{q}(t,q,p))\,. \tag{9.18}$$

$\diamondsuit$

Dimostrazione. Prima di tutto notiamo che per (9.9), l'equazione (9.18) può essere riscritta banalmente come:

$$\mathscr{H}(t,q,p) = \sum_{k=1}^{n} p_k\,\dot{q}^k\,(t,q,\,\dot{q}(t,q,p)) - \mathscr{L}(t,q,\,\dot{q}(t,q,p))\,. \tag{9.19}$$

Consideriamo una curva in $F(\mathbb{V}^{n+1})$ di classe C^1, parametrizzata nel tempo assoluto $t \mapsto (t, q(t), p(t))$. Ad essa corrisponderà una analoga curva in $F(\mathbb{V}^{n+1})$: $t \mapsto (t, q(t),\, \dot{q}(t, q(t), p(t)))$. Calcoliamo la derivata di $\mathscr{H}(t, q(t), p(t))$ usando (9.19). Derivando ambo i membri si trova che:

$$\frac{\partial \mathscr{H}}{\partial t} + \sum_k \frac{\partial \mathscr{H}}{\partial q^k}\frac{dq^k}{dt} + \sum_k \frac{\partial \mathscr{H}}{\partial p_k}\frac{dp_k}{dt} = \sum_{k=1}^{n} \frac{dp_k}{dt}\,\dot{q}^k + \sum_{k=1}^{n} \frac{d\,\dot{q}^k}{dt} p_k - \sum_k \frac{\partial \mathscr{L}}{\partial q^k}\frac{dq^k}{dt} - \sum_k \frac{\partial \mathscr{L}}{\partial\,\dot{q}^k}\frac{d\,\dot{q}^k}{dt} - \frac{\partial \mathscr{L}}{\partial t}.$$

I due termini:

$$\sum_{k=1}^{n} \frac{d\,\dot{q}^k}{dt} p_k - \sum_k \frac{\partial \mathscr{L}}{\partial\,\dot{q}^k}\frac{d\,\dot{q}^k}{dt}$$

potrebbero ancora essere sviluppati esplicitando le derivate $d\,\dot{q}^k/dt$, ma non conviene dato che si elidono reciprocamente, in virtù della (9.9). Tenendo conto di ciò l'identità trovata sopra si può riscrivere:

$$\frac{\partial \mathscr{H}}{\partial t} + \frac{\partial \mathscr{L}}{\partial t} + \sum_k \left(\frac{\partial \mathscr{H}}{\partial q^k} + \frac{\partial \mathscr{L}}{\partial q^k}\right)\frac{dq^k}{dt} + \sum_k \left(\frac{\partial \mathscr{H}}{\partial p_k} - \dot{q}^k\right)\frac{dp_k}{dt} = 0.$$

Dato che per ogni fissato punto (t, q, p) possiamo definire una curva, $t \mapsto (t, q(t), p(t))$, che passa per tale punto per $t = 0$ e possiamo scegliere le componenti $\frac{dp^k}{dt}|_{t=0}$ e $\frac{dq^k}{dt}|_{t=0}$ del tutto arbitrariamente, possiamo per esempio scegliere nulle tutte queste componenti. L'identità di sopra si riduce allora a:

$$\frac{\partial \mathscr{H}}{\partial t}\bigg|_{(t,q,p)} + \frac{\partial \mathscr{L}}{\partial t}\bigg|_{(t,q,p)} = 0\,, \quad \text{per ogni punto rappresentativo } (t,q,p)\,.$$

Ma allora deve essere, ovunque nella regione considerata di $F(\mathbb{V}^{n+1})$:

$$\sum_k \left(\frac{\partial \mathscr{H}}{\partial q^k} + \frac{\partial \mathscr{L}}{\partial q^k}\right)\frac{dq^k}{dt} + \sum_k \left(\frac{\partial \mathscr{H}}{\partial p_k} - \dot{q}^k\right)\frac{dp_k}{dt} = 0.$$

Scegliendo tutti i numeri $\frac{dp_k}{dt}|_{t=0}$ e $\frac{dq^k}{dt}|_{t=0}$ nulli eccetto $\frac{dp_{k_0}}{dt}|_{t=0} = 1$, per un fissato $k_0 = 1, \ldots, n$, concludiamo immediatamente che:

$$\left.\frac{\partial \mathscr{H}}{\partial p_{k_0}}\right|_{(t,q,p)} - \dot{q}^{k_0}|_{(t,q,p)} = 0\,, \quad \text{per ogni punto rappresentativo } (t,q,p)\,.$$

Dato che k_0 è arbitrario, abbiamo anche che l'identità iniziale si riduce a:

$$\sum_k \left(\frac{\partial \mathscr{H}}{\partial q^k} + \frac{\partial \mathscr{L}}{\partial q^k}\right)\frac{dq^k}{dt} = 0.$$

Reiterando la procedura arriviamo a dimostrare tutte le identità (9.15)-(9.17). □

Abbiamo come conseguenza immediata il seguente fondamentale teorema che definisce le *equazioni di Hamilton.*

Teorema 9.1. **(Equazioni di Hamilton.)** *Si consideri un sistema fisico $\mathfrak{S}$ descritto dalla lagrangiana di classe C^1, $\mathscr{L} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$, con struttura (9.10) in ogni sistema di coordinate naturali di $j^1(\mathbb{V}^{n+1})$ e dove la matrice dei coefficienti a_{hk} è ovunque strettamente definita positiva.*
Riferendoci a due sistemi di coordinate locali naturali: $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ su $j^1(\mathbb{V}^{n+1})$ e $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ su $F(\mathbb{V}^{n+1})$, connessi dalle trasformazioni

$$\begin{aligned} t &= t\,, \\ q^k &= q^k\,, \quad \textit{per } k = 1, \ldots, n\,, \\ p_k &= \frac{\partial \mathscr{L}(t, q,\ \dot{q})}{\partial\, \dot{q}^k}\,, \quad \textit{per } k = 1, \ldots, n\,. \end{aligned}$$

La curva $I \ni t \mapsto (t, q(t),\ \dot{q}(t))$ di classe C^1 soddisfa le equazioni di Eulero-Lagrange (6.44) (per qualche intervallo aperto $I \subset \mathbb{R}$):

$$\left\{\begin{array}{ll} \dfrac{d}{dt}\dfrac{\partial \mathscr{L}}{\partial \dot{q}^k} & = \dfrac{\partial \mathscr{L}}{\partial q^k}\,, \\ \dfrac{dq^k}{dt} & = \dot{q}^k(t) \end{array}\right. \quad \textit{per } k = 1, \ldots, n,$$

se e solo la curva $I \ni t \mapsto (t, q(t), p(t))$ ottenuta trasformando la curva precedente tramite le (9.12)-(9.9), soddisfa le **equazioni di Hamilton***:*

$$\left\{\begin{array}{ll} \dfrac{dp_k}{dt} & = -\dfrac{\partial \mathscr{H}}{\partial q^k}\,, \\ \dfrac{dq^k}{dt} & = \dfrac{\partial \mathscr{H}}{\partial p_k}\,, \end{array}\right. \quad \textit{per } k = 1, \ldots, n, \tag{9.20}$$

dove l'hamiltoniana $\mathscr{H}$ associata a $\mathscr{L}$ nelle coordinate dette è stata espressa nelle variabili di Hamilton (9.18).
Infine vale l'identità , se $I \ni\mapsto (t, q(t), p(t))$ è una soluzione delle equazioni di Hamilton:

$$\left.\frac{d\mathscr{H}}{dt}\right|_{(t,q(t),p(t))} = \left.\frac{\partial\mathscr{H}}{\partial t}\right|_{(t,q(t),p(t))} , \tag{9.21}$$

e quindi l'Hamiltoniana è un integrale primo se non dipende esplicitamente dal tempo. $\diamondsuit$

Dimostrazione. La prima equazione di Eulero-Lagrange, si trascrive in coordinate di Hamilton, tenendo conto che $\frac{\partial\mathscr{L}}{\partial\,\dot{q}^k} = p_k$ e usando la (9.17):

$$\frac{dp_k}{dt} = -\frac{\partial\mathscr{H}}{\partial q^k} .$$

La seconda equazione di Eulero-Lagrange, usando (9.16) si trascrive in coordinate di Hamilton:

$$\frac{dq^k}{dt} = \frac{\partial\mathscr{H}}{\partial p_k} .$$

Per costruzione è allora chiaro che una curva $I \ni t \mapsto (t, q(t),\ \dot{q}(t))$ soddisfa le equazioni di Eulero-Lagrange se e solo se $I \ni t \mapsto (t, q(t), p(t, q(t),\ \dot{q}(t))$ soddisfa le equazioni di Hamilton. Questo conclude la dimostrazione della prima parte della tesi.
L'ultima affermazione della tesi segue dal calcolo diretto tenendo conto delle equazioni di Hamilton:

$$\frac{d\mathscr{H}}{dt} = \frac{\partial\mathscr{H}}{\partial t} + \sum_k \frac{\partial\mathscr{H}}{\partial q^k}\frac{dq^k}{dt} + \sum_k \frac{\partial\mathscr{H}}{\partial p_k}\frac{dp_k}{dt} = \frac{\partial\mathscr{H}}{\partial t} + \sum_k \left(\frac{\partial\mathscr{H}}{\partial q^k}\frac{\partial\mathscr{H}}{\partial p_k} - \frac{\partial\mathscr{H}}{\partial p_k}\frac{\partial\mathscr{H}}{\partial q^k}\right) = \frac{\partial\mathscr{H}}{\partial t} .$$

□

Osservazioni 9.2.
(1) Le equazioni di Hamilton sono già in forma normale e pertanto la validità del teorema di esistenza ed unicità è assicurata purché la funzione hamiltoniana sia sufficientemente regolare.
(2) L'identità (9.21) insieme alla (9.15) implica immediatamente la prima affermazione del teorema di Jacobi 8.3.
(3) Le equazioni di Hamilton implicano immediatamente che se $\mathscr{H}$ non dipende esplicitamente da q^i allora la coordinata coniugata p_i è un integrale primo del sistema fisico.
(4) Al contrario della lagrangiana (su $j^1(\mathbb{V}^{n+1})$), la funzione di Hamilton *non è un campo scalare (su $F(\mathbb{V}^{n+1})$)*! Se, per una stessa lagrangiana $\mathscr{L} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$, $\mathscr{H}$ e $\mathscr{H}'$ sono le funzioni hamiltoniane definite su due carte locali naturali con intersezione ed associate a coordinate naturali locali $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ e $(t', q'^1, \ldots, q'^n, p'_1, \ldots, p'_n)$, tali che, sull'intersezione delle

carte vaga:

$$\begin{aligned} t' &= t + c\,, \\ q'^k &= q'^k(t, q^1, \ldots, q^n)\,, \\ p'_k &= \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} p_h\,, \end{aligned}$$

allora si ha anche:

$$\mathscr{H}'(t'(t,q,p), q'(t,q,p), p'(t,q,p)) = \mathscr{H}(t,q,p) + \sum_{k=1}^{n} \frac{\partial q'^k}{\partial t}\bigg|_{(t,q,p)} p'_k(t,q,p)\,. \tag{9.22}$$

Lo si provi per esercizio.

Il teorema 9.1 ha un'immediata ma importante conseguenza.

Corollario. *Nelle ipotesi del teorema 9.1, se una curva C^1, parametrizzata nel tempo assoluto, $\gamma : I \ni t \mapsto F(\mathbb{V}^{n+1})$ soddisfa le equazioni di Hamilton nell'aperto $V \subset F(\mathbb{V}^{n+1})$ dotato di coordinate naturali $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ con hamiltoniana $\mathscr{H}$, tale curva soddisferà le equazioni di Hamilton in ogni altro sistema di coordinate naturali, diverso dal precedente, $(t', q'^1, \ldots, q'^n, p'_1, \ldots, p'_n)$ con corrispondente hamiltoniana $\mathscr{H}'$ (associata alla stessa lagrangiana) sullo stesso V.* $\diamondsuit$

Dimostrazione. La curva γ soddisfa le equazioni di Eulero-Lagrange sull'aperto $U \subset j^1(\mathbb{V}^{n+1})$, dotato di coordinate $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$. Pertanto, per la proposizione 6.3, essa soddisferà le equazioni di Eulero-Lagrange sul sistema di coordinate $(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)$, sempre definito su U associato alle $(t', q'^1, \ldots, q'^n, p'_1, \ldots, p'_n)$ su V. Per il teorema 9.1 γ deve soddisfare le equazioni di Hamilton anche rispetto alle coordinate locali naturali $(t', q'^1, \ldots, q'^n, p'_1, \ldots, p'_n)$. $\square$

Osservazioni 9.3.* Il risultato ha un'importante conseguenza. Noi abbiamo fino ad ora costruito la teoria delle equazioni di Hamilton in un sistema di coordinate locali (naturali), sulla varietà differenziabile $F(\mathbb{V}^{n+1})$, scelto arbitrariamente. Tuttavia, in generale $F(\mathbb{V}^{n+1})$ non è ricopribile con una sola carta locale e considerando una soluzione delle equazioni di Hamilton in una carta locale (U, ψ), ci aspettiamo che tale soluzione possa arrivare fino al bordo ∂U di tale carta. Dal punto di vista fisico ci si aspetta che la soluzione sia estendibile fuori da tale carta, dato che la carta scelta non ha in generale alcun significato fisico privilegiato. Possiamo provare ad usare diversi sistemi di coordinate incollando di volta in volta le varie soluzioni, ma a priori non è detto che la teoria complessiva risulti coerente. Il corollario dimostrato ci dice che i vari ingredienti usati nella formulazione delle equazioni di Hamilton, cambiando coordinate, si comportano coerentemente. Se infatti (V, ϕ) è un'altra carta locale con $V \cap U \neq \emptyset$, la soluzione γ_U suddetta, raggiungerà un punto $\gamma_U(t_1) \in U \cap V$. Possiamo allora reimpostare il problema di Cauchy delle equazioni di Hamilton ora *trascritte nelle coordinate in* V, usando come condizioni iniziali al

tempo t_1, $\gamma_U(t_1)$ e $\dot{\gamma}_U(t_1)$. L'unica soluzione γ_V che otteniamo si estenderà in generale su V *anche fuori da* U. Le curve γ_U e γ_V, per costruzione si raccorderanno differenziabilmente in $U \cap V$. È chiaro che per questa via si ottiene alla fine una soluzione delle equazioni di Hamilton definita, potenzialmente, su *tutto* $F(\mathbb{V}^{n+1})$.

Esempi 9.2.
1. Consideriamo un sistema di N punti materiali, $P_1, \ldots, P_N$ di masse $m_1, \ldots, m_N$ rispettivamente, non sottoposti a vincoli ma interagenti tramite forze dedotte da un'energia potenziale totale $\mathscr{U}|_{\mathscr{I}}$. $\mathscr{I}$ è un sistema di riferimento inerziale nel quale i punti dono individuati da vettori posizione $\mathbf{x}_1, \ldots, \mathbf{x}_n$. La lagrangiana del sistema sarà allora data da, usando le componenti di tutti i vettori $\mathbf{x}_i$ come coordinate libere:

$$\mathscr{L}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_N) = \sum_{i=1}^{N} \frac{m_i}{2} \dot{\mathbf{x}}_i^2 - \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,.$$

L'hamiltoniana risulta avere forma:

$$\mathscr{H}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_N) = \sum_{i=1}^{N} \frac{m_i}{2} \dot{\mathbf{x}}_i^2 + \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,.$$

Come già visto abbiamo anche che:

$$\mathbf{p}_i = \nabla_{\dot{\mathbf{x}}_i} \mathscr{L}(t, \mathbf{x}_1, \ldots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_N) = m_i \dot{\mathbf{x}}_i \,.$$

da cui:

$$\dot{\mathbf{x}}_i = \frac{\mathbf{p}_i}{m_i} \,.$$

Possiamo allora esprimere l'hamiltoniana nelle variabili di Hamilton ottenendo:

$$\mathscr{H}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,.$$

Dato che nel caso in esame l'hamiltoniana non dipende esplicitamente dal tempo essa si deve conservare sulle equazioni del moto di Hamilton (in effetti l'hamiltoniana, sul moto, coincide con l'energia meccanica totale del sistema). Le equazioni di Hamilton risultano essere, per $i = 1, \ldots, N$:

$$\begin{aligned} \frac{d\mathbf{p}_i}{dt} &= (-\nabla_{\mathbf{x}_i} \mathscr{H} =) \; - \nabla_{\mathbf{x}_i} \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,, \\ \frac{d\mathbf{x}_i}{dt} &= (\; \nabla_{\mathbf{p}_i} \mathscr{H} =) \; \frac{\mathbf{p}_i}{m_i} \,. \end{aligned}$$

Le due equazioni insieme forniscono le N equazioni di Newton (per $i = 1, \ldots, N$):

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = -\nabla_{\mathbf{x}_i} \mathscr{U}|_{\mathscr{I}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \,.$$

2. Consideriamo ora il caso di una particella carica di massa m e carica e immersa in un campo elettromagnetico assegnato tramite i potenziali elettromagnetici $\mathbf{A}$ e φ. Sappiamo dalla sezione 7.2.1 che la sua lagrangiana è , in un riferimento inerziale ed usando le componenti del vettore posizione della particella come coordinate libere:

$$\mathscr{L}(t, \mathbf{x}, \dot{\mathbf{x}}) = \frac{m}{2}\dot{\mathbf{x}}^2 - e\varphi(t, \mathbf{x}) + \frac{e}{c}\mathbf{A}(t, \mathbf{x}) \cdot \dot{\mathbf{x}} \, .$$

L'Hamiltoniana vale, in questo caso, per computo diretto:

$$\mathscr{H}(t, \mathbf{x}, \dot{\mathbf{x}}) = \frac{m}{2}\dot{\mathbf{x}}^2 + e\varphi(t, \mathbf{x}) \, .$$

Notare che nella funzione hamiltoniana non appare più il potenziale magnetico: in formulazione lagrangiana l'Hamiltoniana *non* contiene l'informazione della presenza del campo magnetico. Tuttavia, come già visto:

$$\mathbf{p} = m\dot{\mathbf{x}} + \frac{e}{c}\mathbf{A}(t, \mathbf{x}) \, ,$$

e dunque:

$$\dot{\mathbf{x}} := \frac{\mathbf{p}}{m} - \frac{e}{mc}\mathbf{A}(t, \mathbf{x}) \, .$$

Riscrivendo l'hamiltoniana in funzione delle variabili di Hamilton l'informazione riguardante il campo magnetico riappare a causa del fatto che i momenti coniugati non sono semplicemente le componenti dell'impulso :

$$\mathscr{H}(t, \mathbf{x}, \mathbf{p}) = \frac{\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})\right)^2}{2m} + e\varphi(t, \mathbf{x}) \, .$$

Le equazioni di Hamilton risultano:

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= (-\nabla_{\mathbf{x}}\mathscr{H} =) \, \frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) \cdot \frac{e}{c}\nabla_{\mathbf{x}}\mathbf{A} - e\nabla\varphi \, , \\ \frac{d\mathbf{x}}{dt} &= (\ \nabla_{\mathbf{p}}\mathscr{H} =) \, \frac{\mathbf{p}}{m} - \frac{e}{mc}\mathbf{A} \, . \end{aligned}$$

Dalle due equazioni si ricava:

$$m\frac{d^2\mathbf{x}}{dt^2} = -\frac{e}{c}\frac{\partial \mathbf{A}}{\partial t} - e\nabla_x\varphi - \frac{e}{c}\left(\mathbf{v} \cdot \nabla_{\mathbf{x}}\right)\mathbf{A}(t, \mathbf{x}) + \frac{e}{c}\nabla_{\mathbf{x}}\left(\mathbf{v} \cdot \mathbf{A}\right) \, .$$

Il secondo membro non è altro che la forza di Lorentz, come visto nella sezione 7.2.1. Di conseguenza abbiamo trovato le corrette equazioni del moto di una particella carica in un campo elettromagnetico.

3. Consideriamo una particella di massa m, vincolata a stare su una sfera liscia di raggio $R > 0$ e centro O in quiete in un riferimento inerziale $\mathscr{I}$, e sottoposta ad una forza con energia potenziale $\mathscr{U}|_{\mathscr{I}}$. In coordinate angolari sferiche θ e ϕ la lagrangiana della particella è :

$$\mathscr{L}(\theta, \phi, \dot{\theta}, \dot{\phi}) = \frac{mR^2}{2}\left(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta\right) - \mathscr{U}|_{\mathscr{I}}(\theta, \phi) \, ,$$

per cui l'hamiltoniana è :

$$\mathscr{H}(\theta,\phi,\dot{\theta},\dot{\phi}) = \frac{mR^2}{2}\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + \mathscr{U}|_{\mathscr{I}}(\theta,\phi)\,.$$

Dato che i due momenti coniugati p_ϕ e ϕ_θ hanno la struttura:

$$p_\phi = mR^2\dot{\phi}\sin^2\theta\,, \quad p_\theta = mR^2\dot{\theta}\,,$$

si ha che l'hamiltoniana in variabili di Hamilton è scritta come:

$$\mathscr{H}(\theta,\phi,p_\theta,p_\phi) = \frac{p_\theta^2}{2mR^2} + \frac{p_\phi^2}{2mR^2\sin^2\theta} + \mathscr{U}|_{\mathscr{I}}(t,\theta,\phi)\,.$$

Lasciamo al lettore di scrivere esplicitamente le equazioni di Hamilton.
(4) Studiando sistemi fisici generali $\mathfrak{S}$ composti da due (o più) sottosistemi $\mathfrak{S}_1, \mathfrak{S}_2$, l'hamiltoniana ha spesso una struttura della forma, in nriferimento ad un sistema di coordinate locali naturali di $F(\mathbb{V}^{n+1})$:

$$\mathscr{H} = \mathscr{H}_1 + \mathscr{H}_2 + \mathscr{H}_I\,,$$

dove $\mathscr{H}_1$ e $\mathscr{H}_2$ sono le hamiltoniane dei due sistemi pensati come non interagenti: ognuna di tali hamiltoniane contiene *solo* coordinate riferite al corrispondente sottosistema, e $\mathscr{H}_I$ è l'*hamiltoniana d'interazione* che contiene le coordinate di *entrambi* i sottosistemi. L'esempio più semplice è dato da un sistema composto da due particelle di masse m_1 e m_2 rispettivamente, con coordinate di Hamilton $(\mathbf{x}_1, \mathbf{p}_1)$ e $(\mathbf{x}_2, \mathbf{p}_2)$ usando le stesse convenzioni che nell'esempio (1) di sopra rispetto ad un sistema di riferimento inerziale $\mathscr{I}$. In tal caso le hamiltoniane libere sono, per esempio:

$$\mathscr{H}_1 := \frac{\mathbf{p}_1^2}{2m_1}\,, \quad \mathscr{H}_2 := \frac{\mathbf{p}_2^2}{2m_2}\,,$$

mentre una possibile hamiltoniana d'interazione è quella data da una forza conservativa associata ad un'energia potenziale che dipende dalle posizioni dei due punti materiali:

$$\mathscr{H}_I := \mathscr{U}(||\mathbf{x}_1 - \mathbf{x}_2||)\,.$$

Considerando sistemi composti da più sottosistemi l'hamiltoniana d'interazione può essere spesso decomposta in hamiltoniane cosiddette: "a due corpi", "a tre corpi" e via di seguito. Per esempio con 3 particelle, si può avere:

$$\mathscr{H}_I = \mathscr{U}^{(2)}(\mathbf{x}_1,\mathbf{x}_2) + \mathscr{U}^{(2)}(\mathbf{x}_2,\mathbf{x}_3) + \mathscr{U}^{(2)}(\mathbf{x}_1,\mathbf{x}_3) + \mathscr{U}^{(3)}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)\,,$$

dove $\mathscr{U}^{(2)}$ rappresenta l'interazione (conservativa in questo caso) a due corpi e $\mathscr{U}^{(3)}$ quella a tre corpi. Nelle applicazioni pratiche, per esempio in meccanica statistica, si usano a volte metodi approssimati in cui in prima approssimazione si trascura l'hamiltoniana d'interazione o solo una sua parte (es. quella a tre corpi) e se ne tiene conto solo nelle approssimazioni successive.

Esercizi 9.1.

1. Si consideri un punto materiale P di massa m vincolato alla superficie liscia data dal toro $\mathbb{T}^2$ in quiete in un riferimento inerziale $\mathscr{I}$. Fissato un sistema di coordinate cartesiane ortonormali solidali con $\mathscr{I}$, di assi $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ ed origine O, il toro $\mathbb{T}^2$ ha equazione $P(\theta,\phi) = O + (R + r\cos\theta)\cos\phi\,\mathbf{e}_x + (R + r\cos\theta)\sin\phi\,\mathbf{e}_y + r\sin\phi\,\mathbf{e}_z$, dove $R > r > 0$ sono costanti. Oltre alla reazione vincolare $\boldsymbol{\phi}$, su P agisce la forza di una molla ideale (senza massa e con lunghezza nulla a riposo) di costante elastica $k > 0$ e con un estremo attaccato in O ed infine la forza viscosa $-\gamma\mathbf{v}_P$ dove $\mathbf{v}_P$ è la velocità del punto P nel riferimento $\mathscr{I}$ e $\gamma \geq 0$ è una costante. Si risolvano i seguenti quesiti usando le stesse coordinate $\phi \in (-\pi,\pi)$, $\theta \in (-\pi,\pi)$ per descrivere il moto del punto P.

(a) Scrivere le equazioni di Eulero-Lagrange del punto P.

(b) Dimostrare che se $\gamma = 0$, allora su ogni moto del sistema si conserva la componente z del momento angolare rispetto al polo O in $\mathscr{I}$ e l'energia meccanica totale in $\mathscr{I}$ e scrivere la forma esplicita di tali integrali primi in funzione delle coordinate libere usate.

(c) Per $\gamma = 0$ si passi in formulazione di Hamilton: scrivere l'espressione esplicita dell'hamiltoniana in coordinate di Hamilton e scrivere esplicitamente le equazioni di Hamilton.

(d) Nel caso $\gamma > 0$, usando le equazioni di Eulero-Lagrange, determinare la soluzione del problema del moto con condizioni iniziali $\theta(0) = 0$, $\dot\theta(0) = 0$, $\phi(0) = 0$, $\dot\phi(0) = v > 0$.

2. Si consideri un sistema di coordinate cartesiane ortonormali solidali con il riferimento non inerziale $\mathscr{I}$, di origine O e assi $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. Una particella P di massa m è vincolata a scorrere sulla curva liscia di equazione $z = \cosh x$. Una seconda particella Q di massa m è libera di scorrere su un'asta rigida ideale liscia che esce da $O + x\,\mathbf{e}_x$ (il punto proiezione di P sull'asse x) e rimane nel piano perpendicolare a $\mathbf{e}_x$ ed è libera di ruotare attorno a $O + x\,\mathbf{e}_x$. Le due particelle sono connesse l'un l'altra da una molla ideale di costante elastica k e sono sottoposte alla forza di gravità con accelerazione costante $-g\,\mathbf{e}_z$. Il riferimento $\mathscr{I}$ ruota attorno all'asse z rispetto ad un riferimento inerziale $\hat{\mathscr{I}}$, dotato di assi solidali $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$, $\hat{\mathbf{e}}_z \equiv \mathbf{e}_z$ uscenti dall'origine $\hat{O} \equiv O$, con $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = \Omega\,\mathbf{e}_z$ e $\Omega > 0$ costante. Si risolvano i seguenti questiti usando come coordinate libere del sistema le tre coordinate x, θ, r individuate come segue. La coordinata x è la prima componente in $\mathscr{I}$ del vettore posizione $P - O$, l'angolo θ è quello che l'asta $Q - P$ forma con l'asse y (trasportato parallelamente in P) ed è orientato positivamente rispetto a $\mathbf{e}_x$ ed infine r è la distanza tra Q e $O + x\,\mathbf{e}_x$.

(a) Scrivere le equazioni di Eulero-Lagrange del sistema.

(b) Dimostrare che si conserva l'hamiltoniana associata alla lagrangiana nel riferimento $\mathscr{L}|_{\hat{\mathscr{I}}}$ con coordinate libere suddette e provare che tale hamiltoniana, sui moti del sistema, coincide con l'energia meccanica totale nel riferimento non inerziale $\mathscr{I}$.

(c) Passare in formulazione di Hamilton: scrivere l'espressione esplicita dei momenti coniugati in funzione delle coordinate lagrangiane e la trasformazione inversa, scrivere l'hamiltoniana precedentemente trovata in funzione delle variabili di Hamilton e scrivere le equazioni di Hamilton.

9.2 Sistemi hamiltoniani su $\mathbb{R} \times \mathbb{R}^{2n}$.

In questa sezione ci occuperemo di alcune proprietà elementari ma importanti dei sistemi hamiltoniani su $\mathbb{R} \times \mathbb{R}^{2n}$, in particolare vedremo una prima versione del *teorema di Liouville*.

9.2.1 Sistemi hamiltoniani su $\mathbb{R} \times \mathbb{R}^{2n}$.

Consideriamo un sistema fisico $\mathfrak{S}$ che ammette descrizione Hamiltoniana nello spazio delle fasi $F(\mathbb{V}^{n+1})$. Se fissiamo Se fissiamo un sistema di coordinate naturali sullo spazio delle fasi: $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$, possiamo identificare lo spazio delle fasi, almeno per quanto concerne la regione aperta coperta dalle coordinate, con un aperto dello spazio $\mathbb{R} \times \mathbb{R}^{2n}$, dove il primo fattore è l'asse del tempo. Ci metteremo ora nella situazione semplificata in cui le coordinate $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ varino su *tutto* $\mathbb{R} \times \Omega$ dove Ω è un aperto di $\mathbb{R}^{2n}$ che si identifica con ogni fibra $\mathbb{F}_t$ e pertanto possiamo pensare $F(\mathbb{V}^{n+1})$ come dato dal prodotto cartesiano $\mathbb{R} \times \Omega$. Nella situazione detta le equazione di Hamilton prendono la seguente forma:

$$\frac{d\mathbf{x}}{dt} = S\nabla\mathscr{H}(t, \mathbf{x}) \,. \tag{9.23}$$

dove: $\mathbf{x}$ è il vettore colonna $(q^1, \ldots, q^n, p_1, \ldots, p_n)^t$ è un generico punto in $\mathbb{R}^n$, $\mathscr{H} : \mathbb{R} \times \mathbb{R}^{2n} \to \mathbb{R}$ è la funzione hamiltoniana del sistema $\mathfrak{S}$ e assumertemo che $\nabla\mathscr{H}$ sia $C^1(\mathbb{R} \times \Omega)$ per assicurare (con abbondanza di ipotesi) la validità dei teroremi di esistenza ed unicità delle soluzioni[1], la matrice S, detta **matrice simplettica**, è una matrice reale $2n \times 2n$ con struttura

$$S := \left[\begin{array}{c|c} 0 & I \\ \hline -I & 0 \end{array}\right] \,, \tag{9.24}$$

dove la matrice I è la matrice identità $n \times n$. Infine ∇f è il vettore colonna

$$\nabla f := \left(\frac{\partial f}{\partial q^1}, \ldots \frac{\partial f}{\partial q^n}, \frac{\partial f}{\partial p_1}, \ldots \frac{\partial f}{\partial p_n}\right)^t \,.$$

La forma (9.23) delle equazioni di Hamilton implica importanti conseguenze sulle proprietà fisico-matematiche delle soluzioni. In particolare i fondamentali *teorema di Liouville* e *teorema del ritorno di Poincaré*.
Una questione importante è allora quella di determinare criteri matematici per stabilire se un fissato sistema di equazioni differenziali del prim'ordine su qualche aperto $\mathbb{R} \times \Omega$, $\Omega \subset \mathbb{R}^{2n}$:

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t, \mathbf{x}) \,,$$

si possa ricondurre, almeno localmente, ad un sistema hamiltoniano della forma (9.23), definendo una opportuna funzione hamiltoniana.
Per dare una risposta a tale domanda dobbiamo introdurre alcuni rudimenti del *formalismo simplettico*.

[1] Come sappiamo dal capitolo 3, sarebbe sufficiente richiedere che $\nabla\mathscr{H} \in C^0(\mathbb{R} \times \Omega)$ e che valga la condizione di Lipschitz locale nella variabile $\mathbf{x}$ per ogni $t \in \mathbb{R}$.

9.2.2 Il gruppo simplettico ed i sistemi hamiltoniani.

Definizione 9.2. (**Forma simplettica e spazio simplettico**.) Se V è uno spazio vettroriale sul campo $\mathbb{R}$, un'applicazione $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ che associa alla coppia di vettori (u, v) il numero reale $\langle u, v\rangle$ è detta **forma simplettica** su V se gode delle seguenti proprietà :
(a) antisummetria: $\langle u, v\rangle = -\langle v, u\rangle$ per ogni $u, v \in V$,
(b) bilinearità : $\langle\cdot,\cdot\rangle$ è lineare separatamente in ognuno dei due argomenti,
(c) non degeneratezza: se $\langle u, v\rangle = 0$ per ogni $v \in V$ allora $u = 0$.
La coppia $(V, \langle,\rangle)$ è detta **spazio simplettico**. $\diamondsuit$

La rilevanza della nozione di spazio simplettico in meccanica hamiltoniana è dovuta al fatto che la matrice S è naturalmente associata ad una forma simplettica. L'applicazione

$$\langle\cdot,\cdot\rangle_S : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$$

definita da

$$\mathbb{R}^{2n} \to \mathbb{R} \ni (u, v) \mapsto u^t S v \tag{9.25}$$

definisce infatti una forma simplettica su $\mathbb{R}^2$. Infatti la bilinearità è ovvia, l'antisimmetria segue dal fatto che $S^t = -S$ come si verifica immediatamente dalla definizione di S, per cui:

$$\langle u, v\rangle_S = u^t S v = (u^t S v)^t = v^t S^t u = -v^t S u = -\langle v, u\rangle_S \, .$$

La non degeneratezza si prova come segue notando che, dalla definizione di S: $SS = -I$. Se $u^t S v = 0$ per ogni $v \in \mathbb{R}^{2n}$, possiamo scegliere $v = Su$ e pertanto deve valere $u^t S^t S u = 0$ che significa $-u^t S S u = 0$ cioé $u^t u = 0$ e quindi $u = 0$.
Si osservi per inciso che $SS = -I$ implica anche che $\det S = \pm 1$. Infatti: $(\det S)^2 = (\det S^2) = (-1)^{2n} = 1$ da cui, infine $\det S = \pm 1$.
In realtà si può provare che *ogni* forma simplettica su V spazio vettoriale reale di dimensione pari si può ridurre alla forma (9.25), scegliendo un sistema di coordinate cartesiane opportune. Vale infatti il seguente teorema dovuto a Darboux la cui prova è lasciata per sercizio.

Teorema 9.2. (**Primo teorema di Darboux**.) *Se $(V, \langle\cdot,\cdot\rangle)$ è uno spazio simplettico con $dim V = 2n$, $n = 1, 2, \ldots$ fissato, esiste una base $\{\mathbf{s}_i\}_{i=1,\ldots,2n}$ di V in cui la forma simplettica è rappresentata dalla matrice S in (9.24). In altre parole, se u e v sono rispettivamente i vettori colonna delle componenti dei vettori $\mathbf{u}$ e $\mathbf{v}$ in V rispetto alla base $\{\mathbf{s}_i\}_{i=1,\ldots,2n}$:*

$$\langle \mathbf{u}, \mathbf{v}\rangle = u^t S v \, . \tag{9.26}$$

$\diamondsuit$

Osservazioni 9.4.
(1) La base $\{s_i\}_{i=1,\ldots,2n}$ non è l'unica base in cui la forma simplettica si riduce alla (9.25). Esiste una classe di basi, dette **canoniche** rispetto alla forma simplettica, in cui vale la (9.26).

Supponiamo che lo stesso vettore $\mathbf{u} \in V$ sia rappresentato dal vettore colonna $u' \in \mathbb{R}^{2n}$ rispetto alla base $\{\mathbf{s}'_i\}_{i=1,\ldots,2n}$ e dal vettore colonna u rispetto alla base canonica $\{\mathbf{s}_i\}_{i=1,\ldots,2n}$. Valga $u = Au'$ dove la matrice $A \in GL(2n, \mathbb{R})$ (il gruppo delle matrici reali invertibili $2n \times 2n$) è la matrice di passaggio da una base all'altra. Abbiamo allora che:

$$\langle \mathbf{u}, \mathbf{v} \rangle = u^t S v = (Au')^t A v' = u'^t (A^t S A) v' \ .$$

da cui concludiamo che: *la base* $\{\mathbf{s}'_i\}_{i=1,\ldots,2n} \subset V$ *è canonica se e solo se* $A^t S A = S$.
(2) La precedente osservazione ci ha portato a considerare le matrici $A \in M(2n, \mathbb{R})$ (l'algebra delle matrici reali $2n \times 2n$) che soddisfano il vincolo $A^t S A = S$. Mostriamo che queste matrici formano un gruppo sottogruppo di $GL(2n, \mathbb{R})$. Per fare ciò è sufficiente mostrare che le matrici dette (i) formano un sottoinsieme di $GL(2n, \mathbb{R})$, (ii) tale sottoinsieme contiene l'identità , (iii) tale sottoinsieme è chiuso rispetto al prodotto matriciale, (iv) tale sottoinsieme è chiuso rispetto al calcolo dell'inversa[2].
(i) Se $A^t S A = S$, calcolando il determinante e facendo uso della regola di Binet, troviamo che $(\det A^t) \det S \det A = \det S$, da cui, tenendo conto che $\det S \neq 0$ come verificato precedentemente e che $\det A^t = \det A$, si ha $(\det A)^2 = 1$ e pertanto $\det A \neq 0$. Quindi effettivamente le matrici che soddisfano il vincolo $A^t S A = S$ sono un sottoinsieme di $GL(2n, \mathbb{R})$.
(ii) I soddisfa il vincolo: $I^t S I = S$.
(iii) Se $A^t S A = S$ e $B^t S B = S$ allora $(AB)^t S (AB) = B^t (A^t S A) B = B^t S B = S$, per cui $(AB)^t S (AB) = S$.
(iv) Cominciamo a provare che l'insieme delle matrici che soddisfano $A^t S A = S$ è *chiuso sotto l'operazione di trasposizione.* Se $A^t S A = S$ allora $S A^t S A = -I$, dove si è usato il fatto che $SS = -I$ (e quindi $S^{-1} = -S$). Moltiplicando per AS a sinistra e $A^{-1}S$ a destra, troviamo $A S A^t = S$. Facendo uso di ciò , concludiamo che se $A^t S A = S$ allora $A S A^t = S$ e quindi, calcolando l'inverso di entrambi i membri: $A^{t-1}(-S)A^{-1} = -S$, da cui: vale $A^{t-1} S A^{-1} = S$ se $A^t S A = S$

Abbiamo provato la seguente proposizione.

Proposizione 9.2. *Per ogni fissato* $n = 1, 2, \ldots,$ *l'insieme di matrici*

$$Sp(n, \mathbb{R}) := \left\{ A \in M(2n, \mathbb{R}) \ \middle| A^t S A = S \right\} \ ,$$

è un gruppo che è sottogruppo di $GL(n, \mathbb{R})$. *Inoltre vale* $A \in Sp(n, \mathbb{R})$ *se e solo se* $A^t \in Sp(n, \mathbb{R})$.
$\diamondsuit$

Definizione 9.3. (**Gruppo simplettico**) Il **gruppo simplettico reale di ordine** n è definito dal sottogruppo di $GL(2n, \mathbb{R})$

$$Sp(n, \mathbb{R}) := \left\{ A \in M(2n, \mathbb{R}) \ \middle| A^t S A = S \right\} \ .$$

[2] il fatto che l'insieme contenga l'indentità segue in realtà da (iii) e (iv).

$\diamondsuit$

Se $E \in M(m, \mathbb{R})$ e $s \in \mathbb{R}$, definiamo (dove $E^0 := I$):

$$e^{sE} := \sum_{k=0}^{+\infty} \frac{s^k}{k!} E^k \ .$$

La serie converge uniformente nella topologia metrica di $M(m, \mathbb{R}) \equiv \mathbb{R}^{m^2}$ in ogni intervallo finito per ogni scelta delle matrice E.
È un fatto noto dalla teoria del gruppo $GL(m, \mathbb{R})$ (più in generale della teoria dei *gruppi di Lie*) che

$$e^{sE} \in GL(m, \mathbb{R}) \quad \text{per ogni } s \in \mathbb{R} \text{ e ogni } E \in M(m, \mathbb{R}) \ . \tag{9.27}$$

La dimostrazione segue immediatamente dalla nota identità , che prova che il determinante della matrice a primo membro dell'identità di sopra non può mai annullarsi,

$$\det e^F = e^{trF} \quad \text{per ogni } F \in M(m, \mathbb{R}).$$

Il risultato in (9.27) può essere parzialmente invertito nel senso che segue. Pensando $M(m, \mathbb{R})$ come $\mathbb{R}^{m^2}$, $GL(m, \mathbb{R})$ altro non è che la controimmagine di $\mathbb{R} \setminus \{0\}$ della funzione determinante. Dato che questa è continua, $GL(m, \mathbb{R})$ è un sottoinsieme aperto di $\mathbb{R}^{n^2}$ che contiene la matrice indentica I. Calcolando il determinante jacobiano per $F = 0$ dalla funzione $\mathbb{R}^{m^2} \ni F \mapsto e^F \in \mathbb{R}^{m^2}$ si scopre che esso è non nullo. In tal modo questa applicazione definisce un diffeomorfismo locale di $\mathbb{R}^{n^2}$ che manda un intorno aperti della matrice 0 nel dominio in un intorno aperto della matrice I nell'mmagine. Questo porta a concludere che, in un intorno aperto $\mathcal{O}_I$ sufficientemente piccolo dell'identità di $GL(m, \mathbb{R})$, ogni elemento $A \in GL(m, \mathbb{R})$ si può scrivere come e^F ed F è univocamente determinata da A se si lavora in un intorno dufficientemente piccolo di 0.
Dato che $Sp(n, \mathbb{R})$ è un sottogruppo di $GL(2n, \mathbb{R})$, ci si può chiedere se ci sia qualche sottoinsieme dell'intorno $\mathcal{O}_I$ la cui esponenziazione produce elementi di $Sp(n, \mathbb{R})$. La questione si inquadra più in generale nella teoria delle algebre di Lie e nella ben nota corrispondenza locale tra gruppi di Lie ed algebre di Lie associate, ma non ci addentreremo in essa rimamendo su un piano molto elementare. Vale la seguente proposizione.

Proposizione 9.3. *Per ogni fissato $n = 1, 2, \ldots,$ vale:*

$$e^{sE} \in Sp(n, \mathbb{R}) \quad \forall s \in \mathbb{R}$$

se e solo se $E \in sp(n, \mathbb{R})$ dove

$$sp(n, \mathbb{R}) := \left\{ E \in M(2n, \mathbb{R}) \ \middle| E^t S + SE = 0 \right\} \ ,$$

è detto insieme delle **matrici hamiltoniane di ordine** n. $\diamondsuit$

Dimostrazione. Dallo sviluppo di e^{sE} segue facilmente che:

$$(e^{sE})^t = e^{sE^t}$$

ed ancora, derivando sotto il segno di serie, data la uniforme convergenza della serie delle derivate,

$$\frac{de^{sE}}{ds} = Ee^{sE} = e^{sE}E\,.$$

Se $(e^{sE})^t S e^{sE} = S$, derivando per $s = 0$ troviamo subito che $E^t S + SE = 0$. Questo prova che $e^{sE} \in Sp(n, \mathbb{R})$ per ogni $t \in \mathbb{R}$, implica $E \in sp(n, \mathbb{R})$. Sia ora $E \in sp(n, \mathbb{R})$, mostriamo che $e^{sE} \in Sp(n, \mathbb{R})$ per ogni $s \in \mathbb{R}$. Consideriamo la funzione, per $t \in \mathbb{R}$: $f(s) := (e^{sE})^t S e^{sE}$. Calcolando la derivata in s troviamo immediatamente che f soddisfa il problema di Cauchy autonomo e lineare su $\mathbb{R}^{n^2}$:

$$\frac{df}{ds} = E^t f(s) + f(s)E \quad \text{con } f(0) = S\,.$$

La funzione $\mathbb{R} \ni s \mapsto f(s) := S$ *costantemente* è soluzione di tale problema dato che $E^t S + SE = 0$ per ipotesi. Pertanto questa è l'unica soluzione (massimale e completa) del problema di Cauchy. In altre parole:

$$f(s) = (e^{sE})^t S e^{sE} = S \quad \forall s \in \mathbb{R}\,,$$

cioé $e^{sE} \in Sp(n, \mathbb{R})$ per ogni $s \in \mathbb{R}$ se $E \in sp(n, \mathbb{R})$. $\square$

Osservazioni 9.5. $GL(m, \mathbb{R})$ e $Sp(n, \mathbb{R})$ sono gruppi di Lie, cioé hanno una struttura di varietà differenziabile (in questo caso indotta da $\mathbb{R}^{m^2}$ e $\mathbb{R}^{(2n)^2}$ rispettivamente) compatibile con le operazioni di prodotto e calcolo dell'inversa (che risultano pertanto essere differenziabili). $M(m, \mathbb{R})$ è l'*algebra di Lie* di $GL(m, \mathbb{R})$, cioé lo spazio tangente al punto identità $I \in GL(m, \mathbb{R})$, dotato della struttura di parentesi di Lie individuata (come al solito nel caso di gruppi matriciali) dal **commutatore matriciale**: $[A, B] := AB - BA$ con A e B matrici di $M(m, \mathbb{R})$. Per definizione di algebra di Lie il commutatore di due elementi dell'algebra deve rimanere nell'algebra.
Similmente $sp(n, \mathbb{R})$ è l'*algebra di Lie di* $Sp(n, \mathbb{R})$. In particolare, per verifica diretta si prova facilmente che $sp(n, \mathbb{R})$ è un sottospazio vettoriale di $M(2n, \mathbb{R})$ che soddisfa la relazione di chiusura:

$$[E, E'] \in sp(n, \mathbb{R}) \quad \text{se } E, E' \in sp(n, \mathbb{R}).$$

Siamo ora in grado di enunciare e provare il teorema che connette le matrici hamiltoniane ai sistemi di equazioni differenziali hamiltoniani.

Teorema 9.3. *Per $n = 1, 2, \ldots$ fissato, si consideri il sistema di mequazioni differenziali su $\mathbb{R} \times \Omega$, dove Ω è un aperto (non vuoto) di $\mathbb{R}^{2n}$:*

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t, \mathbf{x})\,, \tag{9.28}$$

dove $\mathbf{F} \in C^1(\mathbb{R} \times \Omega)$. *Vale quanto segue.*
(a) *Se tale sistema si può scrivere in forma di Hamilton (9.23) per qualche funzione di hamilton, allora le matrici* $D_{(t,\mathbf{x})}$ *i cui coefficienti sono le derivate:*

$$\left.\frac{\partial F^i}{\partial x^j}\right|_{(t,\mathbf{x})}$$

sono in $sp(n, \mathbb{R})$ *per ogni* $(t, \mathbf{x}) \in \mathbb{R} \times \Omega$.
(b) *Viceversa se* $D_{(t,\mathbf{x})} \in sp(n, \mathbb{R})$ *per ogni* $(t, \mathbf{x}) \in \mathbb{R} \times \Omega$, *allora, per ogni sottoinsieme aperto semplicemente connesso* $B \subset \Omega$, *il sistema (9.28) si può scrivere in forma hamiltoniana su* $\mathbb{R} \times B$. $\diamondsuit$

Dimostrazione. (a) Il sistema è hamiltoniano se e solo se possiamo definire una funzione (il cui gradiente sarà di classe C^1 visto che $\mathbf{F}$ è tale per ipotesi): $(t, \mathbf{x}) \mapsto \mathscr{H}(t, \mathbf{x})$ tale che $\mathbf{F}(t, \mathbf{x}) = S\nabla\mathscr{H}(t, \mathbf{x})$. Questo è equivalente a dire, moltiplicando ambo membri per $-S$, che $\mathbf{G} := -S\mathbf{F}$ è il gradiente di una funzione per ogni fissato istante t. Dato che $\mathbf{G}$ è di classe C^1, il fatto di essere gradiente implica la condizione di irrotazionalità

$$\frac{\partial G^k}{\partial x^h} = \frac{\partial G^h}{\partial x^k} \, .$$

Ossia, passando in componenti:

$$-\frac{\partial \sum_r S_{kr} F^r}{\partial x^h} = -\frac{\partial \sum_r S_{hr} F^r}{\partial x^k} \, ,$$

che può riscriversi:

$$-\sum_r S_{kr} \frac{\partial F^r}{\partial x^h} + \sum_r S_{hr} \frac{\partial F^r}{\partial x^k} = 0 \, ,$$

e ancora, ricordando che $S_{kr} = -S_{rk}$:

$$\sum_r \frac{\partial F^r}{\partial x^h} S_{rk} + \sum_r S_{hr} \frac{\partial F^r}{\partial x^k} = 0 \, .$$

In forma compatta questa identità dice che:

$$D^t_{(t,\mathbf{x})} S + S D_{(t,\mathbf{x})} = 0 \, ,$$

cioé $D_{(t,\mathbf{x})} \in sp(n, \mathbb{R})$.
(b) Per dimostrare il viceversa, fissiamo l'aperto $B \subset \Omega$ semplicemente connesso e $t \in \mathbb{R}$. La condizione sopra scritta è equivalente alla condizione di irrotazionalità per il campo di classe C^1: $B \ni \mathbf{x} \mapsto -S\mathbf{F}(t, \mathbf{x})$. Come ben noto dai corsi di analisi su B è allora definito un potenziale

$$\mathscr{H}(t, \mathbf{x}) := \int_{\mathbf{x}_0}^{\mathbf{x}} (-S\mathbf{F}(t, \mathbf{x})) \cdot d\mathbf{x} \, , \tag{9.29}$$

il cui gradiente è $-S\mathbf{F}(t,\mathbf{x})$. L'integrale è calcolato su un qualsiasi cammino (indipendente da t) di classe C^1, che congiunge $\mathbf{x} \in B$ al punto di riferimento fissato arbitrariamente $\mathbf{x}_0 \in B$, ed è tutto contenuto in B. Per costruzione, esattamente al tempo t su B:

$$\mathbf{F}(t,\mathbf{x}) = S\nabla\mathscr{H}(t,\mathbf{x})\,.$$

La funzione $\mathscr{H}$ dipende parametricamente dal tempo dato che il secondo membro di (9.29) dipende dal tempo. Per costruzione l'espressione a secondo membro di (9.29) determina, ad ogni tempo $t \in \mathbb{R}$, una funzione il cui gradiente è $-S\mathbf{F}$ su B ed è pertanto la funzione di Hamilton cercata su $\mathbb{R} \times B$. $\square$.

Osservazioni 9.6. Il toerema prova più generalmente che una funzione a valori vettoriali $\mathbf{F} := \mathbf{F}(t,\mathbf{x})$ di classe $C^1((a,b)\times\Omega)$ con Ω semplicemente connesso è scrivibile come $\mathbf{F}(t,\mathbf{x}) = S\nabla\mathscr{K}(t,\mathbf{x})$ per qualche funzione scalare $\mathscr{K}$ definita sullo stesso dominio, se e solo se le matrici i cui coefficienti sono

$$\left.\frac{\partial F^i}{\partial x^j}\right|_{(t,\mathbf{x})}$$

appartengono a $sp(n,\mathbb{R})$ per ogni $(t,\mathbf{x}) \in \mathbb{R}\times\Omega$.

9.2.3 Il teorema di Liouville su $\mathbb{R}\times\mathbb{R}^{2n}$.

Il teorema di Lioville è sicuramente uno dei più importanti teoremi di tutta la meccanica Hamiltoniana, della meccanica statistica e di tutta la fisica matematica in generale. In sostanza tale teorema afferma che l'evoluzione di un gruppo di punti rappresentativi del sistema nello *spaziotempo* delle fasi è tale da conservare nel tempo il volume individuato dai punti in ogni fibra a tempo fissato (spazio delle fasi al tempo t, $\mathbb{F}_t$). La misura da usare per calcolare il volume in questione è quello individuato dai sistemi di coordinate naturali. Per il momento non ci occuperemo di tale nozione di volume e lavoreremo ancora con sistemi Hamiltoniani su $\mathbb{R}\times\Omega$ in un fissato sistema di coordinate naturali $(t,\mathbf{x})$, pensando la misura da usare sulle fibre $\mathbb{F}_t \equiv \Omega$ a tempo costante come quella standard in $\mathbb{R}^{2n}$.

Consideriamo il sistema hamiltoniano su $\mathbb{R}\times\Omega$ descritto da (9.23) con $\nabla\mathscr{H} \in C^1(\mathbb{R}\times\Omega)$. Estendiamo il sistema a $\mathbb{R}\times\mathbb{R}\times\Omega$ facendolo diventare un sistema autonomo:

$$\begin{cases} \dfrac{d\mathbf{x}}{ds} &= S\nabla\mathscr{H}\left(t(s),\mathbf{x}(s)\right), \\[2ex] \dfrac{dt}{ds} &= 1\,. \end{cases} \tag{9.30}$$

Indichiamo con $I_{(t_0,\mathbf{x}_0)} \ni t \mapsto (t_0+s,\mathbf{x}(s|t_0,\mathbf{x}_0))$ la soluzione massimale, con dominio massimale dato dall'intervallo aperto $I_{(t_0,\mathbf{x}_0)} \subset \mathbb{R}$, del sistema (9.30) e con condizioni iniziali $\mathbf{x}(0) = \mathbf{x}_0$ e $t(0) = t_0$. Come sappiamo viene a definirsi un gruppo locale ad un parametro di diffeomorfismi $\Phi^{(\mathscr{H})}$ di classe C^1 congiuntamente nelle sue variabili con la sua derivata prima in s, definito dal fatto che

$$\Phi^{(\mathscr{H})}_s(t_0,\mathbf{x}_0) := (t_0+s,\mathbf{x}(s|t_0,\mathbf{x}_0)) \in \mathbb{R}\times\Omega\,.$$

Osservazioni 9.7.
(1) Si deve ricordare che il range di variazione di s dipende dall'argomento $(t_0, \mathbf{x}_0)$. Tuttavia, come abbiamo provato nella sezione 3.5.3 l'applicazione $(s, t_0, \mathbf{x}_0) \mapsto \Phi_s^{(\mathscr{H})}(t_0, \mathbf{x}_0)$ è ben definita su un sottoinsieme aperto non vuoto di $\mathcal{D} \subset \mathbb{R} \times \mathbb{R} \times \Omega$.
(2) Ricordiamo che valgono le relazioni:

$$\Phi_0^{(\mathscr{H})}(t, \mathbf{x}) = (t, \mathbf{x}) \,, \quad \Phi_s^{(\mathscr{H})} \circ \Phi_u^{(\mathscr{H})}(t, \mathbf{x}) = \Phi_{s+u}^{(\mathscr{H})}(t, \mathbf{x}) \,,$$

la seconda ha senso dove entrambi i membri sono ben definiti. Usando il fatto che $\mathcal{D}$ è aperto si ha che (vedi la Sezione 3.5.3), in un intorno sufficientemente piccolo di ogni $(t_0, \mathbf{x}_0)$ che contiene l'argomento $(t, \mathbf{x})$, e se s e u cadono in un intorno di 0 sufficientemente piccolo, la seconda identità di sopra è sicuramente valida. In base ad essa, se l'insieme aperto $A \subset \mathbb{R} \times \Omega$ è sufficientemente piccolo e s è abbastanza vicino a 0, $\Phi_s^{(\mathscr{H})}\restriction_A$ è un diffeomorfismo da A all'aperto $\Phi_s^{(\mathscr{H})}(A)$. Il diffeomorfismo inverso s ottiene restringendo $\Phi_{-s}^{(\mathscr{H})}$ a $\Phi_s^{(\mathscr{H})}(A)$.
(3) Se tutte le soluzioni del sistema (9.30) sono complete, cioé definite da $-\infty$ a $+\infty$, allora $\Phi_s^{(\mathscr{H})}$ è definito globalmente su $\mathbb{R} \times \Omega$ per ogni $s \in \mathbb{R}$, e $\{\Phi_s^{(\mathscr{H})}\}_{s \in \mathbb{R}}$ è propriamente un gruppo ad un parametro (s) di diffeomorfismi definiti su tutto $\mathbb{R} \times \Omega$ e surgettivi su tale insieme (vedi la Sezione 3.5.3). In tal caso vale:

$$\Phi_0^{(\mathscr{H})} = id \,, \quad \Phi_s^{(\mathscr{H})} \circ \Phi_u^{(\mathscr{H})} = \Phi_{s+u}^{(\mathscr{H})} \quad \forall s, t \in \mathbb{R} \,.$$

Questa situazione si ha, in particolare, quando $\mathscr{H}$ è solo funzione di $\mathbf{x}$, e dunque il sistema Hamiltoniano è detto *autonomo*, ed il supporto di $S\nabla\mathscr{H}$ in Ω è un compatto.

$\Phi^{(\mathscr{H})}$ è detto **flusso di fase** associato al sistema hamiltoniano (9.23).
Fissiamo un istante t_0 ed un insieme aperto a chiusura compatta $D_{t_0} \subset \mathbb{F}_{t_0}$. Possiamo fare evolvere D_{t_0} attraverso il gruppo locale $\Phi^{(\mathscr{H})}$ fino al tempo $t = t_0 + s$. Se D_{t_0} è sufficientemente piccolo $\{t\} \times D_t := \Phi_s^{(\mathscr{H})}(\{t_0\} \times D_{t_0})$, con $D_t \subset \mathbb{F}_t$, è ben definito. Per costruzione D_t contiene il punto rappresentativo al tempo t di ogni soluzione delle equazioni di Hamilton che al tempo t_0 passano per D_{t_0}.

Teorema 9.4. (**Teorema di Liouville in** $\mathbb{R} \times \mathbb{R}^{2n}$.) *In riferimento al sistema Hamiltoniano (9.23) su $\mathbb{R} \times \Omega$ con $\Omega \subset \mathbb{R}^{2n}$ aperto non vuoto e $\nabla\mathscr{H} \in C^1(\mathbb{R} \times \Omega)$, il suo flusso di fase $\Phi^{(\mathscr{H})}$ preserva il volume delle fibre nello spaziotempo delle fasi. In altre parole, se $D_{t_0} \subset \mathbb{F}_{t_0}$ è una aperto sufficientemente piccolo e $t > t_0$ è sufficientemente vicino a t_0 in modo tale che $\{t\} \times D_t := \Phi_s^{(\mathscr{H})}(\{t_0\} \times D_{t_0})$ sia ben definito, allora:*

$$\int_{D_{t_0}} d^{2n}\mathbf{x} = \int_{D_t} d^{2n}\mathbf{x} \,. \tag{9.31}$$

dove $d^{2n}\mathbf{x}$ denota la misura di Lebesgue standard sulle fibre $\mathbb{F}_t$, in coordinate $\mathbf{x} = (q^1, \ldots, q^n, p_1, \ldots, p_n)$ data da $d^{2n}\mathbf{x} = dq^1 \cdots dq^n dp_1 \cdots dp_n$.

Se tutte le soluzioni del sistema (9.23) sono complete, l'identità di sopra vale comunque si scelga $D_{t_0} \subset \mathbb{F}_{t_0}$ aperto a chiusura compatta e comunque si prenda $t \in \mathbb{R}$. $\diamondsuit$

Dimostrazione. Prima di tutto notiamo che D_t è aperto: per $\delta > 0$ sufficientemente piccolo, l'aperto $(t_0 - \delta, t_0 + \delta) \times D_{t_0}$ viene trasformato nell'insieme $(t - \delta, t + \delta) \times D_t$ dal diffeomorfismo $\Phi_t^{(\mathscr{H})}$ e pertanto è anch'esso aperto ed in particolare D_t deve essere aperto. Di conseguenza D_t è misurabile. Possiamo scrivere, cambiando variabile nel secondo integrale:

$$\int_{D_t} d^{2n}\mathbf{x} = \int_{D_{t_0}} |\det J_s| d^{2n}\mathbf{x}$$

dove $\det J_t$ è il determinante della matrice jacobiana della trasformazione $\mathbf{x}_0 \mapsto \mathbf{x}(s|t_0, \mathbf{x}_0)$. Possiamo derivare in s sotto il segno di integrale in quanto $(s, \mathbf{x}) \mapsto J_s$ è continua e $\overline{D_{t_0}}$ è compatto. Per provare la tesi è sufficiente mostrare che:

$$\frac{\partial |\det J_s|}{\partial s} = 0 \,, \quad \forall s \in I_{(t_0, \mathbf{x}_0)} \,. \tag{9.32}$$

Notiamo che per $s = 0$, $\det J_s = 1 > 0$ e pertanto il segno di $\det J_s$ deve rimanere positivo per ragioni di continuità : non può annullarsi dato che $(t_0, \mathbf{x}_0) \mapsto (t_0 + s, \mathbf{x}(s|t_0, \mathbf{x}_0))$ è un diffeomorfismo (quindi ha determinante jacobiano non nullo) e il determinante della matrice jacobiana di tale trasformazione coincide con quello della trasformazione $\mathbf{x}_0 \mapsto \mathbf{x}(s|t_0, \mathbf{x}_0)$. Dobbiamo quindi dimostrare che:

$$\frac{\partial \det J_s}{\partial s} = 0 \,, \quad \forall s \in I_{(t_0, \mathbf{x}_0)} \,.$$

Per eseguire il calcolo introduciamo la notazione semplificata, se $s > s'$, ed omettendo di scrivere la dipendenza da t banale:

$$\mathbf{x}_{s'}(\mathbf{x}_s) := \mathbf{x}(s'|t(s'), \mathbf{x}(s)) \,.$$

Abbiamo allora che:

$$\frac{\partial \det J_s}{\partial s} = \lim_{h \to 0} \frac{1}{h} \left(\det \left[\frac{\partial \mathbf{x}_{s+h}(x_0)}{\partial \mathbf{x}_0} \right] - \det \left[\frac{\partial \mathbf{x}_s(x_0)}{\partial \mathbf{x}_0} \right] \right)$$

$$= \lim_{h \to 0} \frac{1}{h} \left(\det \left[\frac{\partial \mathbf{x}_{s+h}(\mathbf{x}_s)}{\partial \mathbf{x}_s} \right] \det \left[\frac{\partial \mathbf{x}_s(\mathbf{x}_0)}{\partial \mathbf{x}_0} \right] - \det \left[\frac{\partial \mathbf{x}_s(\mathbf{x}_0)}{\partial \mathbf{x}_0} \right] \right) =$$

$$= \lim_{h \to 0} \frac{1}{h} \left(\det \left[\frac{\partial \mathbf{x}_{s+h}(\mathbf{x}_s)}{\partial \mathbf{x}_s} \right] - 1 \right) \det \left[\frac{\partial \mathbf{x}_s(\mathbf{x}_0)}{\partial \mathbf{x}_0} \right] = (\det J_s) \lim_{h \to 0} \frac{1}{h} \left(\det \left[\frac{\partial \mathbf{x}_{s+h}(\mathbf{x}_s)}{\partial \mathbf{x}_s} \right] - 1 \right) \,.$$

Possiamo calcolare il limite sviluppando la funzione nel determinante con Taylor nella variabile h per $h = 0$. Abbiamo

$$\frac{\partial x^i_{h=0}(\mathbf{x})}{\partial x^j} = \delta^i_j \,,$$

mentre

$$\frac{\partial}{\partial h}|_{h=0} \frac{\partial x^i_h(\mathbf{x})}{\partial x^j} = \frac{\partial}{\partial x^i} \frac{\partial x^j_h}{\partial h}|_{h=0} = \frac{\partial}{\partial x^j} \left(S \nabla \mathscr{H}(t, \mathbf{x}) \right)^i \,,$$

e pertanto, dove $O_{t,\mathbf{x}}(h) \to 0$ per $h \to 0$,

$$\frac{\partial x_h^i(\mathbf{x})}{\partial x^j} = \delta_j^i + h\frac{\partial}{\partial x^j}\left(S\nabla\mathscr{H}(t,\mathbf{x})\right)^i + hO_{t,\mathbf{x}}(h)\ .$$

Il determinante della matrice i cui coefficienti sono il secondo membro dello sviluppo di sopra si ottiene dal seguente lemma tecnico.

Lemma 9.1. *Se $A(h) = I + hB + hO(h)$ dove $A(h), B, O(h) \in M(m,\mathbb{R})$ con $h \in (-\delta,\delta)$ e $O(h) \to 0$ per $h \to 0$, allora*

$$detA(t) = 1 + h\,trB + h\Omega(h)\ , \tag{9.33}$$

dove $\Omega : (-\delta,\delta) \to \mathbb{R}$ soddisfa $\Omega(h) \to 0$ se $h \to 0$. $\diamondsuit$

Traccia di dimostrazione. La dimostrazione si ottiene per via diretta applicando la formula, valida per $C \in M(m,\mathbb{R})$,

$$\det C = \sum_{i_1,i_2,\ldots,i_m=1}^{m} \epsilon_{i_1 i_2 \ldots i_n}\, C_{1i_1}\cdots C_{mi_m}\ ,$$

dove i numeri $\epsilon_{i_1 i_2 \ldots i_m}$ sono completamente individuati dalle richieste: $\epsilon_{123\ldots m-1m} := 1$ e

$$\epsilon_{i_1\ldots i_{h-1}i_h i_{h+1}\ldots i_{k-1}i_k i_{k+1}\ldots i_n} = -\epsilon_{i_1\ldots i_{h-1}i_k i_{h+1}\ldots i_{k-1}i_h i_{k+1}\ldots i_n} \quad \text{per ogni scelta di } h \neq k,$$

cioè si ha antisimmetria per scambio di ogni coppia di indici arbitrariamente scelti. $\square$

Usando il lemma precedente abbiamo imediatamente che:

$$\det\left[\frac{\partial \mathbf{x}_h(\mathbf{x})}{\partial \mathbf{x}}\right] = 1 + h\sum_{i=1}^{2n}\frac{\partial}{\partial x^i}\left(S\nabla\mathscr{H}(t,\mathbf{x})\right)^i + hO_{t,\mathbf{x}}(h)\ .$$

In definitiva:

$$\lim_{h\to 0}\frac{1}{h}\left(\det\left[\frac{\partial \mathbf{x}_h(\mathbf{x})}{\partial \mathbf{x}}\right] - 1\right) = \lim_{h\to 0}\left\{\sum_{i=1}^{2n}\frac{\partial}{\partial x^i}\left(S\nabla\mathscr{H}(t,\mathbf{x})\right)^i + O_{t,\mathbf{x}}(h)\right\} = \sum_{i=1}^{2n}\frac{\partial}{\partial x^i}\left(S\nabla\mathscr{H}(t,\mathbf{x})\right)^i.$$

Abbiamo quindi trovato che:

$$\frac{\partial|\det J_s|}{\partial s} = J_s\sum_{i=1}^{2n}\frac{\partial}{\partial x^i}\left(S\nabla\mathscr{H}(t,\mathbf{x})\right)^i\ .$$

Esplicitamente:

$$\sum_{i=1}^{2n}\frac{\partial}{\partial x^i}\left(S\nabla\mathscr{H}(t,\mathbf{x})\right)^i = \sum_{i=1}^{2n}\frac{\partial}{\partial x^i}\sum_{j=1}^{2n}S_{ij}\frac{\partial}{\partial x^j}\mathscr{H}(t,\mathbf{x}) = \sum_{i,j=1}^{2n}S_{ij}\frac{\partial^2\mathscr{H}(t,\mathbf{x})}{\partial x^i\partial x^j}\ .$$

Ora tenendo conto del teorema di Schwartz:

$$\sum_{i,j=1}^{2n} S_{ij} \frac{\partial^2 \mathscr{H}(t,\mathbf{x})}{\partial x^i \partial x^j} = \sum_{i,j=1}^{2n} S_{ij} \frac{\partial^2 \mathscr{H}(t,\mathbf{x})}{\partial x^j \partial x^i} \ ,$$

cambiando nome agli indici nel secondo membro otteniamo:

$$\sum_{i,j=1}^{2n} S_{ij} \frac{\partial^2 \mathscr{H}(t,\mathbf{x})}{\partial x^i \partial x^j} = \sum_{i,j=1}^{2n} S_{ji} \frac{\partial^2 \mathscr{H}(t,\mathbf{x})}{\partial x^i \partial x^j} = -\sum_{i,j=1}^{2n} S_{ij} \frac{\partial^2 \mathscr{H}(t,\mathbf{x})}{\partial x^i \partial x^j} \ .$$

e pertanto:

$$\sum_{i,j=1}^{2n} S_{ij} \frac{\partial^2 \mathscr{H}(t,\mathbf{x})}{\partial x^i \partial x^j} = 0$$

che implica la validità di (9.32). Questo conclude la dimostrazione della prima affermazione. La seconda es ultima affermazione nella tesi del teorema è immediata conseguenza del fatto che se tutte le soluzioni del sistema (9.23) sono complete lo sono anche quelle di (9.30) e dunque $\{\Phi_s^{(\mathscr{H})}\}_{s\in\mathbb{R}}$ è un gruppo ad un parametro di diffeomorfismi che quindi agisce globalmente su $\mathbb{R}\times\Omega$ e per ogni $s\in\mathbb{R}$. $\square$

9.3 *La struttura di varietà fibrata di $F(\mathbb{V}^{n+1})$ e le equazioni di Hamilton come equazioni globali.

In questa sezione studieremo più a fondo la struttura geometrico differenziale dello spazio delle fasi ed in particolare la natura globale della trasformazione di Legendre e delle equazioni di Hamilton.

9.3.1 Lo spazio fibrato $F(\mathbb{V}^{n+1})$.

Lo spazio $F(\mathbb{V}^{n+1})$ può essere rigorosamente *definito* come varietà fibrata con base $\mathbb{V}^{n+1}$ nel modo che segue. Come nel caso di $j^1(\mathbb{V}^{n+1})$ si parte da un fibrato standard. In questo caso conviene partire dal fibrato cotangente $T^*\mathbb{V}^{n+1}$ invece che da quello tangente.
Per ogni punto $p\in\mathbb{V}^{n+1}$ si considera il sottospazio $K_p\subset T_p^*\mathbb{V}^{n+1}$ delle 1-forme generate da dT_p. Quindi si considera lo spazio quoziente

$$F_p(\mathbb{V}^{n+1}) := T_p^*\mathbb{V}^{n+1}/\sim$$

dove la relazione di equivalenza in $T_p\mathbb{V}^{n+1}$ è quella associata al sottospazio K_p:

$$\omega\sim\omega' \ , \quad \text{se e solo se } \omega-\omega\in K_p.$$

Come è facile provare questo spazio eredita una struttura di spazio vettoriale reale di dimensione n da quella di $T_p^*\mathbb{V}^{n+1}$. In particolare vale che se $[\omega] \in F_p(\mathbb{V}^{n+1})$ denota la classe di equivalenza di $\omega \in T^*\mathbb{V}^{n+1}$, allora, per ogni coppia $\omega, \omega' \in T_p^*\mathbb{V}^{n+1}$ e $a, b \in \mathbb{R}$:

$$[a\omega + b\omega'] = a[\omega] + b[\omega'] \,, \tag{9.34}$$

mentre

$$[dt|_p] = 0 \,. \tag{9.35}$$

Con la definizione data si ottiene quanto segue per ogni sistema di coordinate locali naturali $(t, q^1, \ldots, q^n)$ definite attorno a $p \in \mathbb{V}^{n+1}$ con conseguente base associata $dt|_p, dq^1|_p, \ldots dq^n|_p \in T_p^*\mathbb{V}^{n+1}$. Due 1-forme

$$\omega = p_t dt|_p + \sum_{k=1}^n p_k dq^k|_p \quad \text{e} \quad \omega' = p'_t dt|_p + \sum_{k=1}^n p'_k dq^k|_p$$

individuano lo stesso elemento di $F_p(\mathbb{V}^{n+1})$ se e solo se differiscono al più nelle componenti temporali p_t e p'_t. Risulta allora chiaro che mentre $[dq^1|_p], \ldots [dq^n|_p] \in F_p(\mathbb{V}^{n+1})$ è una base di tale spazio e, cambiando base (sempre indotta da coordinate naturali su $F_p(\mathbb{V}^{n+1})$) si ha banalmente:

$$[dq'^k|_p] = \sum_{h=1}^n \left.\frac{\partial q'^k}{\partial q^h}\right|_p [dq^h|_p] \,.$$

Questo risultato segue subito dalle (9.34) e (9.35) e dalle relazioni:

$$dq'^k|_p = \sum_{h=1}^n \left.\frac{\partial q'^k}{\partial q^h}\right|_p dq^h|_p + \left.\frac{\partial q'^k}{\partial t}\right|_p dt|_p \,,$$

tenendo anche conto del fatto che la matrice dei coefficienti $\frac{\partial q'^k}{\partial q^h}$ ha determinante non nullo.
Il fibrato $F(\mathbb{V}^{n+1})$ si ottiene con due distinte fibrazioni:

(a) base data da $\mathbb{V}^{n+1}$ e fibre date dagli spazi vettoriali $F_p(\mathbb{V}^{n+1})$;

oppure

(b) base data dall'asse dei tempi $\mathbb{E}^1$ e fibre date dagli *spazi delle fasi* $\mathbb{F}_t$, per ogni tempo $t \in \mathbb{R}$,

$$\mathbb{F}_t = \left\{ (\mathcal{C}_t, p) \mid \mathcal{C}_t \in \mathbb{Q}_t \,, \quad p \in F_{\mathcal{C}_t}(\mathbb{V}^{n+1}) \right\} \,.$$

Coordinate naturali locali adattate ad entrambi i tipi di fibrazione sono quelle che associano ad ogni $(\mathcal{C}_t, p) \in F(\mathbb{V}^{n+1})$ la $2n+1$-pla $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ dove $(q^1, \ldots, q^n)$ sono le coordinate della configurazione al tempo t, $\mathcal{C}_t$ e $p = \sum_{k=1}^n p_k [dq^k|_{\mathcal{C}_t}]$.
la legge di trasformazione tra coordinate locali naturali adattate alle fibre risulta immediatamente essere proprio :

$$t' = t + c \,, \tag{9.36}$$

$$q'^k = q'^k(t, q^1, \ldots, q^n) \,, \,, \tag{9.37}$$

$$p'_k = \sum_{h=1}^n \frac{\partial q^h}{\partial q'^k} p_h \,., \tag{9.38}$$

9.3.2 Trasformazione di Legendre globale come diffeomorfismo da $j^1(\mathbb{V}^{n+1})$ a $F(\mathbb{V}^{n+1})$.

Supponiamo che sia assegnata una funzione lagrangiana $\mathscr{L} : j^1(\mathbb{V}^{n+1}) \to \mathbb{R}$ pensata come campo sclare di classe C^2 che assumeremo della solita forma standard localmente data dalla forma

$$\mathscr{L}(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n) = \sum_{hk} a_{hk}(t, q^1, \ldots, q^n,)\, \dot{q}^k\, \dot{q}^h + g(t, q^1, \ldots, q^n) \tag{9.39}$$

dove la matrice dei coefficienti a_{hk} è ovunque *strettamente definita positiva*.
Fissiamo un sistema locale di coordinate naturali adattate alle fibre di $\mathbb{V}^{n+1}$, $t, q^1, \ldots, q^n$, variabili nell'aperto $\Omega \subset \mathbb{R}^{n+1}$, e quindi, con la procedura detta sopra estendiamo tale sistema di coordinate prima sulla varietà stapzio cotangente e quindi ci riduciamo a $F(\mathbb{V}^{n+1})$, otteniamo una carta locale naturale dotata di coordinate $t, q^1, \ldots, q^n, p_1, \ldots, p_n$ variabili in $\Omega \times \mathbb{R}^n$. Siano anche $t, q^1, \ldots, q^n$, $\dot{q}^1, \ldots, \dot{q}^n$ variabili in $\Omega \subset \mathbb{R}^{n+1}$, le analoghe coordinate locali naturali su $j^1(\mathbb{V}^{n+1})$. Usando questi due sistemi di coordinate definiamo l'applicazione da $\Omega \subset \mathbb{R}^{n+1}$ nello stesso insieme data da

$$t = t\,, \tag{9.40}$$

$$q^k = q^k \quad \text{per } k = 1, \ldots, n\,, \tag{9.41}$$

$$p_k = \left.\frac{\partial \mathscr{L}}{\partial \dot{q}^k}\right|_{(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)} \quad \text{per } k = 1, \ldots, n\,. \tag{9.42}$$

Questa trasformazione, al variare delle carte locali naturali su $\mathbb{V}^{n+1}$, definisce in realtà una funzione globalmente definita su $j^1(\mathbb{V}^{n+1})$ a valori su $F(\mathbb{V}^{n+1})$. La prova di questo fatto è immediata, tenendo conto che cambiando coordinate locali su $\mathbb{V}^{n+1}$ la nuova trasformazione di coordinate corrispondente alle (9.40)-(9.42) coincide con la (9.40)-(9.42) nell'intersezione dei domini delle carte. La prova segue immediatamente dalle (9.36)-(9.38). In definitiva si definisce in questo modo un'applicazione da $j^1(\mathbb{V}^{n+1})$ a valori su $F(\mathbb{V}^{n+1})$ che si chiama **trasformazione di Legendre**. Tale trasformazione è sicuramente biettiva. Si noti infatti che, in ogni carta locale naturale, per ogni fissata $2n+1$-pla $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ che una $2n+1$-pla $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ la cui immagine secondo (9.40)-(9.42) è la prima $2n+1$-pla. Riguardo al fatto di ricoprire tutti i valori $(p_1, \ldots, p_n) \in \mathbb{R}^n$ si osservi che, per $(t, q^1, \ldots, q^n)$ fissati, la (9.42) è sicuramente biettiva per (9.39) in quanto la trasformazione considerata è lineare con matrice a determinante non nullo dato che la matrice dei coefficienti a_{hk} è strettamente positiva. La trasformazione di Legendere globalmente definita deve anche essere iniettiva come ora mostriamo. Si osservi che considernado le fibrazioni $F(\mathbb{V}^{n+1}) \to \mathbb{V}^{n+1}$ e $j^1(\mathbb{V}^{n+1}) \to \mathbb{V}^{n+1}$, se due punti di $j^1(\mathbb{V}^{n+1})$ corrispondono a punti di base differenti, la traformata di Legendre trasformerà tali punti in punti distinti perché in due fibre differenti. Di conseguenza la verifica dell'iniettività della trasformazione di Legendre si riduce alla verifica dell'iniettività per coppie di punti in $j^1(\mathbb{V}^{n+1})$ con lo stesso punto di base. Per tali coppie di punti possiamo lavorare in un'unica carta locale naturale adattata alle fibre di $j^1(\mathbb{V}^{n+1})$. Data la forma local e(9.40)-(9.42) della trasformazione di Legendre, e la banalità delle prime due equazioni, l'iniettività segue subito dal

fatto che la matrice dei coefficienti a_{hk} è strettamente positiva e quindi ha determinante non nullo.
La trasformazione inversa della trasformazione di Legendre, per costruzione, ha una struttura locale data da

$$
\begin{aligned}
t &= t\,, && (9.43)\\
q^k &= q^k \quad \text{per } k = 1, \dots, n\,, && (9.44)\\
\dot{q}^k &= \sum_{k=1}^{n} a_{hk}^{-1}(t, q^1, \dots, q^n) p_k \quad \text{per } k = 1, \dots, n\,. && (9.45)
\end{aligned}
$$

Dato che le funzioni $a_{hk} = a_{hk}(t, q^1, \dots q^n)$ sono C^1 per definizione, lo saranno anche le corrispondenti funzioni $a_{hk}^{-1}(t, q^1, \dots q^n)$. In definitiva, nelle nostre ipotesi, la trasformazione (globale) di Legendre è un diffeomorfismo di classe C^1 da $j^1(\mathbb{V}^{n+1})$ a $F(\mathbb{V}^{n+1})$.

Osservazioni 9.8. Abbiamo dimostrato sopra "a mano" che la trasformazione di Legendre è invertibile sfruttando la particolare forma della lagrangiana. Possiamo dare il seguente risultato generale riguardante l'invertibilità globale di trasformazioni definite su insiemi convessi di $\mathbb{R}^n$.

Proposizione 9.4. *Sia $F : \Omega \to \mathbb{R}^n$ una funzione di classe C^k, $k \geq 1$, con $\Omega \subset \mathbb{R}^n$ aperto e convesso*[3]*. Se la forma quadratica associata alla matrice jacobiana di F è ovunqe strettamente definita positiva allora:*

(a) *$F(\Omega) \subset \mathbb{R}^n$ è aperto,*

(b) *$F : \Omega \to F(\Omega)$ è un diffeomorfismo di classe C^k (cioé F è biettiva e la sua inversa è di classe C^k).* $\diamondsuit$

Dimostrazione. La matrice jacobiana J_p di F in ogni punto p ha determinante non nullo. Se avesse determinante nullo avrebbe anche un autovettore (non nullo) $\mathbf{u}$ con autovalore nullo: $J_p\mathbf{u} = \mathbf{0}$, ma allora sarebbe anche $Q_{J_p}(\mathbf{u}) := \mathbf{u} \cdot J_p\mathbf{u} = 0$, ovvero la forma quadratica associata a J_p si annullerebbe su un vettore non nullo e questo è impossibile dato che Q_{J_p} è strettamente definita positiva per ipotesi. Possiamo allora concludere che, per il teorema del Dini [GiustiII], la funzione F è sicuramente un diffeomorfismo locale di classe C^k nell'intorno di ogni punto di Ω. In altre parole, per ogni $p \in \Omega$, c´è un intorno aperto $U \ni p$ con $U \subset \Omega$ tale che (i) $F(U)$ è un aperto tutto contenuto in $F(\Omega)$ e (ii) $F\restriction_U : U \to F(U)$ è un diffeomorfismo di classe C^k. Quindi in particolare $F(\Omega)$ è aperto essendo unione di insiemi aperti. Per concludere la dimostrazione è sufficiente provare che F è iniettiva su Ω. Supponiamo che esistano $p, q \in \Omega$ con $p \neq q$ e $F(p) = F(q)$. Sia $\mathbf{n}$ il versore del vettore $q - p$. Allora, se $x(t) = p + t(q - p)$, la funzione $f(t) := \mathbf{n} \cdot (F(x(t)) - F(p))$ per $t \in [0, 1]$ (che è ben definita dato che Ω è convesso!) deve avere

[3] Ricordiamo che un insieme $C \subset \mathbb{R}^n$ è detto **convesso** quando, se $p, q \in C$ allora il segmento di estremi p e q è completamente contenuto in C.

derivata nulla in qualche punto $t_0 \in (0,1)$ per il teorema di Rolle:

$$f'(t_0) = \sum_{i,j=1}^{n} n^i \left.\frac{\partial F^i}{\partial x^j}\right|_{x(t_0)} n^j \, ||p-q|| = ||p-q|| Q_{J_{x(t_0)}}(\mathbf{n}) = 0 \, .$$

Questo significa $Q_{J_{x(t_0)}}(\mathbf{n}) = 0$ sul vettore non nullo $\mathbf{n}$. Questo è impossibile come già detto e pertanto i punti p e q suddetti non possono esistere. $\square$

9.3.3 Equazioni di Hamilton assegnate globalmente su $F(\mathbb{V}^{n+1})$ e campo vettoriale dinamico Z.

Quando è assegnata una lagrangiana su $j^1(\mathbb{V}^{n+1})$ viene definita un'hamiltoniana $\mathscr{H}$ ogni sistema di coordinate locali naturali di tramite la **trasformata di Legendre** della lagrangiana

$$\mathscr{H}(t,q,p) := \sum_{k=1}^{n} p_k \, \dot{q}^k(t,q,p) - \mathscr{L}(t,q,\, \dot{q}(t,q,p)) \, .$$

Un'osservazione cruciale è che, al contrario della lagrangiana, la funzione di Hamilton $\mathscr{H}$ non ha comportamento scalare cambiando sistema di coordinate naturali. Direttamente dalla definizione risulta infatti che passando ad un nuovo sistema di coordinate naturali (t', q', p') sullo spaziotempo delle fasi, nell'intersezione dei domini dei nuovo sistemi di coordinate la nuova hamiltoniana risulta essere legata alla precedente dall'equazione:

$$\mathscr{H}'(t'(t,q,p), q'(t,q,p), p'(t,q,p)) = \mathscr{H}(t,q,p) + \sum_{k=1}^{n} \left.\frac{\partial q'^k}{\partial t}\right|_{(t,q,p)} p'_k(t,q,p) \, . \tag{9.46}$$

A dispetto di tale risultato, le equazioni di Hamilton hanno significato globale su $F(\mathbb{V}^{n+1})$. Le equazioni di Hamilton determinano infatti sezioni su $F(\mathbb{V}^{n+1})$ rispetto alla base data dall'asse del tempo assoluto. Tali sezioni sono le curve integrali del campo vettoriale Z su $F(\mathbb{V}^{n+1})$, detto **campo vettoriale dinamico** (o più brevemente **vettore dinamico**), individuato in ogni sistema di coordinate locali naturali da

$$Z(t,q,p) = \frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial p_k} \frac{\partial}{\partial q^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial q^k} \frac{\partial}{\partial p_k} \, . \tag{9.47}$$

Infatti, è evidente che le linee integrali di tale campo sono tutte e sole le curve che soddisfano le equazioni di Hamilton (9.20). Il fatto che il campo Z sia ben definito, indipendentemente dal fatto che, cambiando le coordinate naturali $\mathscr{H}$ si traformi tramite le (9.46), si verifica per computo diretto, mostrando che, cambinado coordinate naturali su $\mathbb{V}^{n+1}$ ed effettuando i corrispondenti cambiamenti di coordinate naturali su $j^1(\mathbb{V}^{n+1})$ e $F(\mathbb{V}^{n+1})$ si ottiene nell'intersezione dei domini delle corodinate naturali su $F(\mathbb{V}^{n+1})$:

$$\frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial p_k} \frac{\partial}{\partial q^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial q^k} \frac{\partial}{\partial p_k} = \frac{\partial}{\partial t'} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}'}{\partial p'_k} \frac{\partial}{\partial q'^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}'}{\partial q'^k} \frac{\partial}{\partial p'_k} \, ,$$

dove le due funzioni di Hamilton sono collegate l'una all'altra tramite la (9.46).
Una verifica indiretta è invece la seguente. Dal corollario al teorema 9.1 segue che: se, per una fissata lagrangiana, una curva integrale soddisfa le equazioni Hamilton in un sistema di coordinate naturali locali definite sull'aperto $V \subset F(\mathbb{V}^{n+1})$, allora le soddisferà rispetto ad ogni altro sistema di coordinate naturali definite su V. Di conseguenza il campo vettoriale Z tangente alla curva è definito indipendentemente dalle coordinate usate. Quindi Z è ben definito su tutto $F(\mathbb{V}^{n+1})$ ed ha senso parlare di soluzione delle equazioni del moto estese a tutto lo spaziotempo delle fasi e non solo ad un sistema di coordinate locali.

Capitolo 10

Alcuni argomenti più avanzati di Meccanica Hamiltoniana.

In questo capitolo ci occuperemo di alcuni argomenti più avanzati di meccanica di Hamilton. In particolare introdurremo le trasformazioni canoniche, le parentesi di Poisson al fine di studiare, in formulazione di Hamilton, il legame tra simmetrie e leggi di conservazione. Nella prima parte introduciamo un atlante particolare dello spaziotempo delle fasi che estende quello delle coordinate naturali. Usando tale atlante riformuleremo il teorema di Liouville e ne otterremo come conseguenza il teorema del ritorno Poincaré.

10.1 Trasformazioni canoniche e loro proprietà fondamentali.

Nel precedente capitolo abbiamo visto che le equazioni di Hamilton hanno la forma:

$$\left\{ \begin{array}{rcl} \dfrac{dp_k}{dt} & = & -\dfrac{\partial \mathscr{H}}{\partial q^k}\,, \\ \dfrac{dq^k}{dt} & = & \dfrac{\partial \mathscr{H}}{\partial p_k}\,, \end{array} \right. \quad \text{per } k = 1, \ldots, n, \tag{10.1}$$

in ogni sistema locale di coordinate naturali $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ costruito sullo spaziotempo delle fasi $F(\mathbb{V}^{n+1})$. Questi sistemi di coordinate che formano un atlante dello spazio delle fasi, sono, in un certo senso "duali" dei sistemi di coordinate locali naturali su $j^1(\mathbb{V}^{n+1})$ (cfr. sezione 9.3) $(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ che formano un corrispondente atlante nello spaziotempo degli atti di moto. Cambiando queste ultime coordinate locali e passando ad un analogo sistema locale $(t', q'^1, \ldots, q'^n, \dot{q}'^1, \ldots, \dot{q}'^n)$ in modo che:

$$\begin{array}{rcl} t' & = & t + c\,, \\ q'^k & = & q'^k(t, q^1, \ldots, q^n)\,, \\ \dot{q}'^k_k & = & \displaystyle\sum_{h=1}^{n} \frac{\partial q'^k}{\partial q^h}\,\dot{q}^h + \frac{\partial q'^k}{\partial t}\,, \end{array}$$

le coordinate nello spaziotempo delle fasi cambiano di conseguenza come:

$$
\begin{aligned}
t' &= t + c\,, && (10.2)\\
q'^k &= q'^k(t, q^1, \ldots, q^n)\,, && (10.3)\\
p'_k &= \sum_{h=1}^{n} \frac{\partial q^h}{\partial q'^k} p_h\,. && (10.4)
\end{aligned}
$$

Nelle nuove coordinate naturali, le equazioni di Hamilton hanno la stessa forma di prima

$$
\begin{cases}
\dfrac{dp'_k}{dt'} &= -\dfrac{\partial \mathscr{H}'}{\partial q'^k}\,,\\
\dfrac{dq'^k}{dt'} &= \dfrac{\partial \mathscr{H}'}{\partial p'_k}\,,
\end{cases}
\quad \text{per } k = 1, \ldots, n,
$$

dove la nuova funzione Hamiltoniana è legata alla vecchia da:

$$
\mathscr{H}'(t'(t,q,p), q'(t,q,p), p'(t,q,p)) = \mathscr{H}(t,q,p) + \sum_{k=1}^{n} \frac{\partial q'^k}{\partial t}\bigg|_{(t,q,p)} p'_k(t,q,p)\,. \tag{10.5}
$$

Ci chiediamo se esistono trasformazioni di coordinate locali che preservano la forma (10.1) delle equazioni di Hamilton indipendentemente dalla struttura dello spaziotempo degli atti di moto e ragionando completamente nello spaziotempo delle fasi senza fare riferimento alle coordinate $(t, q^1, \ldots, q^n,\ \dot{q}^1, \ldots,\ \dot{q}^n)$ nello spaziotempo degli atti di moto. La risposta è positiva ed in particolare trasformazioni siffatte sono le cosiddette *trasformazioni canoniche*.

10.1.1 Trasformazioni canoniche

Definizione 10.1. (**Trasformazioni canoniche e coordinate canoniche**.) Consideriamo un sistema Hamiltoniano descritto su $F(\mathbb{V}^{n+1})$ in coordinate locali arbitrarie $(t, x^1, \ldots, x^{2n})$ dove t è il tempo assoluto e $(x^, \ldots, x^{2n})$ è , per ogni fissato t, un sistema di coordinate sulla fibra $\mathbb{F}_t$ (spazio delle fasi al tempo t).
Un nuovo sistema di coordinate locali $(t', x'^1, \ldots, x'^{2n})$ si dice connesso da una **trasformazione canonica** al precedente se la trasformazione di coordinate, supposta invertibile di classe C^2 con la sua inversa, ha la forma

$$
\begin{aligned}
t' &= t + c\,, && (10.6)\\
x'^k &= x'^k(t, x^1, \ldots, x^{2n})\,, && (10.7)
\end{aligned}
$$

in cui la matrice jacobiana delle *sole coordinate non temporali*, di coefficienti

$$
\frac{\partial x'^k}{\partial x^i}\,, \quad \text{per } i, j = 1, \ldots, 2n,
$$

appartiene al gruppo $Sp(n, \mathbb{R})$ per ogni tempo t. Nel caso in cui non compaia esplicitamente t nel secondo membro in (10.15) la trasformazione di coordinate (10.6)-(10.7) è detta **completamente canonica**.
Ogni sistema di coordinate locali su $F(\mathbb{V}^{n+1})$ connesso ad ogni sistema di coordinate locali naturali tramite una trasformazione canonica è detto sistema locale di **coordinate canoniche** su $F(\mathbb{V}^{n+1})$. $\diamondsuit$

Osservazioni 10.1.
(1) È chiaro che la composizione di trasformazioni canoniche definisce ancora una trasformazione canonica in quanto la matrice jacobiana della trasformazione risultante è semplicemente il prodotto delle due matrici jacobiane delle due trasformazioni fattore e l'insieme delle matrici simplettiche è chiuso per prodotto matriciale essendo un gruppo matriciale. Pertanto un sistema di coordinate costruito con una trasformazione canonica partendo da un sistema di coordinate canonico è anch'esso canonico.
(2) È chiaro che l'inversa di una trasformazione canonica definisce ancora una trasformazione canonica: la matrice jacobiana dell'inversa di una trasformazione canonica, in riferimento alla sottomatrice riguardante solo le coordinate non temporali, è data dall'inversa della matrice jacobiana della trasformazione di partenza in riferimento alla sottomatrice riguardante solo le coordinate non temporali. Dato che l'insieme delle matrici simplettiche è chiuso sotto inversione dei suoi elementi essendo un gruppo matriciale, la trasformazione inversa di una trasformazione canonica risulta essere ancora canonica.

Prima di tutto mostriamo che esistono sistemi di coordinate canonici e l'insieme di essi estende l'atlante dei sistemi di coordinate naturali.

Proposizione 10.1. *Nello spazio delle fasi $F(\mathbb{V}^{n+1})$ ogni sistema locale di coordinate nell'atlante dei sistemi di coordinate naturali (di classe C^2 almeno) è canonico.* $\diamondsuit$

Dimostrazione. Consideriamo la trasformazione di coordinate (10.2)-(10.4). La matrice jacobiana di tale trasformazione, considerando solo la parte relativa alle coordinate non temporali ha struttura:

$$J := \left[\begin{array}{c|c} Q & 0 \\ \hline P & Q^{-1t} \end{array}\right] , \tag{10.8}$$

Dove Q è la matrice jacobiana i cui coefficienti sono

$$Q_{ki} := \frac{\partial q'^k}{\partial q^i} ,$$

P è la matrice i cui coefficienti sono

$$P_{kr} := \sum_{l,s=1}^{n} p_s \frac{\partial^2 q^s}{\partial q'^k \partial q'^l} \frac{\partial q'^l}{\partial q^r} .$$

La forma della matrice J, per computo diretto ed usando il fatto che risulta $Q^t P = P^t Q$ dal teorema di Schwartz, implica immediatamente che:

$$J^t S J = S \, .$$

Questo prova la tesi. $\square$

Osservazioni 10.2. Ci sono sistemi di coordinate canoniche che *non sono* sistemi di coordinate naturali sullo spazio delle fasi. Basta considerare la trasformazione che, partendo dalle coordinate naturali $(t, q^1, \ldots, q^n, p_1, \ldots, p_n)$, definisce le nueve coordinate locali: $(t, Q^1, \ldots, Q^n, P_1, \ldots, P_n)$ individuate dalla trasformazione di classe C^∞ con inversa nella stessa classe:

$$\begin{aligned} t' &= t \, , \\ Q^k &= p_k \, , \quad k = 1, \ldots, n \\ P_k &= -q^k \, , \quad k = 1, \ldots, n \, . \end{aligned}$$

In questo caso la matrice jacobiana delle coordinate non temporali è proprio S e vale $S^t S S = -SSS = -S(-I) = S$. Non possiamo pensare le coordinate $(t, Q^1, \ldots, Q^n, P_1, \ldots, P_n)$ come coordinate naturali sullo spazio delle fasi e quindi connesse ad un sistema di coordinate naturali $(t, Q^1, \ldots, Q^n, \dot{Q}^1, \ldots, \dot{Q}^n)$ nello spaziotempo degli atti di moto, in quanto la legge di trasformazione tra coordinate naturali non ammette trasformazioni in cui compaia la *seconda* equazione scritta sopra. In tale equazione, a secondo membro può solo apparire una funzione delle coordinate $(t, q^1, \ldots, q^n)$.

L'insieme dei sistemi di coordinate locali canonici include l'atlante dei sistemi di coordinate locali naturali ed è esso stesso un atlante per costruzione. Tale atlante verrà detto **atlante canonico** di $F(\mathbb{V}^{n+1})$.

10.1.2 Preservazione della forma delle equazioni di Hamilton.

L'atlante canonico gode della notevole proprietà di preservare la forma delle equazioni di Hamilton. In ogni sistema di coordinate canoniche, le equazioni di Hamilton hanno sempre la forma (10.1) dove abbiamo convenzionalmente indicato con $q^1, \ldots, q^n$ le prime n coordinate non temporali e con $p_1, \ldots, p_n$ le rimanente n coordinate non temporali. La scelta di questa notazione è tuttavia abbastanza arbitraria, visto che le coordinate canoniche non sono in generali naturali e non possiamo pensare le prime n coordinate non temporali, le q^k, come coordinate nello spazio delle configurazioni. Una notazione equivalente per le equazioni di Hamilton è comunque

$$\frac{d\mathbf{x}}{dt} = S \nabla \mathscr{H}(t, \mathbf{x}) \, . \tag{10.9}$$

dove $(t, \mathbf{x}) \equiv (t, x^1, \ldots, x^{2n})$ sono le coordinate canoniche considerate.

Teorema 10.1. *Le trasformazioni canoniche preservano la forma delle equazioni di Hamilton. Più precisamente, se (t, q, p) è in un sistema di coordinate locali nelle coordinate dell'atlante canonico, allora in ogni altro sistema di coordinate locali dell'atlante canonico – ristretto se necessario in modo da prendere valori su un insieme della forma $(a, b) \times \Omega$ con $\Omega \subset \mathbb{R}^{2n}$ aperto semplicemente connesso – le soluzioni di:*

$$\begin{cases} \dfrac{dp_k}{dt} &= -\dfrac{\partial \mathscr{H}}{\partial q^k}\,, \\ \dfrac{dq^k}{dt} &= \dfrac{\partial \mathscr{H}}{\partial p_k}\,, \end{cases} \quad \textit{per } k = 1, \ldots, n, \tag{10.10}$$

per una fissata funzione Hamiltoniana $\mathscr{H}$ (di classe C^1) sono, se trasformate nelle nuove variabili, soluzioni di

$$\begin{cases} \dfrac{dp'_k}{dt} &= -\dfrac{\partial \mathscr{H}'}{\partial q'^k}\,, \\ \dfrac{dq'^k}{dt} &= \dfrac{\partial \mathscr{H}'}{\partial p'_k}\,, \end{cases} \quad \textit{per } k = 1, \ldots, n. \tag{10.11}$$

Vale inoltre:

$$\mathscr{H}'(t', q', p') = \mathscr{H}(t(t'), q(t', q', p'), p(t', q', p')) + \mathscr{K}(t', q', p') \tag{10.12}$$

dove la funzione $\mathscr{K}$ definita su $(a, b) \times \Omega$ e di classe C^1 è individuata dalla trasformazione canonica considerata e non dipende da $\mathscr{H}$ ed è definita a meno di funzioni additive della sola variabile t' e soddisfa:

$$\begin{cases} \dfrac{\partial q'_k}{\partial t} &= \dfrac{\partial \mathscr{K}}{\partial p'_k}\,, \\ \dfrac{\partial p'_k}{\partial t} &= -\dfrac{\partial \mathscr{K}}{\partial q'^k}\,, \end{cases} \quad \textit{per } k = 1, \ldots, n. \tag{10.13}$$

In particolare, se la trasformazione di coordinate tra i due sistemi di coordinate è completamente canonica, si può scegliere:

$$\mathscr{H}'(t', q', p') = \mathscr{H}(t(t'), q(t', q', p'), p(t', q', p'))\,.$$

$\diamondsuit$

Dimostrazione. Per semplicità notazionale faremo la dimostrazione nel caso in cui $t' = t$, dato che il caso generale $t' = t + c$ dove c è una costante non presenta alcun problema in più. Consideriamo una curva differenziabile $\mathbf{x} = \mathbf{x}(t)$ (non supponiamo per ora che sia soluzione delle equazioni di Hamilton (10.10)). Passiamo in coordinate primate ottenendo la curva $\mathbf{x}'(t) = \mathbf{x}'(t, \mathbf{x}(t))$. Abbiamo allora che:

$$\frac{d\mathbf{x}'}{dt} = \frac{\partial \mathbf{x}'}{\partial t} + J\frac{d\mathbf{x}}{dt}\,,$$

dove $J(t,\mathbf{x})$ è la matrice jacobiana i cui coefficienti sono le derivate $\partial x'^i/\partial x^k$, valutata sulla funzione $\mathbf{x} = \mathbf{x}(t)$. Tale curva soddisfa le equazioni di Hamilton nelle coordinate iniziali se e solo se

$$\frac{d\mathbf{x}}{dt} = S\nabla_x \mathscr{H}(t,\mathbf{x}(t)) \,.$$

Pertanto la curva $\mathbf{x}' = \mathbf{x}'(t)$ soddisfa le equazioni di Hamilton nelle coordinate non primate se e solo se

$$\frac{d\mathbf{x}'}{dt} = \frac{\partial \mathbf{x}'}{\partial t} + JS\nabla_{\mathbf{x}} \mathscr{H}(t,\mathbf{x}(t)) \,.$$

Essendo

$$\frac{\partial}{\partial x^k} = \sum_{i=1}^{2n} \frac{\partial x'^i}{\partial x^k}\frac{\partial}{\partial x'^i} \,,$$

deve essere:

$$\frac{d\mathbf{x}'}{dt} = \frac{\partial \mathbf{x}'}{\partial t} + JSJ^t\nabla_{\mathbf{x}'} \mathscr{H}(t,\mathbf{x}(t,\mathbf{x}'(t))) \,.$$

Dato che $JSJ^t = S$ come provato in Proposizione 9.2 (in quanto $J \in Sp(n,\mathbb{R})$ per ipotesi e quindi $J^t \in Sp(n,\mathbb{R})$), la curva $\mathbf{x}' = \mathbf{x}'(t)$ soddisfa le equazioni di Hamilton nelle coordinate iniziali se esolo se:

$$\frac{d\mathbf{x}'}{dt} = \frac{\partial \mathbf{x}'}{\partial t} + S\nabla_{\mathbf{x}'} \mathscr{H}(t,\mathbf{x}(t,\mathbf{x}'(t))) \,.$$

Nel caso la trasformazione di coordinate $\mathbf{x}' = \mathbf{x}'(t,\mathbf{x})$ non dipenda in realtà dal tempo, la dimostrazione è conclusa definendo

$$\mathscr{H}'(t,\mathbf{x}') := \mathscr{H}(t,\mathbf{x}(t,\mathbf{x}')) \,.$$

Per concludere nel caso generale è sufficiente verificare che la funzione:

$$\mathbf{F}(t,\mathbf{x}') := \frac{\partial \mathbf{x}'(t,\mathbf{x})}{\partial t}|_{\mathbf{x}=\mathbf{x}(t,\mathbf{x}')}$$

si possa sempre scrivere, nelle nostre ipotesi, nella forma:

$$\mathbf{F}(t,\mathbf{x}') = S\nabla_{\mathbf{x}'} \mathscr{K}(t,\mathbf{x}')$$

per qualche funzione scalare $\mathscr{K} = \mathscr{K}(t,\mathbf{x}')$. Si osservi che ciò equivale a dimostrare (10.13). In virtù dell'osservazione 9.6 questo è vero se e solo se per ogni $(t,\mathbf{x}')$ la matrice i cui coefficienti sono

$$V_{ij}(t,\mathbf{x}') := \sum_{k=1}^{2n} \frac{\partial^2 x'^i(t,\mathbf{x})}{\partial t \partial x^k}|_{\mathbf{x}=\mathbf{x}(t,\mathbf{x}')} \frac{\partial x^k}{\partial x'^j}|_{(t,\mathbf{x}')}$$

è hamiltoniana. In altre parole dobbiamo verificare che la matrice

$$V(t,\mathbf{x}') := \frac{\partial J}{\partial t} J^{-1}|_{(t,\mathbf{x}')}$$

sia hamiltoniana per ogni $(t, \mathbf{x}')$ nell'ipotesi che la trasformazione di coordinate sia canonica. Derivando in t i due membri dell'identità – che sappiamo essere verificata nelle nostre ipotesi – :

$$J^t S J = S$$

si trova:

$$\frac{\partial J^t}{\partial t} S J + J^t S \frac{\partial J}{\partial t} = 0 \,.$$

Questa si può scrivere come, moltiplicando per J^{-1} a destra e J^{-1t} a sinistra:

$$J^{-1t} \frac{\partial J^t}{\partial t} S + S \frac{\partial J}{\partial t} J^{-1} = 0 \,.$$

Ossia:

$$\left(\frac{\partial J}{\partial t} J^{-1} \right)^t S + S \frac{\partial J}{\partial t} J^{-1} = 0 \,.$$

che è quanto volevamo. Si osservi che $\mathscr{K}$ è indipendente da $\mathscr{H}$ per costruzione ed è definita (dato che si ottiene integrando un campo vettoriale irrotazionale) a meno di costanti additive per ogni fissato istante di tempo. Variando il tempo $\mathscr{K}$ può quindi essere ridefinita al più aggiungendo una funzione arbitraria della sola variabile temporale. □

Osservazioni 10.3.
(1) Dalla dimostrazione si evince che, nel caso in cui la trasformazione canonica sia completamente canonica, possiamo lasciare cadere l'ipotesi della forma $(a, b) \times \Omega$, con Ω aperto semplicemente connesso, dell'insieme su cui prendono i valori le coordinate $(t', \mathbf{x}')$.
(2) Ci si può chiedere se la più generale trasformazione di coordinate che preserva la forma Hamiltoniana delle equazioni del moto sia una trasformazione canonica. La risposta è negativa. A titolo di esmpio, la trasformazione:

$$\begin{aligned} t' &= t \,, \\ q'^k &= \lambda q^k \,, \quad k = 1, \ldots, n \\ p'_k &= \mu p_k \,, \quad k = 1, \ldots, n \,. \end{aligned}$$

dove $\lambda, \mu > 0$ sono costanti, preserva banalmente la forma delle equazioni del moto, pur di definire:

$$\mathscr{H}'(t', q', p') = \lambda \mu \mathscr{H}(t(t'), q(t', q', p'), p(t', q', p')) \,.$$

Tuttavia $J^t S J = \mu \lambda S$, per cui non si ha una trasformazione canonica se $\mu\lambda \neq 1$. In effetti si riesce a dimostrare che una trasformazione differenzibile con inversa differenziabile (entrambe di classe C^2):

$$t' = t + c \,, \tag{10.14}$$

$$x'^k = x'^k(t, x^1, \ldots, x^{2n}) \,, \tag{10.15}$$

preserva la forma hamiltoniana delle equazioni del moto se e solo se la matrice jacobiana delle sole coordinate non temporali associata alla trasformazione soddisfa, per ogni $(t, \mathbf{x})$, $J^t S J = aS$ per qualche costante $a \neq 0$.
(3) Il fatto che la funzione hamiltoniana *non* abbia comportamento scalare sotto trasformazione di coordinate canoniche, ma riceva un addendo della forma $\mathscr{K}$ come in (10.12), suggerisce, per una fissata $\mathscr{H}$, di cercare un sistema di coordinate canoniche in cui $\mathscr{H}'$ sia identicamente nulla. Questo equivale evidentemente a risolvere le equazioni di Hamilton per ogni condizione iniziale, dato che nelle nuove coordinate canoniche le equazioni di Hamilton risultano essere banali ${q'}^k = {q'}_0^k$ costante, ${p'}_k = {p'}_{0k}$ costante. Il metodo che parte da questa osservazione, sviluppato da Jacobi, conduce alla famosa equazione di Hamilton-Jacobi.
(4)* Quando è data su $F(\mathbb{V}^{n+1})$ un'assegnazione di hamiltoniane $\mathscr{H}$ indotta da una funzione lagrangiana $\mathscr{L}$ su $j^1(\mathbb{V}^{n+1})$, le soluzioni delle equazioni di Hamilton sono, come visto nella sezione 9.3 del capitolo precedente, le curve integrali del campo vettoriale Z *globalmente definito* su $F(\mathbb{V}^{n+1})$, detto **campo vettoriale dinamico** (o più brevemente **vettore dinamico**). Z è individuato in ogni sistema di coordinate locali naturali da

$$Z(t, q, p) = \frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial p_k} \frac{\partial}{\partial q^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial q^k} \frac{\partial}{\partial p_k} \,. \tag{10.16}$$

Infatti, è evidente che le linee integrali di tale campo sono tutte e sole le curve che soddisfano le equazioni di Hamilton (10.11). A differenza della situazione descritta nel capitolo precedente i sistemi di coordinate locali che ora possiamo usare sullo spaziotempo delle fasi non sono solo quelli naturali, ma possiamo fare uso anche dei sistemi di coordinate nell'atlante canonico. È chiaro che, per merito del Teorema 10.1 e tenendo conto che le coordinate naturali su $F(\mathbb{V}^{n+1})$ sono una sottoclasse delle coordinate canoniche, il vettore dinamico avrà sempre la stessa espressione (10.16) in ogni sistema locale di coordinate canonico.

Per concludere citiamo un utile corollario tecnico del teorema precedente.

Corollario. *Nelle ipotesi del teorema precedente vale l'identità tra operatori differenziali sull'intersezione delle carte locali canoniche* (t', q', p') *e* (t, q, p)*:*

$$\sum_{k=1}^{n} \left(\frac{\partial \mathscr{K}}{\partial q'^k} \frac{\partial}{\partial p'_k} - \frac{\partial \mathscr{K}}{\partial p'_k} \frac{\partial}{\partial q'^k} \right) = - \sum_{k=1}^{n} \left(\frac{\partial q'^k}{\partial t} \frac{\partial}{\partial q'^k} + \frac{\partial p'_k}{\partial t} \frac{\partial}{\partial p'_k} \right) \,. \tag{10.17}$$

Dimostrazione La dimostrazione di (10.17) è ovvia: la stessa identità si riscrive equivalentemente

$$\sum_{k,h=1}^{n} S^{kh} \frac{\partial \mathscr{K}}{\partial x'^k} \frac{\partial}{\partial x'^h} = - \sum_{k=1}^{n} \frac{\partial x'^k}{\partial t} \frac{\partial}{\partial x'^k} \,,$$

e questa segue immediatamente da:

$$\frac{\partial \mathbf{x}'}{\partial t} = S \nabla_{\mathbf{x}'} \mathscr{K}$$

che è un modo sintetico di scrivere (10.13). □

10.1.3 *Struttura globale delle equazioni di Hamilton in riferimento all'atlante canonico di $F(\mathbb{V}^{n+1})$.

Mostriamo ora come le proprietà delle trasformazioni canoniche consentano di emancipare la formulazione di Hamilton della meccanica da quella di Lagrange. Indipendentemente da (4) in osservazioni 10.3, il fatto che si possa definire un campo vettoriale Z globalmente ben definito su tutto lo spaziotempo delle fasi le cui curve integrali siano tutte e sole le soluzioni delle equazioni di Hamilton e che quindi le equazioni di Hamilton abbiano significato globale, *prescinde dall'assegnazione di una lagrangiana su $j^1(\mathbb{V}^{n+1})$ dalla quale ricavare le funzioni di Hamilton in $F(\mathbb{V}^{n+1})$*. In questo senso è possibile partire direttamente dalla formulazione hamiltoniana della meccanica senza fare riferimento ad una preesistente formulazione lagrangiana ed assegnando direttamente le funzioni di Hamilton sullo spazio delle fasi. Diamo la prova di ciò nel seguente importante teorema che è in realtà un corollario del Teorema 10.1.

Teorema. *Si supponga di assegnare una funzione $\mathscr{H}$ di classe C^2 in ogni carta dell'atlante canonico su $F(\mathbb{V}^{n+1})$ in modo tale che, cambiando coordinate locali in tale atlante le funzioni $\mathscr{H}$ si trasformino secondo l'equazione (10.12) con $\mathscr{K}$ che verifica (10.17). Valgono i seguenti fatti.*
(a) *C'è un campo vettoriale Z di classe C^1 globalmente definito su $F(\mathbb{V}^{n+1})$ (e detto* **campo vettoriale dinamico***) che assume la stessa forma*

$$Z = \frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial p_k}\frac{\partial}{\partial q^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial q^k}\frac{\partial}{\partial p_k}$$

in ogni sistema locale di coordinate canoniche.
(b) *Le curve integrali di tale campo sono le soluzioni dielle equazioni di Hamilton in ogni carta dell'atlante canonico rispetto alla corrispondente funzione hamiltoniana $\mathscr{H}$.* $\diamondsuit$

Dimostrazione. (a) Prima di tutto notiamo che se $(t, q^1, \ldots, q^n, p_1, \ldots, p_n) \equiv (t, \mathbf{x})$, abbiamo che:

$$\frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial p_k}\frac{\partial}{\partial q^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial q^k}\frac{\partial}{\partial p_k} = \frac{\partial}{\partial t} + \sum_{i,j=1}^{n} S^{ij}\frac{\partial \mathscr{H}}{\partial x^i}\frac{\partial}{\partial x^j}$$

è un'espressione equivalente per Z. Consideriamo ora due sistemi locali di coordinate canoniche $(t, \mathbf{x})$ e $(t', \mathbf{x}')$. Usando il fatto che la trasformazione tra sistemi di coordinate canoniche è tale che la matrice jacobiana delle coordinate non temporali preserva la matrice S, si ha subito che:

$$\frac{\partial}{\partial t} + \sum_{i,j=1}^{n} S^{ij}\frac{\partial \mathscr{H}}{\partial x^i}\frac{\partial}{\partial x^j} = \frac{\partial}{\partial t} + \sum_{i,j,h,k=1}^{n} S^{ij}\frac{\partial x'^k}{\partial x^i}\frac{\partial x'^h}{\partial x^j}\frac{\partial \mathscr{H}}{\partial x'^h}\frac{\partial}{\partial x'^k}\,,$$

ossia

$$\frac{\partial}{\partial t} + \sum_{i,j=1}^{n} S^{ij}\frac{\partial \mathscr{H}}{\partial x^i}\frac{\partial}{\partial x^j} = \frac{\partial}{\partial t} + \sum_{h,k=1}^{n} S^{hk}\frac{\partial \mathscr{H}}{\partial x'^h}\frac{\partial}{\partial x'^k}\,,$$

Dove, in conformità con il teorema precedente $\mathscr{H} = \mathscr{H}' - \mathscr{K}$. Usando l'identità :

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial x'^k}{\partial t} \frac{\partial}{\partial x'^k} ,$$

possiamo scrivere che:

$$\frac{\partial}{\partial t} + \sum_{i,j=1}^{n} S^{ij} \frac{\partial \mathscr{H}}{\partial x^i} \frac{\partial}{\partial x^j} = \frac{\partial}{\partial t'} - \sum_{k=1}^{n} \frac{\partial x'^k}{\partial t} \frac{\partial}{\partial x'^k} + \sum_{h,k=1}^{n} S^{hk} \frac{\partial \mathscr{H}'}{\partial x'^h} \frac{\partial}{\partial x'^k} - \sum_{h,k=1}^{n} S^{hk} \frac{\partial \mathscr{K}}{\partial x'^h} \frac{\partial}{\partial x'^k} .$$

Valendo (10.17) abbiamo che:

$$\sum_{k,h=1}^{n} S^{kh} \frac{\partial \mathscr{K}}{\partial x'^k} \frac{\partial}{\partial x'^h} = - \sum_{k=1}^{n} \frac{\partial x'^k}{\partial t} \frac{\partial}{\partial x'^k} ,$$

e pertanto, sostituendo sopra, abbiamo che:

$$\frac{\partial}{\partial t} + \sum_{i,j=1}^{n} S^{ij} \frac{\partial \mathscr{H}}{\partial x^i} \frac{\partial}{\partial x^j} = \frac{\partial}{\partial t'} + \sum_{h,k=1}^{n} S^{hk} \frac{\partial \mathscr{H}}{\partial x'^h} \frac{\partial}{\partial x'^k} .$$

Di conseguenza il campo vettoriale

$$Z := \frac{\partial}{\partial t} + \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial p_k} \frac{\partial}{\partial q^k} - \sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial q^k} \frac{\partial}{\partial p_k}$$

è in realtà ben definito su tutta la varietà spazitempo delle fasi dato che (1) è assegnato su ogni carta locale di un atlante e (2) nelle intersezioni di due carte risulta essere indipendente dalla carta scelta. La prova di (b) è ovvia data la forma del campo Z. □

10.2 *Il teorema di Liouville in forma globale ed il teorema "del ritorno" di Poincaré.

Il formalismo introdotto nella sezione precedente permette di dare la versione globale del teorema di Liouville sullo spaziotempo delle fasi $F(\mathbb{V}^{n+1})$ senza alcuna restrizione alla forma di questo spazio. Una conseguenza di tale formulazione è il teorma del ritorno di Poincaré che esamineremo di seguito.

10.2.1 Teorema di Liouville e l'equazione di Liouville.

Un'altra proprietà fondamentale delle trasformazioni canoniche è che esse preservano il volume di ogni spazio delle fasi. Fissiamo una fibra $\mathbb{F}_{\tau_0}$ al tempo τ_0 e consideriamo un sistema locale di coordinate canoniche $(t, \mathbf{x}) \equiv (t, q^1, \ldots, q^n, p_1, \ldots, p_n)$ definito nell'intorno $\mathcal{O}$ di un punto

$p \in \mathbb{F}_{\tau_0}$. Passiamo ad un altro analogo sistema di coordinate canoniche definite nell'intorno $\mathcal{O}$ dello stesso punto: $(t'\mathbf{x}') \equiv (t', q'^1, \ldots, q'^n, p'_1, \ldots, p'_n)$. Se $D \subset \mathbb{F}_{\tau_0}$ è un insieme misurabile contenuto in $\mathcal{O}$, possiamo ad esso associare due misure di volume:

$$\int_{\mathbf{x}(D)} dq^1 \cdots dq^n dp_1 \cdots dp_n \quad \text{e} \quad \int_{\mathbf{x}'(D)} dq'^1 \cdots dq'^n dp'_1 \cdots dp'_n \,.$$

In virtù della canonicità della trasformazione di corodinate, in realtà le due misure coincidono. Infatti:

$$\int_{\mathbf{x}'(D)} dq'^1 \cdots dq'^n dp'_1 \cdots dp'_n = \int_{\mathbf{x}(D)} |\det J(t, q^1, \ldots, q^n, p_1, \ldots, p_n)| \, dq^1 \cdots dq^n dp_1 \cdots dp_n$$

J è la matrice jacobiana della trasformazione di coordinate che deve avere determinante di modulo 1 in virtù del fatto che $J^t S J = S$ per cui $(\det J)^2 = 1$ come già osservato. In definitiva:

$$\int_{\mathbf{x}(D)} dq^1 \cdots dq^n dp_1 \cdots dp_n = \int_{\mathbf{x}'(D)} dq'^1 \cdots dq'^n dp'_1 \cdots dp'_n \,.$$

I sistemi di coordinate canonici individuano localmente una misura naturale intrinseca sugli spazi delle fasi. Usando tecniche standard di teoria della misura su varietà differenzibili che fanno uso di una partizione dell'unità , si arriva per questa via a definire una **misura canonica** μ_τ globale su ogni fibra $\mathbb{F}_\tau$ dello spazio delle fasi individuata completamente dall'atlante canonico di $F(\mathbb{V}^{n+1})$. Tale misura è una misura di Borel positiva per costruzione che coincide con la misura di Lebesgue di $\mathbb{R}^{2n}$ quando ci si restringe ad ogni sistema di coordinate canoniche. Di conseguenza con μ possiamo misurare in particolare gli insiemi aperti che non possono avere misura nulla ed i compatti che devono avere misura finita.
Usando questa misura canonica si giunge infine, sempre con tecniche standard e facendo uso del teorema di Liouville su $\mathbb{R} \times \mathbb{R}^{2n}$ dimostrato nel precedente capitolo, a dimostrare il teorema di Liouville nella forma generale.

Teorema 10.2. (**Teorema di Liouville**.) *In riferimento ad un sistema fisico descritto nello spazio delle fasi $F(\mathbb{V}^{n+1})$ con fibre (spazi delle fasi al tempo τ) $\mathbb{F}_\tau$, si supponga che le soluzioni delle equazioni di Hamilton – cioé le linee integrali del campo vettoriale globale Z (9.47) – siano complete e pertanto sia definito, da tali soluzioni, un gruppo (globale) ad un parametro $\{\Phi_s\}_{s\in\mathbb{R}}$ su $F(\mathbb{V}^{n+1})$ – il* **flusso di fase**.
Il flusso di fase conserva il volume canonico delle fibre: se $D_{t_0} \subset \mathbb{F}_{t_0}$ è un insieme misurabile e per ogni $s \in \mathbb{R}$, $(\{t\} \times D_t) = \Phi_s(\{t_0\} \times D_{t_0})$ dove $t = t_0 + s$, vale:

$$\int_{D_t} d\mu_t = \int_{D_{t_0}} d\mu_{t_0} \,.$$

$\diamondsuit$

La presenza di una misura naturale nello spazio delle fasi consente di dare senso alla nozione di distribuzione di probabiulità che ha avuto fondamentali applicazioni in fisica. Il teorema enunciato sopra è infatti il punto di partenza della cosiddetta *meccanica statistica*. Si tratta della teoria matematica che descrive un insieme statistico di sistemi fisici identici e quindi soggetti alla stessa dinamica hamiltoniana, ma descritti da punti rappresentativi differenti istante per istante a causa di differenti condizioni iniziali. Un tale insieme di sistemi fisici può descrivere la situazione nella quale noi non conosciamo con precisione lo stato del sistema ad ogni singolo istante. Tale approccio è anche usato per dare una descrizione miscroscopica alla termodinamica dei sistemi all'equilibrio, quando si assume che lo stato macroscopico del sistema all'equilibrio corrisponda ad una descrizione mediata su intervalli di tempo lunghi dello stato microscopico del sistema, e che tale operazione di media temporale sia equivalente all'assegnazione, istante per istante, di una distribuzione di stati microscopici (punti rappresentativi) invece che un unico punto rappresentativo (ipotesi ergodica).
Nella meccanica statistica si assegna una densità di probabilità $\rho = \rho(t,q,p) \geq 0$ (funzione di calsse C^1 almeno) nello spaziotempo delle fasi il cui integrale all'istante t, sull'insieme misurabile $D_t \subset \mathbb{F}_t$, integrale valutato *rispetto alla misura canonica* μ, definisce la probabilità $P_t(D_t)$ che il punto rappresentativo del sistema fisico si trovi in D_t al tempo t:

$$P_t(D_t) := \int_{D_t} \rho(t,q,p)\, d\mu_t(t,q,p)\,. \tag{10.18}$$

La richiesta naturale che si deve imporre sulla funzione ρ è che, individuato all'istante t_0 un insieme misurabile $D_{t_0} \subset \mathbb{F}_{t_0}$ di punti rappresentativi con probabilità $P_{t_0}(D_{t_0})$, la probabilità $P_t(D_t)$, per ogni istante $t \neq t_0$, calcolata su D_t ottenuto da D_{t_0} attraverso il flusso di fase, soddisfi:

$$P_t(D_t) = P_{t_0}(D_{t_0})\,. \tag{10.19}$$

In particolare, se $D_t = \mathbb{F}_t$ tale probabiltà deve valere sempre 1.
La richista (10.19) può essere trascritta equivalentemente come:

$$\frac{d}{dt}\int_{D_t} \rho(t,q,p)\, d\mu_t(t,q,p) = 0\,.$$

Questa equazione, restringendosi a lavorare per tempi ed insiemi D_t a chiusura compatta, comperti da coordinate canoniche, può scriversi equivalentemente:

$$\frac{d}{dt}\int_{D_{t_0}} \rho(t, q(t,q_0,p_0), p(t,q_0,p_0))\, |\mathrm{det} J_t|\, d\mu_t(t_0,q_0,p_0) = 0\,,$$

dove $q(t,q_0,p_0), p(t,q_0,p_0)$ sono la posizione nello spaziotempo delle fasi al tempo t del punto rappresentativo che evolve soddisfacendo le equazioni di Hamilton, partendo dalla condizione iniziale (q_0,p_0) al tempo t_0. Tenendo conto che $|\mathrm{det} J_t|$ è costante, l'identità di sopra si trascrive equivalentemente passando la derivata sotto il segno di integrale:

$$\int_{D_{t_0}} \frac{d}{dt}\left(\rho(t, q(t,q_0,p_0), p(t,q_0,p_0))\right)\, |\mathrm{det} J_t|\, d\mu_t(t_0,q_0,p_0) = 0\,,$$

Per l'arbitrarietà di D_{t_0} (che possiamo scegliere essere una palla coordinata sufficientemente piccola) e usando la continuità della funzione integranda, la richiesta è equivalente a[1]:

$$\frac{d}{dt}\rho(t, q(t, q_0, p_0), p(t, q_0, p_0)) = 0$$

Questa equazione dice che lungo ogni linea di universo che soddisfa le equazioni di Hamilton, il valore di ρ deve rimanere costante (ma può dipendere dalla linea di universo). In forma esplicita ed in coordinate canoniche, l'equazione scritta sopra prende la forma

$$\frac{\partial \rho}{\partial t} = \sum_{k=1}^{n} \left(\frac{\partial \rho}{\partial p_k} \frac{dp_k}{dt} - \frac{\partial \rho}{\partial q^k} \frac{dq^k}{dt} \right) \quad \text{in ogni punto dello spazio delle fasi. .}$$

Facendo infine uso delle equazioni di Hamilton troviamo infine la celebre **equazione di Liouville**:

$$\frac{\partial \rho}{\partial t} = \sum_{k=1}^{n} \left(\frac{\partial \rho}{\partial p_k} \frac{\partial \mathscr{H}}{\partial q^k} - \frac{\partial \rho}{\partial q^k} \frac{\partial \mathscr{H}}{\partial p_k} \right) \quad \text{in ogni punto dello spazio delle fasi .} \tag{10.20}$$

Si può verificare, procedendo in senso contrario ed usando una partizione dell'unità , che la validità dell'equazione di Liouville è anche sufficiente affinché valga la (10.19).

10.2.2 Il teorema "del ritorno" (o "di ricorrenza") di Poincaré .

Consideriamo il caso di un sistema dinamico descritto nello spaziotempo delle fasi $F(\mathbb{V}^{n+1})$ che possa decomporsi canonicamente nel prodotto cartesiano $\mathbb{R} \times \mathbb{F}$ dando luogo ad un flusso di fase completo, nel senso che valgano le seguenti tre richieste.
(R1) Sia possibile ricoprire $F(\mathbb{V}^{n+1})$ con un atlante di sistemi di coordinate canonici in cui la coordinata temporale vari su tutto $\mathbb{R}$ ed in modo tale che le trasformazioni canoniche tra i sistemi di coordinate dell'atlante siano completamente canoniche.
(R2) Le funzioni hamiltoniane del sistema (pensate come funzioni di classe C^2) siano indipendenti dal tempo in ciascuna di tali carte locali. (In questa situazione la funzione hamiltoniana si estende ad un campo scalare indipendente dal tempo su tutto lo spaziotempo delle fasi. Le equazioni di Hamilton risultano essere in questa situazione, un sistema di equazioni differenziali autonomo sulla varietà spazio delle fasi $M := \mathbb{F}$ nel senso discusso nella sezione 3.5).
(R3) Le soluzioni delle equazioni di Hamilton siano complete.

Come chiarito nella sezione 3.5.3, in questo modo viene a definirsi un gruppo ad un parametro di diffeomorfismi di $\mathbb{F}$, detto ancora **flusso di fase**, $\{\phi_t\}_{t\in\mathbb{R}}$ con $\phi_t : \mathbb{F} \to \mathbb{F}$, definito dalla richiesta che $\mathbb{R} \ni t \mapsto \phi_t(\mathbf{x})$ sia l'unica soluzione delle equazioni di Hamilton con condizione iniziale

[1]Se, per una funzione continua $g : D \to \mathbb{R}$ con $D \subset \mathbb{R}^n$ aperto, vale $\int_A g(x)d^n x = 0$ per ogni insieme misurabile $A \subset D$ allora deve essere $g(x) = 0$ ovunque su D. Se fosse $g(x_0) > 0$ in x_0, la funzione dovrebbe verificare $0 < \epsilon_0 < g(x_0) - \epsilon < g(x) < g(x_0) + \epsilon$ in un intorno aperto A_{x_0} di x_0. In tale intorno avremmo: $\int_{A_{x_0}} g(x)d^n x \geq \epsilon_0 \int_{A_{x_0}} d^n x > 0$ che contraddice le ipotesi. Se fosse $g(x_0) < 0$, usando $-g$ si arriverebbe allo stesso assurdo.

$\mathbf{x}_0 \in \mathbb{F}$ al tempo $t = 0$. Sappiamo dalla sezione precedente che i diffeomorfismi ϕ_t preservano il volume canonico dei sottoinsiemi misurabili di $\mathbb{F}$. Vale allora il famoso *teorema "del ritorno" o "di ricorrenza" di Poincaré.*

Teorema 10.3. (**"Teorema del ritorno", Poincaré**.) *Si consideri un sistema fisico descritto in formulazione di Hamilton e che soddisfi le richieste (R1), (R2) e (R3) e $\mathcal{R} \subset \mathbb{F}$ verifichi le seguenti ipotesi:*
(i) $\phi_t(\mathcal{R}) \subset \mathcal{R}$ per ogni $t \in \mathbb{R}$;
(ii) $\mathcal{R}$ sia misurabile secondo la misura (di Borel) canonica μ dello spazio delle fasi ed abbia misura finita: $\mu(\mathcal{R}) < +\infty$.
In questo caso, per ogni intorno aperto arbitrariamente piccolo $\mathcal{G} \subset \mathcal{R}$ valgono i seguenti fatti.

(a) *L'insieme $\mathcal{G}$ "ritorna infinite volte" in $\mathcal{G}$ stesso. Più precisamente, esiste una successione di tempi $0 < \tau_1 < \tau_2 < \cdots \to +\infty$ per cui $\mathcal{G} \cap \phi_{\tau_n}(\mathcal{G}) \neq \emptyset$ dove $\mu\left(\mathcal{G} \cap \phi_{\tau_n}(\mathcal{G})\right) > 0$ per ogni $n \in \mathbb{N}$.*

(b) *I punti di $\mathcal{G}$ che "non ritornano mai più " in $\mathcal{G}$ formano un sottoinsieme di misura nulla:*

$$\mu\left(\{p \in \mathcal{G} \mid \phi_t(p) \notin \mathcal{G} \quad t > 0\}\right) = 0\,. \qquad \diamondsuit$$

Dimostrazione. Dimostriamo il teorema premettendo un lemma.

Lemma 10.1. *Se $A \subset \mathcal{R}$ è misurabile con $\mu(A) > 0$, consideriamo gli insiemi*

$$A_n := \phi_{nt}(A) := (\phi_t)^n(A)$$

dove $n \in \mathbb{N}$ e $t > 0$ è una costante fissata arbitrariamente. Allora, per qualche $k_t \in \mathbb{N} \setminus \{0\}$, deve valere $A \cap A_{k_t} \neq \emptyset$ ed inoltre $\mu(A \cap A_{k_t}) > 0$.

Dimostrazione del Lemma. Gli insiemi A_m sono misurabili perchè controimmagine di insiemi misurabili secondo funzioni continue (che sono misurabili essendo la misura μ di Borel). Definiamo, per $n \in \mathbb{N} \setminus \{0\}$, gli insiemi $B_n := A_n \setminus (\cup_{m \neq n} A_m)$. Supponiamo che $\mu(A_n \cap A_m) = 0$ per ogni coppia m, n con $n \neq m$. Allora abbiamo che $\mu(A_n) = \mu(B_n)$ inoltre, dato che $B_m \cap B_n = \emptyset$ se $n \neq m$:

$$\mu\left(\bigcup_m B_n\right) = \sum_m \mu(B_m) = \sum_m \mu(A_m) = \sum_m \mu(A) = +\infty\,.$$

Questo è impossibile visto che

$$\mu\left(\bigcup_m B_n\right) \subset \mu(\mathcal{R}) < +\infty$$

dal momento che $\cup_m B_n \subset \mathcal{R}$. Allora deve accadere che $\mu(A_n \cap A_m) > 0$ per qualche m, n con $m \neq n$. Se n è il minimo tra m e n, applicando $(\phi_{nt})^{-1}$ e tenendo conto che tale trasformazione preserva il volume, abbiamo che

$$\mu\left(A \cap (\phi_{nt})^{-1}(A_m)\right) = \mu\left((\phi_{nt})^{-1}(A_n \cap A_m)\right) = \mu(A_n \cap A_m) > 0\,.$$

Si osservi che risulta $A_{k_t} = \phi_{nt}^{-1}(A_m)$ dove $k_t := m - n > 0$. Abbiamo quindi trovato che $\mu(A \cap A_{k_t}) > 0$. $\square$

Il lemma precedente prova la proposizione (a) scegliendo $A = \mathcal{G}$. La successione $\tau_1 < \tau_1 < \cdots$ si ottiene applicando il lemma per tempi $t_1 < t_2 < t_3, \ldots \to +\infty$ in modo tale che $t_2 > k_{t_1} t_1$, $t_3 > k_{t_2} t_2$ e via di seguito scegliendo in generale $t_n > k_{t_{n-1}} t_{n-1}$, e definendo infine $\tau_1 := k_{t_1} t_1$, $\tau_2 := k_{t_2} t_2$ e in generale $\tau_n := k_{t_n} t_n$. Per costruzione $0 < \tau_1 < \tau_2 < \cdots \to +\infty$.
Passiamo alla proposizione (b). Consideriamo quindi il sottoinsieme $E \subset \mathcal{G}$ dei punti $p \in \mathcal{G}$ che non ritornano in $\mathcal{G}$ sotto evoluzione discreta: $\phi_{nt}(p) \notin \mathcal{G}$ per ogni $n = 1, 2, \ldots$ per qualche $t > 0$ fissato. L'insieme E è misurabile dato che può essere scritto come intersezione numerabile di insiemi misurabili:

$$E = \bigcap_{n \in \mathbb{N} \setminus \{0\}} (\phi_{nt})^{-1}(\mathcal{R} \setminus \mathcal{G}) \,.$$

Se fosse $\mu(E) > 0$, applicando il lemma precedente per $A = E$ avremmo che $\mathcal{G} \cap \phi_{kt}(E) \supset E \cap \phi_{kt}(E) \neq \emptyset$ per qualche $k \in \mathbb{N} \setminus \{0\}$ e questo è impossibile per definizione di E. Quindi deve essere $\mu(E) = 0$. L'insieme E_0 dei punti di $\mathcal{G}$ che non ritornano mai più in $\mathcal{G}$:

$$E_0 := \{p \in \mathcal{G} \mid \phi_t(p) \notin \mathcal{G} \quad t > 0\}$$

soddisfa banalmente $E_0 \subset E$. Dato che la misura μ è completa per costruzione, deve essere E_0 misurabile e quindi $\mu(E) = \mu(E_0) = 0$ $\square$

Esempi 10.1.
1. Sistemi hamiltoniani che soddisfano le ipotesi del teorema di Poincaré sono abbastanza comuni. Prendiamo un sistema fisico sottoposto a forze indipendenti dal tempo in un riferimento inerziale che ammetta descrizione di Lagrange e quindi di Hamilton. Queste ipotesi, lavorando in coordinate solidali con il riferimento inerziale implicano la validità delle richieste (R1) e (R2). Supponiamo inoltre che il sistema fisico ammetta un integrale primo $H : \mathbb{F} \to \mathbb{R}$, tipicamente l'energia, le cui superfici di livello nello spazio delle fasi, ottenute fissando un valore generico dell'integrale primo

$$\Sigma_h := \{\mathbf{x} \in \mathbb{F} \mid H(\mathbf{x}) = h\}$$

siano insiemi compatti. Sotto queste ipotesi, dato che ogni soluzione delle equazioni di Hamilton deve rimanere confinata all'interno dell'insieme compatto Σ_h (dove h è il valore che assume l'integrale primo H all'istante iniziale), ogni soluzione massimale sarà completa in virtù della proposizione 3.7. In questo caso anche (R3) è soddisfatta.
Se infine accade che il volume della regione

$$\mathcal{R} := \{\mathbf{x} \in \mathbb{F} \mid h_1 < H(\mathbf{x}) < h_2\}$$

ha misura finita – è a tal fine sufficiente che la chiusura di $\mathcal{R}$ sia compatta – allora sono soddisfatte le ipotesi del teorema di Poincaré, visto che $\phi_t(\mathcal{R}) \subset \mathcal{R}$ nelle nostre ipotesi: se $p \in \mathcal{R}$, $H(p) = h_p \in (h_1, h_2)$ e, dato che H è integrale primo $H(\phi_t(\mathcal{R})) = h_p \in (h_1, h_2)$ per cui $\phi_t(p) \in \mathcal{R}$ per

ogni $t \in \mathbb{R}$.
2. La situazione descritta nel precedente esempio si concretizza prendendo un sistema fisico di N particelle con hamiltoniana

$$\mathscr{H} = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \mathscr{U}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$$

dove l'energia potenziale $\mathscr{U}$ di classe C^1 è limitata dal basso – cioé per ogni $(\mathbf{x}_1, \ldots, \mathbf{x}_N) \in \mathbb{R}^{3N}$ vale $\mathscr{U}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \geq C$ per qualche costante $C \in \mathbb{R}$ – e tale che gli insiemi dei punti $(\mathbf{x}_1, \ldots, \mathbf{x}_N) \in \mathbb{R}^{3N}$ con $\mathscr{U}((\mathbf{x}_1, \ldots, \mathbf{x}_N)) \leq h$ siano compatti $K_u \subset \mathbb{R}^{3N}$ per ogni h fissato in un intervallo di valori. (Quest'ultima condizione è per esempio assicurata se $\mathscr{U}$ è una somma di energie potenziali per la singola particelle descriventi buche di potenziale). In questo caso i valori assunti da $\sqrt{\sum_{i=1}^{N} \mathbf{p}_i^2}$ non possono superare $N\sqrt{2M(h-C)}$, dove M è la più grande delle masse m_i, e pertanto anche il vettore $(\mathbf{p}_1, \ldots, \mathbf{p}_N)$ è contenuto nella palla chiusa $B_h \subset \mathbb{R}^{3N}$ centrata nell'origine e di raggio $N\sqrt{2M(h-C)}$. In definitiva la superficie chiusa $\Sigma_h \subset \mathbb{R}^{6N}$ individuata da

$$\Sigma_h := \{(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N) \in \mathbb{R}^{6N} \mid \mathscr{H}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N) = h\}$$

è confinata nel compatto $K_h \times B_h \subset \mathbb{R}^{6N}$ ed è pertanto compatta.
Se infine, al variare di h in (h_1, h_2) c'è un compatto K che include tutti i compatti K_h, l'insieme

$$\mathcal{R} := \{(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N) \in \mathbb{R}^{6N} \mid \mathscr{H}(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N) \in (h_1, h_2)\}$$

ha volume finito essendo incluso nel compatto $K \times B_{h_2}$ ed è invariante sotto il flusso dinamico per costruzione.
L'esempio più semplice della situazione descritta è una particella sottoposta alla forza armonica.
3. Una particella sottoposta al potenziale coulombiano attrattivo invece può non soddisfare le ipotesi, in quanto i valori assumibili dall'impulso possono essere arbitrariamente alti a spese dell'energia potenziale che può essere arbitrariamente piccola nei negativi, pur rimanendo l'energia totale costante. Le superfici ad energia costante nello spazio delle fasi possono non essere limitate e dunque non compatte. Tuttavia se si considera anche l'integrale primo del momento angolare, si vede che è possibile definire, per valori arbitrari dell'energia e valori *non nulli* del momento angolare[2], insiemi compatti contenenti le soluzioni del moto, la cui unione forma un insieme $\mathcal{R}$ invariante con misura canonica finita.

Osservazioni 10.4. Un caso fisicamente interessante nel quale si applica il teorema del ritorno è quello di un gas di particelle confinato in una buca di potenziale limitata dal basso (una

[2]Nella trattazione in coordinate radiali, quando il momento angolare è non nullo, la teoria è descritta da un'energia poteziale effettiva che tiene conto del cosiddetto "potenziale centrifugo" repulsivo in addizione all'energia potenziale coulombiana. A causa di questo termine l'energia potenziale effettiva è limitata dal basso quando il momento angolare è non nullo.

scatola contenente il gas può essere pensata come una buca di potenziale limitatat dal basso) che non interagiscono tra di loro (gas ideale). In questo caso, partendo da ogni stato iniziale – in particolare quello in cui tutte le particelle si trovano in metà scatola – ed aspettando un tempo sufficientemente lungo, il sistema tornerà vicino a tale stato. È chiaro che questa situazione è in contrasto con il secondo principio della termodinamica. Tuttavia si riescono a stimare i tempi "di ritorno" per sistemi fisici come quello descritto e si scopre che hanno ordini di grandezza comparabili con quelli dell'età del sistema solare [Arnold92]. In base a tale risultato si può ritenere la termodinamica dell'equilibrio una descrizione approssimata valida su scale temporali relativamente brevi. È importante, infine mezionare il fatto che esiste una versione quantistica del teorema del ritorno, valida per sistemi hamiltoniani con "spettro discreto" dell'energia (P. Bocchieri e A. Loinger, Phys. Rev. 107, 337 - 338 (1957)). Per un gas ideale quantistico confinato in una scatola lo spettro dell'energia è realmente discreto; pertanto continuano a valere, anche in ambito quantistico, le implicazioni del teorema di ricorrenza nei confronti della validità della termodinamica dell'equilibrio.

10.3 Simmetrie e leggi di conservazione in meccanica di Hamilton.

Introdurremo ora un utile strumento teorico e pratico dato dalle cosiddette *parentesi di Poisson*. Si tratta di funzionali bilineari che agiscono su funzioni definite nello spaziotempo delle fasi. L'importanza di tale strumento matematico è , da un lato che esso consente di studuare agevolmente il legame tra simmetrie e quantità conservate in formulazione di Hamilton, dall'altro che diversi risultati della meccanica hamiltoniana se scritti in termini di parentesi di Poisson hanno un corrispondente quantistico quando si sostituiscono alle parentesi di Poisson i commutatori quantistici (si tratta del cosiddetto *principio di corrispondenza di Dirac*).
In questa sezione useremo in particolare il formalismo delle parentesi di Poisson per riformulare, nel linguaggio hamiltoniano, il legame tra simmetrie e leggi di conservazione già introdotto in formulazione di Lagrange attraverso il teorema di Nöther.

10.3.1 Parentesi di Poisson.

Consideriamo una funzione (campo scalare) $f : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ di classe C^1. Restringiamoci ora a lavorare in un sistema locale di coordinate canoniche $(t, q^1, \ldots, q^n, p_1, \ldots, p_n) \equiv (t, \mathbf{x})$. Alla funzione f possiamo associare il campo vettoriale X_f su $F(\mathbb{V}^{n+1})$ definito localmente da

$$X_f := \sum_{i,j=1}^{2n} S^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j} = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial q^k} \frac{\partial}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial}{\partial q^k} \right) . \tag{10.21}$$

Questo campo definito localmente ha in realtà una natura globale, nel senso che l'espressione sopra scritta vale in ogni sistema locale di coordinate canoniche e pertanto, dato che tali sistemi

di coordinate locali formano un atlante, individua univocamente un campo vettoriale su $F(\mathbb{V}^{n+1})$ che indicheremo ancora con X_f. X_f è pertanto definito come campo vettoriale su $F(\mathbb{V}^{n+1})$ che si riduce all'espressione (10.21) in ogni sistema locale di coordinate canoniche. Per dimostrare che l'espressione (10.21) è invariante cambiando coordinate locali canoniche, supponiamo che $(t', \mathbf{x}')$ siano le coordinate di un secondo sistema di coordinate canoniche con dominio sullo spazio delle fasi che interseca quello delle precedenti. In tale intersezione possiamo scrivere che:

$$\sum_{i,j=1}^{2n} S^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j} = \sum_{i,j,k,r=1}^{2n} S^{ij} \frac{\partial x'^k}{\partial x^i} \frac{\partial x'^r}{\partial x^j} \frac{\partial f}{\partial x'^k} \frac{\partial}{\partial x'^r} = \sum_{k,r=1}^{2n} (JSJ^t)^{kr} \frac{\partial f}{\partial x'^k} \frac{\partial}{\partial x'^r} = \sum_{k,r=1}^{2n} S^{kr} \frac{\partial f}{\partial x'^k} \frac{\partial}{\partial x'^r},$$

dove abbiamo usato la definizione di trasformazione canonica e quindi il fatto che la matrice jacobiana J (e la sua trasposta) della trasformmazione di coordinate non temporali è un elemento del gruppo $Sp(n, \mathbb{R})$ e pertanto $JSJ^t = S$.
Consideriamo ora due funzioni $f, g : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ di classe C^1. Ha senso costruire un terzo campo scalare definito localmente, in coordinate canoniche da:

$$\{f, g\} := X_f(g) = -X_g(f) := \sum_{i,j=1}^{2n} S^{ij} \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j} = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial q^k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q^k} \right) . \tag{10.22}$$

Tale campo scalare sarà in realtà ben definito *su tutto lo spaziotempo delle fasi* per la stessa ragione per la quale il campo X_f è ben definito su tutto lo spaziotempo delle fasi.

Definizione 10.2. Si considerino $f, g : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ di classe C^1. Il **campo vettoriale associato** a f è il campo vettoriale su $F(\mathbb{V}^{n+1})$, X_f, definito in coordinate locali canoniche da (10.21). La **parentesi di Poisson di** f **e** g è il campo scalare $\{f, g\}$ su $F(\mathbb{V}^{n+1})$ definito in coordinate locali canoniche da (10.22). $\diamondsuit$

Osservazioni 10.5.
(1) A titolo di esempio si verifica immediatamente che se (t, q, p) sono coordinate canoniche allora valgono le condizioni:

$$\{q^i, q^j\} = \{p_i, p_j\} = 0 \quad \text{e} \quad \{q^i, p_j\} = \delta^i_j . \quad \text{per ogni } i, j = 1, \ldots, n .$$

Classi di funzioni f^i, g_j, con $i, j = 1, \ldots, n$, definite sullo spaziotempo delle fasi, o solo localmente in qualche sistema di coordinate canoniche, che soddisfano tutte le identità scritte sopra sono dette essere **canonicamente coniugate**. Questa condizione (che rivestirà un ruolo importante in meccanica quantistica dato che è legata al *principio di indeterminazione di Heisenberg*) è in realtà una condizione *sufficiente* per assicurare che le funzioni definiscano un sistema di coordinate canoniche come proviamo sotto.
(2) È importante notare che $X_f = X_{f+h}$ se e solo se $h : \mathbb{R} \to \mathbb{R}$ è una *qualsiasi* funzione del solo tempo e pertanto il campo vettoriale X_f *non* porta le stesse informazioni del campo scalare f.

Proposizione 10.2. *Si considerino due sistemi di coordinate locali su $F(\mathbb{V}^{n+1})$, (t,q,p) e (t',q',p') definite su un dominio comune in cui $t' = t + c$ è il tempo assoluto e le rimanenti coordinate definiscono sistemi di coordinate (per ogni t e t' fissati, sulle fibre di $F(\mathbb{V}^{n+1})$). Si supponga infine che (t,q,p) sia un sistema locale di coordinate canoniche. I seguenti fatti sono equivalenti.*
(a) *Le coordinate locali (t',q',p') sono canoniche;*
(b) *per ogni scelta di $f,g : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ di classe C^1 risulta:*

$$\{f,g\} = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial q'^k} \frac{\partial g}{\partial p'_k} - \frac{\partial f}{\partial p'_k} \frac{\partial g}{\partial q'^k} \right) ;$$

(c) *Valgono le condizioni:*

$$\{q'^i, q'^j\} = \{p'_i, p'_j\} = 0 \quad e \quad \{q'^i, p'_j\} = \delta^i_j \,, \quad \textit{per ogni } i,j = 1,\ldots,n \,. \tag{10.23}$$

Dimostrazione. Il fatto che (a) impliche (b) è immediato dal fatto che il campo X_f ha la stessa struttura in tutti i sistemi di coordinate canoniche come dimostrato precedentemente. Il fatto che (b) impliche (c) si prova per computo diretto come già osservato sotto la definizione 10.2 esprimendo le parentesi di Poisson direttamente nelle coordinate canoniche primate. Proviamo per concludere che (c) implica (a). Le condizioni (10.23) si possono riscrivere tutte insieme:

$$\sum_{k,h=1}^{2n} S^{kh} \frac{\partial x'^i}{\partial x^k} \frac{\partial x'^j}{\partial x^h} = S^{ij} \,.$$

In forma compatta:

$$J^t S J = S \,.$$

Questo significa che la trasformazione di coordinate è canonica e dunque le coordinate (t',q',p') sono canoniche. □

La seguente proposizione illustra alcune proprietà più avanzate delle parentesi di Poisson.

Proposizione 10.3. *Le parentesi di Poisson godono delle seguenti proprietà .*
(1) *$f,g \mapsto \{f,g\}$ è un'applicazione* bilineare *e* antisimmetrica *che associa campi scalari di classe $C^k(F(\mathbb{V}^{n+1}))$ ($k > 0$) ad un campo scalare di classe $C^{k-1}(F(\mathbb{V}^{n+1}))$.*
(2) *Vale la* **regola di Leibniz***:*

$$\{f \cdot g, h\} = f \cdot \{g,h\} + g \cdot \{f,h\}$$

dove $\cdot$ indica il prodotto punto per punto ed f,g,h sono campi scalari di classe C^1 su $F(\mathbb{V}^{n+1})$.
(3) *Vale la* **regola di Jacobi***, se f,g,h sono campi scalari di classe C^2 su $F(\mathbb{V}^{n+1})$:*

$$\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\} = 0 \,.$$

(4) *Se* $[\cdot,\cdot]$ *indica la parentesi di Lie di campi vettoriali su* $F(\mathbb{V}^{n+1})$ *e* f, g *sono campi scalari di classe* C^2 *su* $F(\mathbb{V}^{n+1})$*:*

$$[X_f, X_g] = X_{h+\{f,g\}}$$

dove $h : \mathbb{R} \to \mathbb{R}$ *è una qualsiasi funzione del solo tempo.*
$\diamondsuit$

Dimostrazione. Le proprietà (1) e (2) sono di dimostrazione ovvia. Dimostriamo la proprietà (3). Vale:

$$\{\{f,g\},h\} = \sum_{ijkr} S^{kr} S^{ij} \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial g}{\partial x^j} \frac{\partial h}{\partial x^r} + \sum_{ijkr} S^{kr} S^{ij} \frac{\partial f}{\partial x^i} \frac{\partial^2 g}{\partial x^k \partial x^j} \frac{\partial h}{\partial x^r} \,.$$

Gli altri due addendi a primo membro di $\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\} = 0$ sono ottenuti dall'espressione precedente permutando in essa f, g, h. È sufficiente mostrare che, nella somma $\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\}$, i termini che contendono le derivate seconde di f si annullano reciprocamente, la stessa cosa accadrà per simmetria per i termini che contengono le derivate seconde in g ed in h. Il primo termine con le derivate seconde di f è quello che appare sopra nello sviluppo di $\{\{f,g\},h\}$. Ne compare solo un altro nello sviluppo di $\{\{h,f\},g\}$ che si ottiene dal secondo addendo nello sviluppo di $\{\{f,g\},h\}$ con la permutazione $f, g, h \to h, f, g$. In definita bisogna verificare che:

$$\sum_{ijkr} S^{kr} S^{ij} \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial g}{\partial x^j} \frac{\partial h}{\partial x^r} + \sum_{ijkr} S^{kr} S^{ij} \frac{\partial h}{\partial x^i} \frac{\partial^2 f}{\partial x^k \partial x^j} \frac{\partial g}{\partial x^r} = 0 \,. \tag{10.24}$$

Il secondo addendo a primo membro di (10.24), cambiando i nomi agli indici ($k \to j$ e quindi $i \to r$), può essere riscritto come:

$$\sum_{ijkr} S^{kj} S^{ri} \frac{\partial h}{\partial x^r} \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial g}{\partial x^j} = \sum_{ijkr} S^{kj} S^{ri} \frac{\partial h}{\partial x^r} \frac{\partial^2 f}{\partial x^i \partial x^k} \frac{\partial g}{\partial x^j} \,,$$

dove, nell'ultimo passaggio abbiamo usato il teorema di Schwarz (per questo abbiamo richiesto le funzioni di classe C^2 e non solo differenziabili fino al secondo ordine). Scambiando ancora i nomi di i e k arriviamo a dire che il secondo addendo a primo membro di (10.24):

$$\sum_{ijkr} S^{ij} S^{rk} \frac{\partial h}{\partial x^r} \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial g}{\partial x^j} = -\sum_{ijkr} S^{ij} S^{kr} \frac{\partial h}{\partial x^r} \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial g}{\partial x^j} \,,$$

dove, nell'ultimo passaggio abbiamo uasato l'antisimmetria della matrice S. Abbiamo ottenuto che:

$$\sum_{ijkr} S^{kr} S^{ij} \frac{\partial h}{\partial x^i} \frac{\partial^2 f}{\partial x^k \partial x^j} \frac{\partial g}{\partial x^r} = -\sum_{ijkr} S^{ij} S^{kr} \frac{\partial h}{\partial x^r} \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial g}{\partial x^j} \,.$$

Sostituendo a primo membro in (10.24) si ha immediatamente che (10.24) è verificata.
Per concludere dimostriamo la proprietà (4). Se h è un campo scalare su $F(\mathbb{V}^{n+1})$ di classe C^2 almeno, si ha:

$$[X_f, X_g]h = X_f X_g h - X_g X_f h = X_f\{g,h\} - X_g\{f,h\} = \{f,\{g,h\}\} - \{g,\{f,h\}\}$$
$$= -\{\{g,h\},f\} - \{\{h,f\},g\} = \{\{f,g\},h\} = X_{\{f,g\}}h\,.$$

L'arbitrarietà di h implica che:

$$[X_f, X_g] = X_{\{f,g\}}\,.$$

L'osservazione (2) in osservazioni 10.5 conclude la dimostrazione □

10.3.2 Gruppi locali ad un parametro di trasformazioni canoniche attive.

Mostriamo ora come sia possibile associare ad ogni campo scalare $f : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ un gruppo ad un parametro di trasformazioni canoniche *attive*. Attivo significa che le trasformazioni non si devono intendere come trasformazioni di coordinate, ma come trasformazioni che spostano il punto rappresentativo del sistema fisico. Consideriamo a tal fine $f \in C^2(F(\mathbb{V}^{n+1}))$ ed il campo associato X_f. Vogliamo trovare le curve integrali di tale campo e con esse costruire il gruppo locale ad un parametro di diffeomorfismi associato a X_f. Si tratta quaindi, in ogni sistema locale di coordinate canoniche $(t, \mathbf{x})$, di risolvere l'equazione differenziale del prim'ordine nell'incognita data dalla curva $(t, \mathbf{x}) = (t(u), \mathbf{x}(u))$:

$$X_f(t(u), \mathbf{x}(u)) = \frac{dx^k}{du}\frac{\partial}{\partial x^k} + \frac{dt}{du}\frac{\partial}{\partial t}\,,$$

cioé , esplicitamente:

$$\sum_{k=1}^{n} S^{ik}\left.\frac{\partial f}{\partial x^i}\right|_{(t(u),\mathbf{x}(u))}\frac{\partial}{\partial x^k} = \frac{dx^k}{du}\frac{\partial}{\partial x^k} + \frac{dt}{du}\frac{\partial}{\partial t}\,.$$

Notiamo prima di tutto che la funzione $t = t(u)$ deve essere costante. Pertanto l'equazione differenziale iniziale, *istante per istante*, si riduce al sistema:

$$S^t\nabla_{(t,\mathbf{x}(u))}f = \frac{d\mathbf{x}}{du}\,. \tag{10.25}$$

Le curve soluzioni $(t, \mathbf{x}) = (t(u), \mathbf{x}(u))$ che si ottengono nello spazio delle fasi sono tali da rimanere ciascuna nella stessa fibra individuata all'istante iniziale. In altre parole, il gruppo locale di diffomorfismi ad un parametro $\Phi^{(X_f)}$ associato al campo X_f *preserva le fibre* dello spaziotempo delle fasi. Fissiamo un istante t una volta per tutte e consideriamo il gruppo locale ad un parametro $\phi^{(t,X_f)}$ associato su $\mathbb{F}_t$ all'equazione differenziale autonoma (si noti che t *non* è il parametro che descrive le curve!). Per definizione

$$u \mapsto \phi_u^{(t,X_f)}(\mathbf{x}) =: \mathbf{x}'$$

è l'unica soluzione locale dell'equazione (10.25) che esce dal punto $\mathbf{x}$ per $u = 0$. Vogliamo mostrare che per ogni valore di u la trasformazione $\mathbf{x} \mapsto \mathbf{x}'$ definita localmente (in dipendenza dal valore assunto da u), *è canonica, nel senso che la sua matrice jacobiana appartiene a* $Sp(n, \mathbb{R})$. Definiamo, per t fissato, la matrice Jacobiana J, dipendente da $\mathbf{x}$ e da u, di coefficienti:

$$J^i{}_k := \frac{\partial x'^i}{\partial x^k} .$$

Consideriamo poi l'unica soluzione $\mathbf{x}' = \mathbf{x}'(u)$ di (10.25) che congiunge $\mathbf{x}$, preso a $u = 0$ a $\mathbf{x}'_1$ preso a $u = u_1$. In questo modo $\mathbf{x}' = \phi_{u_1}^{(t,X_f)}(\mathbf{x})$. Vogliamo calcolare, lungo questa curva la derivata in u di

$$u \mapsto (J^t S J)_{\mathbf{x}'(u)} ,$$

e mostrare che è nulla. Questo fatto implica che $J^t S J$ è costante lungo la curva e quindi la trasformazione $\mathbf{x}' = \phi_{u_1}^{(t,X_f)}(\mathbf{x})$ ha matrice jacobiana che soddisfa $J^t S J = S$ dato che ciò accade per $u = 0$. Notiamo che

$$\frac{d}{du}\left(J^t S J\right) = \left(\frac{dJ^t}{du} S J\right) + J^t S \frac{dJ}{du} . \tag{10.26}$$

Tenendo conto che

$$\frac{d}{du}(J_{\mathbf{x}'(u)})^i{}_k = \frac{d}{du}\left.\frac{\partial x'^i}{\partial x^k}\right|_{\mathbf{x}'(u)} = \frac{\partial^2 x'^i}{\partial u \partial x^k} = \frac{\partial}{\partial x^k}\frac{dx'^i}{du} = \frac{\partial X_f^i}{\partial x^k} ,$$

esplicitando in componenti, il secondo memebro di (10.26) è :

$$\sum_{i,k=1}^n \frac{\partial X_f^i}{\partial x^j} S_{ik} \frac{\partial x'^k}{\partial x^l} + \sum_{i,k=1}^n \frac{\partial x'^i}{\partial x^j} S_{ik} \frac{\partial X_f^k}{\partial x^l} = \sum_{i,k,p=1}^n S_{ip} \frac{\partial^2 f}{\partial x^j \partial x'^p} S_{ik} \frac{\partial x'^k}{\partial x^l} + \sum_{i,k,s=1}^n \frac{\partial x'^i}{\partial x^j} S_{ik} S_{ks} \frac{\partial^2 f}{\partial x^l \partial x'^s}$$

$$= \sum_{k,p=1}^n \delta_{kp} \frac{\partial^2 f}{\partial x^j \partial x'^p} \frac{\partial x'^k}{\partial x^l} - \sum_{i,s=1}^n \frac{\partial x'^i}{\partial x^j} \delta_{is} \frac{\partial^2 f}{\partial x^l \partial x'^s} = \sum_{k=1}^n \frac{\partial^2 f}{\partial x^j \partial x'^k} \frac{\partial x'^k}{\partial x^l} - \sum_{s=1}^n \frac{\partial x'^s}{\partial x^j} \frac{\partial^2 f}{\partial x^l \partial x'^s}$$

$$= \frac{\partial^2 f}{\partial x^j \partial x^l} - \frac{\partial^2 f}{\partial x^l \partial x^j} = 0 .$$

Dato che per $u = 0$ si ha $J = I$ e pertanto $J^t S J = S$, concludiamo che, per ogni u deve valere $J^t S J = S$ come già detto all'inizio. In particolare ciò accade per $u = u_1$. Concludiamo che la trasformazione

$$(t, \mathbf{x}) \mapsto (t', \mathbf{x}') = \left(t, \phi_u^{(t,X_f)}(\mathbf{x})\right) ,$$

è canonica ovunque è definita. Abbiamo provato la seguente proposizione.

Proposizione 10.4. *Il gruppo locale ad un parametro di diffeomorfismi* $\Phi^{(X_f)}$ *generato dal campo vettoriale* X_f *associato ad ogni campo scalare* $f \in C^2(F(\mathbb{V}^{n+1}))$ *è costituito da*

trasformazioni canoniche attive. Più precisamente, se l'aperto $A \subset F(\mathbb{V}^{n+1})$ è dotato di coordinate canoniche $(t, \mathbf{x})$ e $\Phi_u^{(X_f)}(A) \subset A$ è ben definita per un certo u, allora l'applicazione $(t, \mathbf{x}) \mapsto (t', \mathbf{x}')$ che corrisponde in coordinate all'azione di $\Phi_u^{(X_f)}$ è tale che $t' = t$ e la matrice jacobiana delle coordinate non temporali appartiene a $Sp(n, \mathbb{R})$. $\diamondsuit$

Se interpretiamo *passivamente* la trasformazione canonica indotta da f possiamo chiederci quale sia l'hamiltoniana $\mathscr{H}'(t', \mathbf{x}')$ associata alle coordinate $(t', \mathbf{x}')$. Si osservi che tuttavia, dato che la trasformazione di coordinate è attiva, la nuova funzione hamiltoniana $\mathscr{H}'$ si deve più correttamente pensare come la funzione hamiltoniana di un *nuovo* sistema fisico, quello che si è ottenuto dal precedente in seguito alla trasformazione attiva generata dalla funzione f. Il punto importante è che questo nuovo sistema ammette ancora una descrizione di Hamilton. Si osservi infine che non è in generale vero che l'hamiltoniana $\mathscr{H}'$ espressa nelle variabili $(t', \mathbf{x}')$ abbia la stessa forma dell'hamiltoniana $\mathscr{H}$ espressa nelle variabili $(t, \mathbf{x})$. In questo senso il nuovo sistema fisico sarà *differente* da quello iniziale.

Quello che faremo ora è determinare $\mathscr{H}'$ a meno di infinitesimi in u del secondo ordine. La sua forma sarà utile nello stabilire le connessioni tra simmetrie di un sistema di Hamilton e grandezze conservate.

Proposizione 10.5. *Se, nelle stesse ipotesi della proposizione precedente, si intepreta la trasformazione $\mathbf{x} \mapsto \mathbf{x}'$ in senso passivo ed è assegnata una funzione di Hamilton $\mathscr{H} = \mathscr{H}(t, \mathbf{x})$ di classe C^2, la funzione hamiltoniana $\mathscr{H} = \mathscr{H}(t', \mathbf{x}')$ da attribuire alle nuove coordinate può sempre essere scelta in modo tale che:*

$$\mathscr{H}'(t', \mathbf{x}'(t, \mathbf{x})) = \mathscr{H}(t, \mathbf{x}) - u\frac{\partial f}{\partial t}(t, \mathbf{x}) + uO_{(t,\mathbf{x})}(u) \,, \tag{10.27}$$

dove $O_{(t,\mathbf{x})}(u) \to 0$ per $u \to 0$. $\diamondsuit$

Dimostrazione. Porremo $t = t'$ ovunque nel seguito dato che la trasformazione canonica considerata non altera la coordinata temporale. Dal Teorema (10.1) sappiamo che vale:

$$\mathscr{H}'(t, \mathbf{x}') = \mathscr{H}(t, \mathbf{x}(t, \mathbf{x}')) + \mathscr{K}(t, \mathbf{x}') \,,$$

dove la funzione $\mathscr{K}$ è determinata, a meno di funzioni della sola variabile t additive, dalla richiesta

$$\frac{\partial \mathbf{x}'}{\partial t} = S\nabla_{\mathbf{x}'}\mathscr{K}(t, \mathbf{x}') \,,$$

cioé

$$\nabla_{\mathbf{x}'}\mathscr{K}(t, \mathbf{x}') = -S\frac{\partial \mathbf{x}'}{\partial t} \,,$$

Sviluppiamo con Taylor nella variabile u la funzione $\partial \mathbf{x}'/\partial t$, dove $\mathbf{x}' = \mathbf{x}'(u, t, \mathbf{x})$. Si osservi che per $u = 0$, vale $\mathbf{x}'(0, t, \mathbf{x}) = \mathbf{x}$ e pertanto la derivata parziale rispetto a t si annulla. Avremo

allora:

$$\frac{\partial \mathbf{x}'}{\partial t} = u \left. \frac{\partial}{\partial u} \frac{\partial \mathbf{x}'}{\partial t} \right|_{u=0} + uO_{(t,\mathbf{x})}(u) \ ,$$

dove $O_{(t,\mathbf{x})}(u) \to 0$ per $u \to 0$. In altre parole, scambiando l'ordine delle derivate ed usando la definizione di X_f:

$$\frac{\partial \mathbf{x}'}{\partial t} = u \frac{\partial}{\partial t} S^t \nabla_{\mathbf{x}} f + uO_{(t,\mathbf{x})}(u) \ .$$

Applicando $-S$ ad entrambi i membri, ricordando che $SS^t = I$ e notando che $SO_{(t,\mathbf{x})}(u)$ è ancora infinitesimo per $u \to 0$ (nel seguito rinominiamo pertanto $-SO$ semplicemente O), si trova che deve essere:

$$\nabla_{\mathbf{x}'} \mathscr{K}(t, \mathbf{x}') = -u \frac{\partial}{\partial t} \nabla_{\mathbf{x}} f(t, \mathbf{x}) + uO_{(t,\mathbf{x})}(u) \ .$$

Nel secondo membro possiamo sostituire $\nabla_{\mathbf{x}} f(t, \mathbf{x})$ con $\nabla_{\mathbf{x}'} f(t, \mathbf{x}')$ commettendo un errore di ordine superiore al primo in u come è facile verificare[3]. In definitiva abbiamo trovato che:

$$\nabla_{\mathbf{x}'} \mathscr{K}(t, \mathbf{x}') = -\nabla_{\mathbf{x}'} u \frac{\partial f(t, \mathbf{x}')}{\partial t} + uO_{(t,\mathbf{x})}(u) \ .$$

A meno di infinitesimi del secondo ordine in u siamo liberi di scegliere:

$$\mathscr{K}(t, \mathbf{x}) = -u \frac{\partial f(t, \mathbf{x})}{\partial t} \ .$$

□

Definizione 10.3. Nelle ipotesi della proposizione precedente, l'hamiltoniana $\mathscr{H}$ si dirà essere **invariante in forma al prim'ordine** sotto il gruppo locale ad un parametro di trasformazioni canoniche generate dalla funzione f quando:

$$\mathscr{H}'(t', \mathbf{x}') = \mathscr{H}(t', \mathbf{x}') + uO_{(t',\mathbf{x}')}(u) \ , \tag{10.28}$$

dove $O_{(t',\mathbf{x}')}(u) \to 0$ se $u \to 0$. ♢.

Osservazioni 10.6. Il significato della definizione dovrebbe essere fisicamente ovvio: se l'hamiltoniana è invariante in forma il sistema fisico associato all'hamiltoniana $\mathscr{H}'$ che si ottiene sottoponendo il sistema fisico inziale, associato all'hamiltoniana $\mathscr{H}$, ha la stessa forma (nelle nuove coordinate) della precedente hamiltoniana. La trasformazione canonica attiva che abbiamo usato è in questo senso una **simmetria hamiltoniana** del sistema fisico. Vedremo nella prossima sezione che in questa situazione abbiamo nuovamente uno strettissimo legame tra presenza di simmetria ed esistenza di quantità fisiche conservate (integrali primi).

[3] Sviluppando con Taylor per $u = 0$, abbiamo che $u\nabla_{\mathbf{x}} f(t, \mathbf{x}') = u\left(\nabla_{\mathbf{x}} f(t, \mathbf{x}) + uH(t, \mathbf{x})\right) = u\nabla_{\mathbf{x}} f(t, \mathbf{x}) + uO(u)$. Accade la stessa cosa rimpiazzando $\nabla_{\mathbf{x}}$ con $\nabla_{\mathbf{x}'}$.

10.3.3 Simmetrie e leggi di conservazione.

Consideriamo una funzione $f : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ quando sullo spazio delle fasi è stato assegnato un campo vettoriale dinamico Z individuato da funzioni Hamiltoniane $\mathscr{H} = \mathscr{H}(t, q, p)$ in ogni sistema di coordinate dell'atlante canonico. La condizione che f sia un integrale primo per il sistema hamiltoniano $\mathfrak{S}$ la cui dinamica è descritta da Z, in coordinate canoniche si scrive:

$$\frac{d}{dt} f(t, \mathbf{x}(t)) = 0 \,, \tag{10.29}$$

dove $\mathbf{x} = \mathbf{x}(t)$ è una qualsiasi soluzione delle equazioni di Hamilton:

$$\frac{d\mathbf{x}}{dt} = S \nabla \mathscr{H}(t\mathbf{x}(t)) \,.$$

Sviluppando il secondo membro di (10.29), la stessa condizione si può riscrivere:

$$\left.\frac{\partial f}{\partial t}\right|_{(t,\mathbf{x}(t))} + \left.\frac{d\mathbf{x}}{dt} \cdot \nabla f\right|_{(t,\mathbf{x}(t))} = 0 \,,$$

cioé, usando le equazioni di Hamilton:

$$\left.\frac{\partial f}{\partial t}\right|_{(t,\mathbf{x}(t))} + \left.\sum_{i,j=1}^{2n} S^{ij} \frac{\partial \mathscr{H}}{\partial x^j} \frac{\partial f}{\partial x^i}\right|_{(t,\mathbf{x}(t))} = 0 \,.$$

Abbiamo trovato che f è un integrale primo se e solo se, su ogni soluzione $\mathbf{x} = \mathbf{x}(t)$ delle equazioni di Hamilton vale:

$$\left.\left[\frac{\partial f}{\partial t} + \{f, \mathscr{H}\}\right]\right|_{(t,\mathbf{x}(t))} = 0 \,.$$

D'altra parte, assumendo che l'hamiltoniana sia una funzione di classe C^2 e pertanto valendo il teorema di esistenza (ed unicità) delle soluzioni delle equazioni di Hamilton, possiamo concludere che per ogni evento $(t, \mathbf{x})$ nello spazio delle fasi, c'è sicuramente una soluzione che ha quell'evento come condizione iniziale. Questo significa che affinché f sia integrale primo è necessario e sufficiente che *ovunque* sullo spaziotempo delle fasi (in ogni sistema locale di coordinate canoniche) valga l'identità :

$$\frac{\partial f}{\partial t} + \{f, \mathscr{H}\} = 0 \,.$$

Osservazioni 10.7. In tale identità non compaiono più le soluzioni delle equazioni del moto e pertanto, in viortù di tale identità , non è più necessario risolvere il problema del moto per stabilire se una funzione è un integrale primo.

Esprimiamo questo risultato formalmente aggiungendo un ulteriore importante risultato.

Teorema 10.4. *Si consideri un sistema fisico hamiltoniano $\mathfrak{S}$ descritto sullo spazio delle fasi da un campo vettoriale dinamico Z individuato da funzioni Hamiltoniane $\mathscr{H} = \mathscr{H}(t, q, p)$ di classe C^2 in ogni sistema di coordinate dell'atlante canonico. Valgono i seguenti fatti.*
(a) *Una funzione $f : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ di classe C^1 è integrale primo per il sistema hamiltoniano $\mathfrak{S}$ se e solo se vale la condizione, ovunque su $F(\mathbb{V}^{n+1})$:*

$$\frac{\partial f}{\partial t} + \{f, \mathscr{H}\} = 0\,. \tag{10.30}$$

(b) *Se f è di classe C^2 e $g : F(\mathbb{V}^{n+1}) \to \mathbb{R}$ è un ulteriore integrale primo di classe C^2, allora il campo scalare su $F(\mathbb{V}^{n+1})$ dato dalla parentesi di Poisson $\{f, g\}$ è un ulteriore integrale primo.* $\diamondsuit$

Dimostrazione. La prima proposizione è stata provata sopra. Per quanto riguarda la seconda osserviamo che:

$$\frac{\partial\{f,g\}}{\partial t} + \{\{f,g\},\mathscr{H}\} = \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\} - \{\{\mathscr{H}, f\}, g\} - \{\{g, \mathscr{H}\}, f\}\,,$$

dove, nell'ultimo passaggio abbiamo usato l'identità di Jacobi. Usando l'antisimmetria e la linearità della parentesi di Poisson, possiamo ancora scrivere che:

$$\frac{\partial\{f,g\}}{\partial t} + \{\{f,g\},\mathscr{H}\} = \left\{\frac{\partial f}{\partial t} + \{f,\mathscr{H}\}, g\right\} - \left\{f, \frac{\partial g}{\partial t} + \{g,\mathscr{H}\}\right\} = 0\,.$$

Notiamo ora che, essendo f e g integrali primi, deve essere:

$$\frac{\partial g}{\partial t} + \{g, \mathscr{H}\} = 0 \quad \text{e} \quad \frac{\partial f}{\partial t} + \{f, \mathscr{H}\} = 0\,.$$

Per linearità $\{f, g\}$ soddisfa la condizione (10.30) ed è dunque anch'esso un integrale primo. $\square$

Osservazioni 10.8.
(1) Il primo enunciato del teorema implica immediatamente che l'hamiltoniana stessa è un integrale primo se nel sistema di coordinate canoniche usato essa non dipende esplicitamente dal tempo, essendo evidentemente $\{\mathscr{H}, \mathscr{H}\} = 0$ per antisimmetria.
(2) Il numero massimo possibile di integrali primi su $F(\mathbb{V}^{n+1})$ funzionalmente indipendenti è $2n + 1$. Questo segue dal fatto che sistemi di funzioni a valori in $\mathbb{R}$ che siano funzionalmente indipendenti possono essere usati localmente come sistemi di coordinate su $F(\mathbb{V}^{n+1})$ – oppure su sottovarietà – a causa della loro funzionale indipendenza, e del fatto che $F(\mathbb{V}^{n+1})$ ha esattamente $2n + 1$ dimensioni. Questo significa che, calcolando ricorsivamente parentesi di Poisson di integrali primi, ad un certo punto si troveranno gli stessi integrali primi gà ottenuti oppure funzioni di essi.
(3) L'equazione di Liouville per la densità di probabilità ρ nello spaziotempo delle fasi può essere scritta:

$$\frac{\partial \rho}{\partial t} + \{\rho, \mathscr{H}\} = 0\,.$$

Diamo infine il teorema hamiltoniano che, giocando lo stesso ruolo del teorema di Nöther in meccanica lagrangiana, connette la presenza di simmetrie di un sistema hamiltoniano all'esistenza di integrali primi. In questo contesto hamiltoniano abbiamo un ulteriore risultato, assente nella formulazione di Lagrange, l'integrale primo associato ad una simmetria risulta essere proprio la funzione che genera la simmetria nel senso dei gruppi locali ad un parametro di trasformazioni canoniche.

Teorema 10.5. *Sia $f \in C^2(F(\mathbb{V}^{n+1}))$ un campo scalare che genera un gruppo locale ad un parametro di trasformazioni canoniche $\Phi^{(X_f)}$. Sia $\mathscr{H}$ una funzione di Hamilton riferita ad un sistema locale di coordinate canoniche su $F(\mathbb{V}^{n+1})$ il cui dominio è invariante sotto l'azione di $\Phi^{(X_f)}$. I seguenti tre fatti sono equivalenti.*

(1) *f è un integrale primo.*

(2) *Vale ovunque nelle coordinate considerate:*

$$\{f, \mathscr{H}\} + \frac{\partial f}{\partial t} = 0$$

(3) *La funzione hamiltoniana $\mathscr{H}$ è in variante in forma al prim'ordine sotto l'azione del gruppo locale ad un parametro di diffeomorfismi generato da X_f.*
Se valgono le condizioni (a), (b) e (c), allora vale anche

$$[X_f, Z] = 0 \tag{10.31}$$

sulla carta considerata e pertanto $\Phi^{(X_f)}$ trasforma soluzioni delle equazioni di Hamilton associate a $\mathscr{H}$ in analoghe soluzioni. $\diamondsuit$

Dimostrazione. L'equivalenza tra (1) e (2) è già stata provata nel teorema precedente. Mostriamo l'equivalenza di tali condizioni con (3). La condizione di invarianza in forma al prim'ordine si scrive (si ricordi che $t = t'$):

$$\mathscr{H}'(t, \mathbf{x}') = \mathscr{H}(t, \mathbf{x}') + uO'_{(t,\mathbf{x}')}(u) \,. \tag{10.32}$$

Tenendo conto che deve anche valere, per la Proposizione 10.5:

$$\mathscr{H}'(t, \mathbf{x}') = \mathscr{H}(t, \mathbf{x}(t, \mathbf{x}')) - u\frac{\partial f}{\partial t} + uO_{(t',\mathbf{x}')}(u) \,, \tag{10.33}$$

dalla precedente formula si trova che:

$$\mathscr{H}(t, \mathbf{x}') = \mathscr{H}(t, \mathbf{x}) - u\frac{\partial f}{\partial t} + uO''_{(t',\mathbf{x}')}(u) \,. \tag{10.34}$$

Dividendo per u e facendo il limite per $u \to 0$ otteniamo che:

$$\left.\frac{d}{du}\right|_{u=0} \mathscr{H}(t, \mathbf{x}') + \frac{\partial f}{\partial t} = 0 \,.$$

Esplicitamente:

$$\sum_{k=1}^{n} \frac{\partial \mathscr{H}}{\partial x^k} \left. \frac{dx'^k}{du} \right|_{u=0} + \frac{\partial f}{\partial t} = 0 \,,$$

ossia

$$\sum_{k,i=1}^{n} \frac{\partial \mathscr{H}}{\partial x^k} S^{ik} \frac{\partial f}{\partial x^i} + \frac{\partial f}{\partial t} = 0 \,,$$

che può finalmente riscriversi nella forma (2):

$$\{f, \mathscr{H}\} + \frac{\partial f}{\partial t} = 0 \,.$$

Procedendo al contrario abbiamo che questa identità implica (10.34). Quest'ultima, in vitrù di (10.33) implica (10.32) cioé la (1). Abbiamo provato in questo modo che (1),(2) e (3) sono equivalenti.
Per quanto riguarda l'ultima affermazione, si osservi che il vettore dinamico Z può scriversi formalmente:

$$Z = \frac{\partial}{\partial t} - X_{\mathscr{H}} \,,$$

pertanto, usando il fatto che $\left[X_f, \frac{\partial}{\partial t}\right] = X_{-\frac{\partial}{\partial t}}$ e facendo inparticolare uso di (4) della Proposizione 10.3 si trova:

$$[X_f, Z] = \left[X_f, \frac{\partial}{\partial t} - X_{\mathscr{H}}\right] = \left[X_f, \frac{\partial}{\partial t}\right] + [X_f, X_{-\mathscr{H}}] = X_{-\frac{\partial}{\partial t}} + X_{-\{f,\mathscr{H}\}} = -X_{\frac{\partial}{\partial t}+\{f,\mathscr{H}\}} = 0 \,.$$

Questo conclude la dimostrazione. □

Esempi 10.2. Consideriamo una particella di massa m che si muov in un piano individuata da corodinate polari (r, θ) ed è sottoposta ad una forza conservativa con energia potenziale centrale $\mathscr{U} = \mathscr{U}(r)$. L'hamiltoniana della particella sarà allora:

$$\mathscr{H}(r, \theta, p_r, p_\theta) = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \mathscr{U}(r) \,. \tag{10.35}$$

Consideriamo la funzione $f(r, \theta, p_r, p_\theta) := p_\theta$ ed il gruppo ad un parametro di trasformazioni canoniche ad essa associato. Tale gruppo si ottiene integrando l'equazione differenziale, se $\mathbf{x} := (r, \theta, p_r, p_\theta)$:

$$S^t \nabla_{\mathbf{x}} f = \frac{d\mathbf{x}}{du} \,.$$

Il calcolo diretto mostra che tale equazione si riduce semplicemente a:

$$\frac{dr}{du} = 0 \,, \quad \frac{d\theta}{du} = -1 \,, \quad \frac{dp_r}{du} = 0 \,, \quad \frac{dp_\theta}{du} = 0 \,.$$

In definitiva il gruppo $\phi^{X_{p_\theta}}$ agisce come:

$$\phi_u^{(X_{p_\theta})}(r, \theta, p_r, p_\theta) = (r, \theta - u, p_r, p_\theta) =: (r', \theta', p'_r, p'_\theta) \,, \tag{10.36}$$

e quindi, nello spazio delle configurazioni, corrisponde a rotazioni di un angolo $-u$ del punto attorno all'origine. Dato che la trasformazione canonica $\phi_u^{(X_{p_\theta})}$ non dipende dal tempo ed è quindi completamente canonica, l'hamiltoniana $\mathscr{H}'(\mathbf{x}')$ può essere definito come $\mathscr{H}'(\mathbf{x}') = \mathscr{H}(\mathbf{x}(\mathbf{x}'))$ cioé , da (10.36):

$$\mathscr{H}'(r', \theta', p'_r, p'_\theta) = \mathscr{H}(r', \theta' + u, p'_r, p'_\theta) \,.$$

Abbiamo trovato che:

$$\mathscr{H}'(r', \theta', p'_r, p'_\theta) = \frac{p'^2_r}{2m} + \frac{p'^2_\theta}{2mr'^2} + \mathscr{U}(r') \,.$$

Per confronto con (10.35) abbiamo che:

$$\mathscr{H}'(r', \theta', p'_r, p'_\theta) = \mathscr{H}(r', \theta', p'_r, p'_\theta)$$

e quindi l'hamiltoniana è invariante in forma (non solo al prim'ordine) sotto l'azione di $\phi^{X_{p_\theta}}$. Di conseguenza la funzione p_θ si conserva sui moti del sistema. Tale grandezza è il momento angolare del punto materiale considerato. Similmente deve accadere che se $t \mapsto \mathbf{x}(t)$ descrive un moto del punto materiale, anche $t \mapsto \phi_u^{X_{p_\theta}}(\mathbf{x}(t))$ deve descrivere un moto del sistema fisico per ogni fissato valore dell'angolo u.

Appendice A

Alcune nozioni matematiche rilevanti.

Vediamo in questa appendice alcune nozioni di grande importanza in meccanica classica e più in generale in fisica matematica. Si tratta delle nozioni elementari di geometria differenziale e di teoria degli spazi affini ed euclidei.
Ricordiamo prima alcuni fatti di topologia generale elementare. Una coppia $(X, \mathscr{T})$, dove X è un insieme e $\mathscr{T}$ una classe di sottoinsiemi di X, si dice **spazio topologico** se valgono i seguenti fatti: (i) $\emptyset, X \in \mathscr{T}$, (ii) l'unione di elementi (in quantità arbitraria) di $\mathscr{T}$ è ancora un elemento di $\mathscr{T}$, (iii) l'intersezione di una quantità finita di elementi di $\mathscr{T}$ è ancora un elemento di $\mathscr{T}$. In questo caso gli insiemi di $\mathscr{T}$ vengono detti **(insiemi) aperti**, $\mathscr{T}$ viene detta **topologia** su X. Gli **insiemi chiusi** di X sono, per definizione, i complementi rispetto a X degli aperti.
In $\mathbb{R}^n$ c'è una topologia naturale detta **topologia euclidea** o anche **topologia standard**, i cui aperti sono, oltre all'insieme vuoto, gli insiemi dati dalle unioni arbitrarie di palle aperte di centro e raggio arbitrario. Le nozioni topologiche elementari di $\mathbb{R}^n$ studiate nei corsi di analisi sono sottocasi delle nozioni generali date sopra.
Se $A \subset X$ è un insieme arbitrario e $(X, \mathscr{T})$ è uno spazio topologico, è possibile indurre su A una topologia $\mathscr{T}_A$ tramite $\mathscr{T}$. Ponendo infatti: $\mathscr{T}_A := \{B \cap A \mid B \in \mathscr{T}\}$, si verifica subito che $(A, \mathscr{T}_A)$ è ancora uno spazio topologico. $\mathscr{T}_A$ è detta **topologia indotta (da $\mathscr{T}$) su** A.
Se $p \in X$, dove $(X, \mathscr{T})$ è uno spazio topologico, un **intorno aperto** di p è un qualsiasi $A \in \mathscr{T}$ tale che $p \in \mathscr{T}$.
Se $(X, \mathscr{T})$ e $(X', \mathscr{T}')$ sono spazi topologici, una funzione $f : X \to X'$ è detta **continua**, se è tale che $f^{-1}(A) \subset \mathscr{T}$ quando $A \in \mathscr{T}'$ (cioé la controimmagine di un aperto dello spazio di arrivo è un aperto nello spazio di partenza). Se $(X, \mathscr{T})$ e $(X', \mathscr{T}')$ sono spazi topologici e $p \in X$, una funzione $f : X \to X'$ è detta **continua in** p se, per ogni intorno aperto B di $f(p)$, esiste un intorno aperto A di p tale che $f(A) \subset B$. Si verifica facimente che $f : X \to X'$ è continua se e solo se è continua in ogni punto $p \in X$. Se $(X, \mathscr{T})$ e $(X', \mathscr{T}')$ sono spazi topologici, una funzione $f : X \to X'$ è detta **omeomorfismo** quando è continua, biettiva e l'inversa è continua.
Si prova che se i due spazi X, X' usati sopra sono $\mathbb{R}^n$ ed $\mathbb{R}^m$ rispettivamente, oppure sottoinsiemi

di tali spazi, dotati delle rispettive topologie euclidee, la definizioni classiche di funzione continua e continua in un punto sono equivalenti a quelle generali date sopra.
Uno spazio topologico $(X, \mathscr{T})$ è detto **spazio di Hausdorff** o **spazio** T_2 o **spazio separabile di secondo tipo**, se per ogni coppia di punti $p, q \in X$ ci sono due intorni aperti $U \ni p$ e $V \ni q$ con $U \cap V = \emptyset$.
Uno spazio topologico $(X, \mathscr{T})$ è detto essere **a base numerabile** se esiste una classe numerabile $\mathcal{B} \subset \mathscr{T}$ (appunto la "base numerabile") tale che ogni insieme aperto di X si possa ottenere come unione di elementi di $\mathcal{B}$.
$\mathbb{R}^n$ dotato della topologia standard è sia di Hausdorff che a base numerabile.

A.1 Elementi di geometria differenziale.

In tutte le dispense è stata adottata la seguente definizione tecnica che enunciamo qui una volta per tutte.

Definizione A.1. Siano $n, m = 1, 2, \ldots$ e $k = 0, 1, \ldots$ fissati e sia $\Omega \subset \mathbb{R}^n$ un insieme aperto e non vuoto.
(a) Una funzione $f : \Omega \to \mathbb{R}^m$ è detta essere **di classe** C^k, e si scrive in tal caso $f \in C^k(\Omega; \mathbb{R}^n)$, se tutte le derivate parziali (incluse quelle miste) delle componenti di f esistono e sono continue fino all'ordine k incluso. Si pone $C^k(\Omega) := C^k(\Omega; \mathbb{R})$.
(b) $f : \Omega \to \mathbb{R}^m$ è detta di classe C^∞ se è di classe C^k per ogni $k = 0, 1, \ldots$ e si definisce:

$$C^\infty(\Omega; \mathbb{R}^n) := \bigcap_{k=0,1,\ldots} C^k(\Omega; \mathbb{R}^n) \, .$$

Si pone $C^\infty(\Omega) := C^\infty(\Omega; \mathbb{R})$. $\diamondsuit$

Solitamente quando non è menzionata esplicitamente la classe di differenziabilità k di una funzione (oppure di una varietà) si sottointende che $k = \infty$, noi seguiamo questa convenzione in tutte le dispense.

A.1.1 Varietà differenziabili.

Lo strumento matematico più generale e potente atto a descrivere le proprietà generali dello spaziotempo, dello spazio fisico tridimensionale e dello spazio astratto in cui descrivere i sistemi fisici delle teorie classiche, è la nozione di *varietà differenziabile*. Si tratta in essenza di un insieme di oggetti arbitrari, indicati con il nome generico di *punti*, che può essere ricoperto localmente con sistemi di coordinate i quali identificano i punti dell'insieme con n-ple di $\mathbb{R}^n$.

Definizione A.2. Una **varità differenziabile di dimensione** n e **classe** C^k, con $n = 1, 2, 3, \cdots$ e $k = 1, 2, ..., \infty$ fissati, è un insieme M i cui elementi sono detti **punti**, dotato di

alcune strutture geometriche con proprietà che precisiamo di seguito.
(1) L'insieme M deve essere dotato di una **struttura differenziabile di classe** C^k e **dimensione** n, $\mathcal{A} = \{(U_i, \phi_i)\}_{i\in I}$ cioé , una collezione di coppie (U_i, ϕ_i), dette **carte locali** o **sistemi di coordinate locali**, in cui U_i è sottoinsieme di M e ϕ_i è un'applicazione con dominio U_i a valori in $\mathbb{R}^n$ e vale:

(i) $\cup_{i\in I} U_i = M$, ogni ϕ_i è iniettiva e $\phi_i(U_i)$ aperto in $\mathbb{R}^n$;

(ii) le carte locali in $\mathcal{A}$ devono essere *k-compatibili* a due a due. Due applicazioni iniettive $\phi : U \to \mathbb{R}^n$ e $\psi : V \to \mathbb{R}^n$ con $U, V \subset M$ sono dette k-**compatibili** se vale $U \cap V \neq \emptyset$ e le funzioni $\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$ e $\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ sono entrambe di classe C^k, oppure se vale $U \cap V = \emptyset$;

(iii) $\mathcal{A}$ è **massimale** ossia soddisfa: se $U \subset M$ è aperto e $\phi : U \to \mathbb{R}^n$ è compatibile con ogni carta di $\mathcal{A}$, allora $(U, \phi) \in \mathcal{A}$.
(2) Dal punto di vista topologico, si richiede che:

(i) M sia uno spazio topologico di Hausdorff a base numerabile;

(ii) lo spazio topologico M sia, tramite le carte di $\mathcal{A}$, *localmente omeomorfo a* $\mathbb{R}^n$. In altre parole, se $(U, \phi) \in \mathcal{A}$, allora U è aperto e $\phi : U \to \phi(U)$ è un omeomorfismo. $\diamondsuit$

Osservazioni A.1.
(1) Ogni carta locale (U, ϕ) permette di assegnare biunivocamente una n-pla di numeri reali $(x_p^1, \cdots, x_p^n) = \phi(p)$ ad ogni punto p di U. Gli elementi della n-pla sono le **coordinate** di p nella carta (U, ϕ). I punti in U sono quindi in corrispondenza biunivoca con le n-ple di $\phi(U) \subset \mathbb{R}^n$. Una carta locale con dominio dato da tutto M è detta **carta globale** o **sistema di coordinate globale**.
(2) Nell'ipotesi $U \cap V \neq \emptyset$, la k-compatibilità di carte locali (U, ϕ) e (V, ψ) implica che la matrice jacobiana di $\phi \circ \psi^{-1}$, essendo invertibile, abbia determinante ovunque non nullo. Viceversa, se $\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$ è biettiva, di classe C^k, con determinante della matrice jacobiana non nullo su $\psi(U \cap V)$, allora $\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ è anch'essa C^k e quindi le due carte locali sono k-compatibili. La prova di ciò (esercizio A.1.**5**) si basa sul noto [GiustiII]:
Teorema A.1. (**Teorema della funzione inversa**)
Sia $f : D \to \mathbb{R}^n$, con $D \subset \mathbb{R}^n$ aperto non vuoto, una funzione di classe C^k, con $k = 1, 2, \ldots, \infty$ fissato. Se la matrice jacobiana di f, valutata in $p \in D$, ha determinante non nullo allora esistono un intorno aperto $U \subset D$ di p ed un intorno aperto V di $f(p)$ tali che: (i) $f \restriction_U : U \to V$ sia biettiva (ii) la sua inversa $f \restriction_U^{-1} : V \to U$ sia di classe C^k. $\diamondsuit$
(3) Le due richieste sul *tipo* di topologia in (2)(i) (che valgono per la topologia standard di $\mathbb{R}^n$) sono di carattere tecnico e assicurano, rispettivamente, l'unicità delle soluzioni di problemi basati su equazioni differenziali su M (necessaria dal punto di vista fisico quando queste equazioni descrivono l'evoluzione di sistemi fisici) e la fattibilità della teoria dell'intergazione su M. La richiesta in (2)(ii) corrisponde invece al requisito intuitivo che M "sia continuo come" $\mathbb{R}^n$ nell'intorno di ogni punto. Classici controesempi mostrano che la proprietà di Hausdorff di $\mathbb{R}^n$ non è trasportata su M dagli omeomorfismi locali dati e pertanto deve essere imposta separatamente.
(4) Sia M spazio topologico di Hausdorff a base numerabile. Una collezione di carte locali $\mathcal{A}$

su M che soddisfi (i) e (ii) in (1) ma non necessariamente (iii), e che soddisfi (ii) in (2), è detto **atlante** su M **di dimensione** n e **classe** C^k. Si dimostra facilmente che per ogni atlante $\mathcal{A}$ su M esiste un unico atlante massimale che lo include. Si osservi che due atlanti su M tali che ogni carta di uno sia compatibile con ogni carta dell'altro, inducono la stessa struttura differenziabile su M. Quindi per assegnare una struttura differenziabile è sufficiente assegnare un atlante non massimale, uno dei possibili che la individua. L'unica struttura differenziabile associata nel modo detto ad un fissato atlante si dice essere **indotta** dall'atlante.
(5) Si deve comunque notare che possono esistere più strutture differenziabile inequivalenti sulla stesso spazio topologico di Hausdorff a base numerabile[1]. Ciò accade in dimensione ≥ 4. Pertanto, per assegnare una varietà differenziabile N di dimensione $n \geq 4$ non è sufficiente specificare il solo insieme N, anche se è stata specificata la topologia appropriata. Fanno eccezione i casi in cui N è un sottoinsieme di una varietà di dimensione maggiore M assegnata, di cui N è sottovarietà embedded, come specificato sotto.
(6) Si può provare che se $1 \leq k < \infty$, si possono eliminare alcune carte dalla struttura differenziabile (un numero infinito di carte!), in modo tale che l'insieme risultate sia ancora un atlante con $k = \infty$. Si possono considerare varietà *analitiche* (in simboli si scrive C^ω), in cui tutte le funzioni $\phi \circ \psi^{-1}$ e $\psi \circ \phi^{-1}$ sono assunte essere funzioni analitiche reali.

Esempi A.1.
1. L'esempio più semplice di varietà differenziabile, di classe C^∞ e dimensione n, è ogni sottoinsieme non vuoto e aperto di $\mathbb{R}^n$ (includendo $\mathbb{R}^n$ stesso) con una struttura differenziabile standard individuata dalla funzione identità (che da sola definisce un atlante).
2. Si consideri sfera unitaria $\mathbb{S}^2$ (dotata della topologia ereditata da $\mathbb{R}^3$) in $\mathbb{R}^3$, centrata nell'origine e quindi di equazione, in coordinate canoniche x^1, x^2, x^3 di $\mathbb{R}^3$:

$$\mathbb{S}^2 := \left\{ (x^1, x^2, x^3) \in \mathbb{R}^3 \;\middle|\; (x^1)^2 + (x^2)^2 + (x^3)^2 = 1 \right\} .$$

$\mathbb{S}^2$ acquista una struttura di varietà differenziabile, di dimensione 2 e classe C^∞, da quella di $\mathbb{R}^3$, definendo un atlante su $\mathbb{S}^2$ costituito da 6 carte locali $(\mathbb{S}^2_{(i)\pm}, \phi^{(i)}_\pm)$ $(i = 1, 2, 3)$ ottenute come segue. Considerato l'asse x^i $(i = 1, 2, 3)$ e la coppia di emisferi aperti $\mathbb{S}^2_{(i)\pm}$ con asse sud-nord dato dall'asse x^i, si considerano le carte locali $\phi^{(i)}_\pm : \mathbb{S}^2_{(i)\pm} \to \mathbb{R}^2$ che associano ad ogni $p \in \mathbb{S}^2_{(i)\pm}$ le coordinate di esso sul piano a $x^i = 0$. Si può provare (vedi sotto) che è impossibile dotare $\mathbb{S}^2$ di una carta globale a differenza di $\mathbb{R}^3$ (o di ogni suo sottoinsieme aperto). Questo fatto dimostra che la classe delle varietà differenziabili non si riduce ai soli sottoinsiemi non vuoti aperti degli $\mathbb{R}^n$ ed è pertanto interessante. Un'esempio analogo è quello di una circonferenza in $\mathbb{R}^2$.

[1] Tecnicamente parlando strutture, differenziabili non diffeomorfe. Si rimanda ai corsi di geometria differenziale per approfondimenti.

A.1.2 Varietà prodotto.

Date due varietà differenziabili M ed N, di dimensione m ed n rispettivamente, ed entrambe di classe C^k, si può costruire una terza varietà differenziabile di classe C^k e dimensione $m+n$ sull'insieme di punti $M \times N$ dotato della topologia prodotto. (Tale spazi topologico è quindi ancora di Hausdorff ed a base numerabile.) Tale varietà è detta *varietà prodotto* di M e N e si indica semplicemente con $M \times N$. La struttura differenziabile di $M \times N$, detta *struttura differenziabile prodotto*, è quella che si ottiene come segue. Se (U, ϕ) e (V, ψ) sono due carte locali su M ed N rispettivamente, è immediato verificare che

$$U \times V \ni (p, q) \mapsto (\phi(p), \psi(q)) =: \phi \oplus \psi(p, q) \in \mathbb{R}^{m+n} \tag{A.1}$$

è un omeomorfismo locale. Inoltre, se (U', ϕ') e (V', ψ') sono altre due carte locali su M ed N rispettivamente, k-compatibili con le rispettive precedenti, le carte $(U \times V, \phi \oplus \psi)$ e $(U' \times V', \phi' \oplus \psi')$ risultano essere banalmente k-compatibili. Infine, al variare delle carte (U, ϕ) e (V, ψ) nelle strutture differenziabili di M e N, le carte $(U \times V, \phi \oplus \psi)$ definiscono un atlante su $M \times N$. la struttura differenziabile da esso generata è , per definizione, la struttura differenziabile prodotto.

Definizione A.3. Date due varietà differenziabili M ed N, di dimensione m ed n rispettivamente, ed entrambe di classe C^k, la **varietà prodotto** $M \times N$ è la varietà sull'insieme $M \times N$, dotato della topologia prodotto, e con struttura differenziabile indotta dalle carte locali $(U \times V, \phi \oplus \psi)$ definite in (A.1), quando (U, ϕ) e (V, ψ) variano nelle strutture differenziabili di M e N. $\diamondsuit$

A.1.3 Funzioni differenziabili.

Dato che una varietà differenziabile, localmente è indistiguibile da $\mathbb{R}^n$, la struttura di varietà differenziabile permette di dare senso alla nozione di *funzione differenziabile* definita su un insieme che non sia un $\mathbb{R}^n$ oppure un suo sottoinsieme, ma che abbia la struttura di varietà differenziabile. L'idea è banalmente quella di ridursi, localmente, alla definizione standard in $\mathbb{R}^n$ di funzione differenziabile, usando la struttura di carte locali che ricoprono ogni varietà differenziabile. Se M è una varietà differenziabile di dimensione n e classe C^k, diremo che $f : M \to \mathbb{R}$ è differenziabile con continuità fino all'ordine ordine $p \leq k$, oppure più brevemente, che f è di *classe* C^p, se le funzioni $f \circ \phi^{-1}$ sono di classe C^p come funzioni da $\mathbb{R}^n$ in $\mathbb{R}$ per ogni carta locale (U, ϕ) su M. Tenendo conto che $\mathbb{R}$ è una varietà differenziabile, possiamo dare la seguente definizione del tutto generale che include il caso appena considerato (se $N = \mathbb{R}$ con struttura differenziabile standard).

Definizione A.4. Siano M, N varietà differenziabili di dimensione m, n e classe C^p e C^q $(p, q \geq 1)$. Una funzione continua $f : M \to N$ è di **classe** C^k $(0 \leq k \leq p, q$ eventualmente $k = \infty)$ se $\psi \circ f \circ \phi^{-1}$ è di classe C^k, come funzione da $\mathbb{R}^m$ in $\mathbb{R}^n$, per ogni scelta delle carte

locali $(U, \phi), (V, \psi)$, rispettivamente in N e M.
La classe delle funzioni differenziabili di classe $k = 0, 1, 2, \ldots, \infty$ da M ad N è indicata con $C^p(M; N)$; se $N = \mathbb{R}$ si scrive semplicemente $C^k(M)$.
Un **k-diffeomorfismo** $f : M \to N$ tra due varietà M, N è una funzione di classe C^k, biettiva, con inversa di classe C^k. Se M ed N sono connesse da un k-diffeomorfismo f si dicono varietà **k-diffeomorfe** (tramite f). $\diamondsuit$

Osservazioni A.2.
(1) Si noti che abbiamo ammesso il caso di funzioni differenziabili di classe C^0, che in realtà corrisponde a funzioni solamente continue ed a omeomorfismi nel caso di 0-diffeomorfismi. Ovviamente ogni k-diffeomorfismo è anche un omeomorfismo per cui, per esempio, non ci possono essere diffeomorfismi tra $\mathbb{S}^2$ e $\mathbb{R}^2$ (oppure ogni suo sottoinsieme non vuoto e aperto), essendo il primo compatto ed il secondo no. Questo fatto prova, come menzionato sopra, che non possono esistere carte globali su $\mathbb{S}^2$.
(2) Si dimostra facilmente che, affinché $f : M \to N$ sia C^p, è sufficiente che $\psi \circ f \circ \phi^{-1}$ siano funzioni C^k al variare delle carte locali $(U, \phi), (V, \psi)$ in due atlanti, rispettivamente su M ed N, senza dover controllare la validità di tale condizione per *tutte* le possibile carte locali delle due varietà .
(3) Se $f : M \to N$ è una funzione differenziabile (di classe C^k) e sono assegnate carte locali $(U, \phi), (V, \psi)$, rispettivamente in N e M, la funzione $\psi \circ f \circ \phi^{-1}$ è detta **rappresentazione in coordinate** di f.

A.1.4 Sottovarietà embedded.

Un altro utile concetto è quello di *sottovarietà embedded.* $\mathbb{R}^n$ è una sottovarietà embedded di $\mathbb{R}^m$ con $m > n$. In coordinate canoniche $x^1, \cdots, x^m$ di $\mathbb{R}^m$, $\mathbb{R}^n$ si identifica con il sottoinsieme individuato dalle condizioni $x^{n+1} = \cdots = x^m = 0$ e le prime n coordinate di $\mathbb{R}^m$, $x^1, \cdots, x^n$, sono identificate con le coordinate standard di $\mathbb{R}^n$. L'idea è quella di generalizzare, in senso locale, questa situazione usando coordinate locali e considerando varietà N ed M in luogo di $\mathbb{R}^n$ e $\mathbb{R}^m$.

Definizione A.5. Sia data una varietà differenziabile M, di dimensione $m > n$ e classe C^k $(k \geq 1)$. Una **sottovarietà embedded di M di dimensione n e classe C^k**, N, è una varietà differenziabile (di dimensione n e classe C^k) costituita come segue.
(a) N è un sottoinsieme di M dotato della topologia ereditata da M.
(b) La struttura differenziabile di N è quella generata da un atlante $\{(U_i, \phi_i)\}_{i \in I}$ in cui:
(i) $U_i = V_i \cap N$ e $\phi_i = \psi\restriction_{V_i \cap N}$ per una opportuna carta locale (V_i, ϕ_i) su M;
(ii) nelle coordinate $x^1, \cdots, x^m$ associate a (V_i, ϕ_i), l'insieme $V_i \cap N$ è individuato dalla richiesta $x^{n+1} = \cdots = x^m = 0$ e le rimanenti coordinate $x^1, \cdots, x^n$ sono coordinate locali associate a ϕ_i. $\diamondsuit$

Osservazioni A.3.

(1) La topologia di N è ancora, per costruzione, di Hausdorff ed a base numerabile.
(2) La struttura di sottovarietà embedded su N, se può essere assegnata, è univocamente individuata dalla struttura di varietà differenziabile di M come segue dall'esercizio **4** sotto.

Esercizi A.1.

1. Mostrare che $\mathbb{S}^2$ nell'esempio (2) di esempi A.1 è una sottovarietà embedded di $\mathbb{R}^3$ di dimensione 2.

2. L'insieme N di $\mathbb{R}^2$ ottenuto unendo la circonferenza di raggio unitario centrata in $(0,1)$ con l'asse x^1 può essere dotato della struttura di sottovarietà embedded di $\mathbb{R}^2$? Può essere dotata della struttura di varietà differenziabile di dimensione 1?

3. Considerare la superficie conica C in $\mathbb{R}^3$ di equazione $0 \leq x^3 = \sqrt{(x^1)^2 + (x^2)^2}$. Si dimostri che C *non* può essere dotata della struttura di sottovarietà embedded di $\mathbb{R}^3$ di dimensione 2, ma che può essere dotata della struttura di varietà differenziabile di dimensione 2. Cosa succede se si considera $C^* := C \setminus \{(0,0,0)\}$ in luogo di C?

4. Provare che se M è una varietà differenziabile e $N \subset M$, con topologia ereditata da M, ammette due distinti atlanti $\{(U_i, \phi_i)\}_{i \in I}$ e $\{(U'_j, \phi'_j)\}_{j \in J}$ che soddisfano (b) della definizione A.5 allora le carte dei due atlanti sono tra di loro compatibili.

5. Siano $U \cap V \neq \emptyset$ con (U, ϕ) e (V, ψ) carte locali su M, tali che: $\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$ è biettiva, di classe C^k $(k = 1, 2, \ldots, \infty)$, con determinante della matrice jacobiana non nullo su $\psi(U \cap V)$. Provare che $\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ è anch'essa C^k e quindi le due carte locali sono k-compatibili.

Per concludere citiamo, senza provarlo (vedi per es. [doCarmo92, Westenholtz78]), un importante teorema che permette di stabilire quando un sottoinsieme di una varietà differenziabile può essere dotato della struttura di sottovarietà embedded. La dimostrazione del teorema è una diretta conseguenza del teorema del Dini [GiustiII].

Teorema A.2. (**Teorema dei valori regolari**.) *Sia M una varietà differenziabile di dimensione m e classe C^k. Si consideri l'insieme N determinato da $c(< m)$ costanti, v_j, e da c funzioni di classe C^k, $f_j : M \to \mathbb{R}$*

$$N := \{p \in M \mid f_j(p) = v_j \ , \ j = 1, \cdots, c\} \ .$$

Se, nell'intorno di ogni punto $p \in N$, esiste una carta locale (U, ϕ) di M tale che la matrice jacobiana di coefficienti $\partial(f_j \circ \phi^{-1})/\partial x^i|_{\phi(p)}$ (con $j = 1, \cdots, c$ e $i = 1, \cdots, m$) abbia rango r, allora l'insieme N può essere dotato della struttura di sottovarietà differenziabile embedded di M di dimensione $n := m - c$ e classe C^k.
Se in particolare, la matrice quadrata $c \times c$ di elementi

$$\frac{\partial f_j \circ \phi^{-1}}{\partial x^k} \ , \quad j = 1, \ldots, c \text{ e } k = m - c + 1, m - c + 2, \ldots, m$$

è non singolare in $\phi(p)$ con $p \in N$, le prime n coordinate $x^1, \ldots, x^n$ definiscono un sistema di

coordinate della struttura differenziabile di N in un intorno di p in N. $\diamondsuit$

Osservazioni A.4. La seconda parte del teorema dei valori regolari può essere enunciata rispetto ad una qualsiasi scelta di c coordinate rispetto alle quali la corrispondente matrice quadrata $c \times c$ è non singolare. Le rimanenti n coordinate definiscono, localmente un sistema di coordinate ammissibili su N. La prova di tale fatto è ovvia tenendo conto del fatto che se $\phi : M \ni p \mapsto (x^1(p), x^2(p), \ldots, x^m(p)) \in \mathbb{R}^m$ è un sistema di coordinate locali su M, lo è anche ogni permutazione delle coordinate, per es. $\phi : M \ni p \mapsto (x^2(p), x^1(p), \ldots, x^m(p)) \in \mathbb{R}^m$.

Definizione A.6. Sia $f : M \to \mathbb{R}$ di classe C^k sulla varietà M di classe C^k. f è detta **non singolare in** $p \in M$ se esiste una carta locale (U, ϕ), con $U \ni p$, tale che la matrice riga jacobiana $\partial(f \circ \phi^{-1})/\partial x^i|_{\phi(p)}$ sia non nulla (cioé abbia rango 1). $\diamondsuit$

Se $f : M \to \mathbb{R}$ è una funzione di classe C^k ovunque non singolare sulla varietà M di classe C^k e dimensione m, per il teorema dei valori regolari gli insiemi:

$$\Sigma_t := \{p \in M \mid f(p) = t\} ,$$

sono sottovarietà embedded di classe C^k e dimensione $m - 1$. Per costruzione $\Sigma_t \cap \Sigma_{t'} = \emptyset$ se $t \neq t'$. Infine $\cup_{t \in f(M)} \Sigma_t = M$. In tale situazione si dice che M è una **varietà fogliata** e che le sue **foglie** sono le sottovarietà Σ_t con $t \in f(M)$.

Esempi A.2.
1. L'esercizio A.1.**1** si risolve immediatamente applicando il teorema A.2. La sfera unitaria $\mathbb{S}^2$ in $\mathbb{R}^3$ ha equazione $f(x^1, x^2, x^3) = 0$ dove $f(x^1, x^2, x^3) := \sum_{i=1}^3 (x^i)^2$. f è differenziabile, la matrice rettangolare di coefficienti $\partial f/\partial x^i$ vale $(2x^1, 2x^2, 2x^3)$ e non si annulla su $\mathbb{S}^2$ (dove vale $(2, 2, 2)$). Di conseguenza il suo rango è 1.
2. Il teorema dei valori regolari non può essere direttamente applicato nell'esercizio A.1.**3** nel caso di C in quanto la funzione $f(x^1, x^2, x^3) := x^3 - \sqrt{(x^1)^2 + (x^2)^2}$ non è differenziabile in $(0, 0, 0)$, cioé nel vertice del cono. Se si usa la funzione $g(x^1, x^2, x^3) := (x^3)^2 - ((x^1)^2 + (x^2)^2)$ il teorema non può ancora essere applicato perché il rango della matrice riga di elementi $\partial g/\partial x^i$ è nullo nell'origine.
3. Indichiamo con x, y, z le coordinate standard di $\mathbb{R}^3$ e consideriamo una superficie S che possa vedersi come l'immagine di $z = g(x, y)$, con $(x, y) \in D$ insieme aperto del piano $z = 0$ e con f di classe C^1. Questa superficie è una sottovarietà embedded di dimensione 2 e classe C^1 in $D \times \mathbb{R}$ dotato della struttura di varietà differenziabile standard. Inoltre il sistema di coordinate x, y appartiene alla struttura differenziabile della sottovarietà . Eccone la prova. Possiamo dotare $D \times \mathbb{R}$ del sistema di coordinate $y^1 = x^1, y^2 := x^2, y^3 := z - g(x, y)$: questa trasformazione di coordinate è infatti invertibile, differenziabile con inversa differenziabile come si prova immediatamente, per cui le coordinate y^1, y^2, y^3 appartengono alla struttura differenziabile di $D \times \mathbb{R}$ (e definiscono una carta locale della struttura differenziabile di $\mathbb{R}^3$). S è determinata da $y^3 = 0$ ed è quindi una sottovarietà embedded direttamente dalla definizione.

Alternativamente, si può usare il teorema dei valori regolari, considerando la funzione $f(x,y,z) := z - g(x,y)$ con $(x,y,z) \in D \times \mathbb{R}$, che determina S con la condizione $f(x,y,z) = 0$: in coordinate cartesiane la matrice jacobiana di f è il vettore riga $(\partial f/\partial x, \partial f/\partial y, 1) \neq 0$.
4. Studiando le superfici in $\mathbb{R}^3$ (o equivalentemente in $\mathbb{E}^3$), riferendosi a coordinate cartesiane ortonormali $\mathbf{x} := (x,y,z)$, si è soliti definire una **superficie regolare** come l'immagine S di un'applicazione di classe (almeno) C^1, $\mathbf{x} = \mathbf{x}(u,v)$ con $(u,v) \in D$ sottoinsieme aperto di $\mathbb{R}^2$, in modo tale che

$$\frac{\partial \mathbf{x}}{\partial u} \wedge \frac{\partial \mathbf{x}}{\partial v} \neq \mathbf{0}\,, \quad \text{per ogni } (u,v) \in D\,. \tag{A.2}$$

In effetti una superficie regolare definita in questo modo è una sottovarietà embedded di $\mathbb{R}^3$ di dimensione 2 e classe C^1 e u, v sono coordinate (globali) della sua struttura differenziabile.
Viceversa, se $S \subset \mathbb{R}^3$ è una sottovarietà bidimensionale embedded di dimensione 2 e classe (almeno) C^1, allora S è unione di superfici regolari.
Dimostriamo queste affermazioni. Si fissi $p = (x_p, y_p, z_p) \in S$ corrispondente a (u_p, v_p) e si consideri la condizione (A.2) per il punto p. L'espressione esplicita di tale condizione dice che almeno uno dei tre determinanti jacobiani delle trasformazioni $x = x(u,v), y = y(u,v)$ oppure $x = x(u,v), z = z(u,v)$ oppure $y = y(u,v), z = z(u,v)$ è non nullo in (u_p, v_p). Senza perdere generalità supponiamo che ciò accada per $x = x(u,v), y = y(u,v)$. Per il teorema della funzione inversa A.1, in un intorno aperto D di (x_p, y_p) in $\mathbb{R}^2$, possiamo invertire la funzione $x = x(u,v), y = y(u,v)$ in una funzione C^1, $u = u(x,y), v = v(x,y)$. Pertanto, nell'aperto $M = D \times \mathbb{R}$, $S \cap M$ è rappresentabile come l'immagine della funzione C^1, $z = f(x,y)$ dove $f(x,y) = z(u(x,y), v(x,y))$. Per l'esempio precedente $S \cap M$ è una sottovarietà embedded di M. Dato che le coordinate usate su M sono coordinate locali su $\mathbb{R}^3$ della sua struttura differenziabile standard, e p è arbitrario, quanto visto dimostra che la superficie è una sottovarietà embedded e anche che u, v sono coordinate compatibili con la struttura differenziabile della sottovarietà .
Supponiamo che $S \subset \mathbb{R}^3$ sia una sottovarietà embedded di dimensione 2 e classe C^1. Se $p \in S$, c'è un intorno aperto $U \ni p$ coperto da coordinate y^1, y^2, y^3 della struttura differenziabile C^1 di $\mathbb{R}^3$, tale che $S \cap U$ è descritto da $y^1 = 0$. $u = y^2, v = y^3$ sono coordinate locali su S. Il fatto che le coordinate y^1, y^2, y^3 siano compatibili con le coordinate cartesiane standard di $\mathbb{R}^3$, $\mathbf{x} = (x^1, x^2, x^3)$, implica che la matrice jacobiana di elementi $\frac{\partial x^i}{\partial y^j}$ abbia determinante non nullo. Questo equivale a dire che i tre vettori $\frac{\partial \mathbf{x}}{\partial y^j}$, $j = 1,2,3$ siano ovunque linearmente indipendenti. Questo sarebbe impossibile se fosse

$$\frac{\partial \mathbf{x}}{\partial y^2} \wedge \frac{\partial \mathbf{x}}{\partial y^3} = 0\,,$$

per cui deve valere ovunque:

$$\frac{\partial \mathbf{x}}{\partial u} \wedge \frac{\partial \mathbf{x}}{\partial v} \neq 0\,.$$

Esercizi A.2.

1. Provare l'affermazione in (2) in osservazioni A.1:
Nell'ipotesi $U \cap V \neq \emptyset$, la k-compatibilità delle carte locali (U, ϕ) e (V, ψ) implica che la matrice jacobiana di $\phi \circ \psi^{-1}$, essendo invertibile, abbia determinante ovunque non nullo.

2. Provare che, se $f : M \to \mathbb{R}$ soddisfa il requisito di non singolarità in $p \in M$ rispetto alla carta locale (U, ϕ), allora soddisfa tale requisito rispetto ad ogni altra carta locale definita in un intorno di p.

3. Mostrare che, se per assegnate funzioni f_j e costanti c_j sono soddisfatte le ipotesi del teorema dei valori regolari per il punto p rispetto alla carta locale (U, ϕ), allora le ipotesi sono valide rispetto ad ogni altra carta locale definita in un intorno di p.

A.1.5 Spazio tangente e cotangente. Campi vettoriali covarianti e controvarianti.

Consideriamo la varietà differenziabile M di dimensione n e classe C^k $(k \geq 1)$. Consideriamo ogni fissato spazio $C^k(M)$, come uno spazio vettoriale sul campo $\mathbb{R}$ rispetto alle combinazioni lineari di funzioni definite come, se $a, b \in \mathbb{R}$ e $f, g \in C^k(M)$:

$$(af + bg)(p) := af(p) + bg(p)\,, \quad \text{per ogni } p \in M\,.$$

Fissato un punto $p \in M$, una **derivazione** in p, è un'applicazione $\mathbb{R}$-lineare $L_p : C^k(M) \to \mathbb{R}$ che gode della *proprietà di Leibniz*:

$$L_p(fg) = f(p)L_p(g) + g(p)L_p(f)\,, \quad \text{per ogni } f, g \in C^k(M). \tag{A.3}$$

Evidentemente una combinazione lineare di derivazioni in p, $aL_p + bL'_p$ $(a, b \in \mathbb{R})$ dove

$$(aL_p + bL'_p)(f) := aL_p(f) + bL'_p(f)\,, \quad \text{per ogni } f, g \in C^k(M),$$

è ancora una derivazione. Pertanto le derivazioni in p formano uno spazio vettoriale sul campo $\mathbb{R}$, che indichiamo con $\mathcal{D}^k_p$. Ogni carta locale (U, ϕ) con $U \in p$ definisce automaticamente n derivazioni in p come segue. Se $x^1, \ldots, x^n$ sono le coordinate associate a ϕ, definiamo la derivazione rispetto alla coordinata k-esima:

$$\left.\frac{\partial}{\partial x^k}\right|_p : f \mapsto \left.\frac{\partial f \circ \phi^{-1}}{\partial x^k}\right|_{\phi(p)}\,, \quad \text{per ogni } f, g \in C^1(M). \tag{A.4}$$

Le n derivazioni in p, $\left.\frac{\partial}{\partial x^k}\right|_p$, sono *linearmente indipendenti*: se 0 indica la derivazione nulla e $c^1, c^2, \cdots, c^n \in \mathbb{R}$ sono tali che:

$$\sum_{k=1}^n c^k \left.\frac{\partial}{\partial x^k}\right|_p = 0\,,$$

allora scegliendo una funzione differenziabile che coincide con la funzione coordinata x^l in un intorno aperto di p (la cui chiusura è inclusa in U) e si annulla fuori di esso, la richiesta

$$\sum_{k=1}^n c^k \left.\frac{\partial}{\partial x^k}\right|_p f = 0\,,$$

implica che $c^l = 0$. Dato che possiamo scegliere l arbitrariamente, concludiamo che ogni coefficiente c^r è nullo per $r = 1, 2, \ldots, n$. In definitiva le n derivazioni $\frac{\partial}{\partial x^k}\Big|_p$ formano una base per un sottospazio di $\mathcal{D}_p^k$ di dimensione n (si può in realtà provare che, nel caso di $k = \infty$, questo sottospazio coincide con $\mathcal{D}_p^\infty$ stesso). Cambiando carta locale ed usando (V, ψ) con $V \ni p$ e con coordinate $y^1, \ldots, y^n$, le nuove derivazioni rispetto alle nuove coordinate sono legate alle vecchi dalla relazione

$$\frac{\partial}{\partial y^i}\bigg|_p = \sum_{k=1}^n \frac{\partial x^k}{\partial y^i}\bigg|_{\psi(p)} \frac{\partial}{\partial x^k}\bigg|_p \,. \tag{A.5}$$

La dimostrazione di questo fatto è immediata dalle definizioni date. Dato che la matrice jacobiana di coefficienti $\frac{\partial x^k}{\partial y^i}\Big|_{\psi(p)}$ è biettiva per definizione di carte locali, concludiamo che il sottospazio di $\mathcal{D}_p^k$ generato dalle derivazioni $\frac{\partial}{\partial y^i}\Big|_p$ coincide con quello generato dalle derivazioni $\frac{\partial}{\partial x^k}\Big|_p$. Tale sottospazio è quindi un oggetto *intrinseco*.

Definizione A.7. (**Spazio tangente**). Data una varietà differenziabile di dimensione n e classe C^k $(k \geq 1)$, si consideri un punto $p \in M$.
Il sottospazio vettoriale delle derivazioni in $p \in M$ generato dalle n derivazioni $\frac{\partial}{\partial x^k}\Big|_p$, con $k = 1, 2, \ldots, n$, riferite ad un qualsiasi sistema di coordinate locali (U, ϕ) con $U \ni p$, è detto **spazio tangente in** p **ad** M e si indica con T_pM. Gli elementi dello spazio tangente in p di dicono **vettori tangenti in** p ad M o **vettori controvarianti in** p.
$\diamondsuit$

Ricordiamo che se V è uno spazio vettoriale sul campo $\mathbb{R}$, lo spazio V^* delle funzioni lineari da V in $\mathbb{R}$ è detto **spazio duale** di V. Se la dimensione di V è finita, è tale anche quella di V^* e le due dimensioni coincidono. In particolare, se $\{e_i\}_{i=1,\ldots,n}$ è una base per V, la **base duale** in V^* è la base di V^*, $\{e^{*j}\}_{j=1,\ldots,n}$, completamente individuata dal requisito di linearità e dalle richieste:

$$e^{*j}(e_i) = \delta_i^j \,, \quad \text{per } i, j = 1, \ldots, n \,.$$

Se $f \in V^*$ e $v \in V$, si usa la notazione:

$$\langle v, f \rangle := f(v) \,.$$

Definizione A.8. (**Spazio cotangente**). Data una varietà differenziabile di dimensione n e classe C^k $(k \geq 1)$, si consideri un punto $p \in M$.
Lo spazio duale di T_pM è detto **spazio cotangente in** p **ad** M e si indica con T_p^*M. Gli elementi dello spazio cotangente in p di dicono **vettori cotangenti in** p a M o **vettori covarianti in** p o **1-forme** in p. Per ogni base di elementi $\frac{\partial}{\partial x^k}\Big|_p$ in T_pM, gli n elementi della base duale in T_p^*M vengono indicati con $dx^i|_p$. Per definizione:

$$\left\langle \frac{\partial}{\partial x^k}|_p, dx^i|_p \right\rangle = \delta_k^i \,.$$

$\diamondsuit$

Possiamo ora dare la nozione di campo vettoriale differenziabile sulla varietà M.

Definizione A.9. Se M è una varietà differenziabile di classe C^k e dimensione n, un **campo vettoriale di classe** C^r o **campo vettoriale controvariante** di classe C^r, con $r = 0, 1, \ldots, k$ è un assegnazione di un vettore $v(p) \in T_pM$ per ogni $p \in M$, in modo tale che, per ogni carta locale (U, ϕ) con coordinate $x^1, \ldots, x^n$, per cui

$$v(q) = \sum_{i=1}^{n} v^i(x_q^1, \ldots, x_q^n) \left.\frac{\partial}{\partial x^i}\right|_q ,$$

le n funzioni $v^i = v^i(x^1, \ldots, x^n)$ sono di classe C^r su $\phi(U)$. un **campo covettoriale di classe** C^r o **campo vettoriale covariante** di classe C^r con $r = 0, 1, \ldots, k$ è un assegnazione di un covettore $\omega(p) \in T_pM$ per ogni $p \in M$, in modo tale che, per ogni carta locale (U, ϕ) con coordinate $x^1, \ldots, x^n$, per cui

$$\omega(q) = \sum_{i=1}^{n} v_i(x_q^1, \ldots, x_q^n) \left. dx^i\right|_q ,$$

le n funzioni $\omega_i = \omega_i(x^1, \ldots, x^n)$ sono di classe C^r su $\phi(U)$. $\diamondsuit$

Osservazioni A.5.
(1) Dimostreremo più avanti che la nozione data di vettore tangente, nel caso di spazi affini e quindi nel caso di $\mathbb{R}^n$ in particolare, coincide con la nozione standard.
(2) Sia $v \in T_pM$ e si considerino due carte locali (U, ϕ) e (V, ψ) con $U \cap V \ni p$ e con coordinate, rispettivamente, $x^1, \ldots, x^n$ e $x'^1, \ldots, x'^n$. In tal caso deve valere:

$$v = \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_p = \sum_{j=1}^{n} v'^j \left.\frac{\partial}{\partial x'^j}\right|_p .$$

Pertanto

$$\sum_{i}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_p = \sum_{j,i=1}^{n} v'^j \left.\frac{\partial x^i}{\partial x'^j}\right|_{\psi(p)} \left.\frac{\partial}{\partial x^i}\right|_p ,$$

da cui

$$\sum_{i=1}^{n} \left(v^i - \sum_{j=1}^{n} \left.\frac{\partial x^i}{\partial x'^j}\right|_{\psi(p)} v'^j \right) \left.\frac{\partial}{\partial x^i}\right|_p = 0 .$$

Dato che le derivazioni $\left.\frac{\partial}{\partial x^i}\right|_p$ sono linearmente indipendenti, concludiamo che vale la legge di trasformazioni delle componenti di uno stesso vettore in T_pM, al variare delle coordinate,

$$v^i = \sum_{j=1}^{n} \left.\frac{\partial x^i}{\partial x'^j}\right|_{\psi(p)} v'^j , \tag{A.6}$$

Con la stessa procedura si ottiene l'analoga formula per i vettori covarianti

$$\omega_i = \sum_{j=1}^{n} \left.\frac{\partial x'^j}{\partial x^i}\right|_{\psi(p)} \omega'_j \,, \tag{A.7}$$

quando

$$\omega = \sum_{i=1}^{n} \omega_i \, dx^i\big|_p = \sum_{j=1}^{n} \omega'_j \, dx'^j\Big|_p \,.$$

Vale la seguente proposizione di dimostrazione ovvia, ma molto importante dal punto di vista delle applicazioni.

Proposizione A.1. *Se M è una varietà differenziabile di classe C^k e dimensione n, assegnare un campo vettoriale controvariante X di classe C^r oppure un campo vettoriale controvariante ω di classe C^r è completamente equivalente all'assegnazione di n-ple di funzioni di classe C^r, $\{X^i_{(r)}\}_{i=1,\ldots,n}$ oppure, rispettivamente, $\{\omega_{(r)j}\}_{j=1,\ldots,n}$ – una n-pla per ogni carta locale $(U_{(r)}, \phi_{(r)})$ di un fissato (arbitrariamente) atlante di M – in modo tale che, cambiando carte nell'atlante, valgano le relazioni:*

$$X^i_{(r)}\left(\phi_{(r)}(p)\right) = \sum_{j=1}^{n} \left.\frac{\partial x^i_{(r)}}{\partial x^j_{(r')}}\right|_{\phi_{(r')}(p)} X^j_{(r')}\left(\phi_{(r')}(p)\right)\big) \,, \tag{A.8}$$

oppure, rispettivamente,

$$\omega_{(r')i}\left(\phi_{(r')}(p)\right) = \sum_{j=1}^{n} \left.\frac{\partial x^j_{(r')}}{\partial x^i_{(r)}}\right|_{\phi_{(r)}(p)} \omega_{(r)j}\left(\phi_{(r)}(p)\right) \,, \tag{A.9}$$

per ogni punto $p \in M$. $\diamondsuit$

A.1.6 Differenziali, curve e vettori tangenti.

Consideriamo un campo scalare $f : M \to \mathbb{R}$ di classe C^r sulla varietà M di classe C^k e dimensione n. Supponiamo esplicitamente che $k \geq r > 1$. Se assegnamo su un atlante e per ogni carta dell'atlante le n funzioni $\{\frac{\partial f}{\partial x^i}\}$ si verifica subito che sono rispettate le condizioni della proposizione A.1, di conseguenza abbiamo definito un campo vettoriale covariante.

Definizione A.10. Consideriamo un campo scalare $f : M \to \mathbb{R}$ di classe C^r sulla varietà M di classe C^k e dimensione n e valga $k \geq r > 1$. Il **differenziale** di f, df è il campo vettoriale covariante di classe C^{r-1} individuato, in ogni carta locale (U, ψ) da:

$$df|_p = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x^i}\right|_{\psi(p)} dx^i|_p \,.$$

$\diamondsuit$.

Consideriamo una curva di classe C^r nella varietà M di classe C^k, cioé un'applicazione di classe C^r $(r = 0, 1, \ldots, k)$, $\gamma : I \to M$, dove $I \subset \mathbb{R}$ è un intervallo aperto pensato come sottovarietà differenziabile di $\mathbb{R}$. Supponiamo esplicitamente che $r > 1$. Se $p \in \gamma(I)$, possiamo definire il *vettore tangente* a γ in p come, se $\gamma(t_p) = p$:

$$\dot{\gamma}(p) := \sum_{i=1}^{n} \left.\frac{dx^i}{dt}\right|_{t_p} \left.\frac{\partial}{\partial x^i}\right|_p ,$$

in una qualsiasi carta locale definita nell'intorno di p. La definizione in realtà *non* dipende dalla carta scelta. Infatti, se definissimo

$$\dot{\gamma}'(p) := \sum_{j=1}^{n} \left.\frac{dx'^j}{dt}\right|_{t_p} \left.\frac{\partial}{\partial x'^j}\right|_p ,$$

in riferimento ad un secondo sistema di coordinate definite nell'intorno di p, attraverso la (A.6) otterremmo subito che:

$$\dot{\gamma}(p) = \dot{\gamma}'(p) .$$

Possiamo allora dare la seguente definizione.

Definizione A.11. Una **curva di classe** C^r, $r = 0, 1, \ldots, k$ nella varietà differenziabile M di dimensione n e classe C^k, è un'applicazione di classe C^r, $\gamma : I \to M$, dove $I \subset \mathbb{R}$ è un intervallo aperto (pensato come sottovarietà differenziabile embedded di $\mathbb{R}$). Se $r > 1$, il **vettore tangente** a γ in $p = \gamma(t_p)$ per qualche $t \in I$, è il vettore $\dot{\gamma}(p) \in T_pM$ definito da

$$\dot{\gamma}(p) := \sum_{i=1}^{n} \left.\frac{dx^i}{dt}\right|_{t_p} \left.\frac{\partial}{\partial x^i}\right|_p , \tag{A.10}$$

in una qualsiasi carta locale definita nell'intorno di p. $\diamondsuit$.

A.1.7 Pushforward, pullback, derivata di Lie.

Siano M ed N sono varietà differenziabili (almeno di classe C^1), di dimensione m e n rispettivamente, e $f : N \to M$ una funzione differenziabile (almeno di classe C^1). Per un punto $p \in N$ consideriamo carte locali (U, ϕ) in N e (V, ψ) in M rispettivamente attorno a p e $f(p)$. Indichiamo con $(y^1, \ldots, y^n)$ le coordinate definite in tal modo in U e con $(x^1, \ldots, x^m)$ le coordinate definite in tal modo in V. Definiamo ancora $f^k(y^1, \ldots, y^n) = y^k(f \circ \phi^{-1})$ per $k = 1, \ldots, m$. Si definiscono allora:
(i) il **pushforward** $df_p : T_pN \to T_{f(p)}M$, data in coordinate da:

$$df_p : T_pN \ni \sum_{i=1}^{n} u^i \left.\frac{\partial}{\partial y^i}\right|_p \mapsto \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \left.\frac{\partial f^j}{\partial y^i}\right|_{\phi(p)} u^i \right) \left.\frac{\partial}{\partial x^j}\right|_p , \tag{A.11}$$

e
(ii) il **pullback** $f_p^* : T^*_{f(p)}M \to T^*_p N$, data in coordinate da:

$$f_p^* : T^*_{f(p)}M \ni \sum_{j=1}^m \omega_j dx^j|_{f(p)} \mapsto \sum_{i=1}^n \left(\sum_{j=1}^m \left.\frac{\partial f^j}{\partial y^i}\right|_{\phi(p)} \omega_j \right) dy^i|_p \,. \tag{A.12}$$

Si verifica immediatamente che le definizioni *non dipendono dalle coordinate usate attorno a p e $f(p)$*. Il pushforward è anche indicato con $f_{p*} : T_p N \to T_{f(p)}M$.

Sia X è un campo vettoriale di classe C^1 sulla varietà differenziabile M di classe C^2 e dimensione n. possiamo considerare il gruppo ad un parametro di diffeomorfismi locali $\phi^{(X)}$ (vedi la sezione 3.5.3) da esso generato. Fissiamo $p \in M$ ed un campo vettoriale differenziabile Y su N. Se $\epsilon > 0$ è sufficientemente piccolo, ha senso il vettore in p:

$$(d\phi_t^{(X)})_{\phi_{-t}^{(X)}(p)} Y(\phi_{-t}^{(X)}(p)) \,, \quad \text{per ogni } t \in (-\epsilon, +\epsilon) \,.$$

La **derivata di Lie** del campo Y in p rispetto al campo X è definita come il vettore di T_pM:

$$\mathcal{L}_X|_p Y := \left.\frac{d}{dt}\right|_{t=0} (d\phi_t^{(X)})_{\phi_{-t}^{(X)}(p)} Y(\phi_{-t}^{(X)}(p)) \,. \tag{A.13}$$

In coordinate locali attorno a p, il gruppo ad un parametro trasforma $(x^1, \ldots, x^n)$ in $(x'^1, \ldots, x'^n) = (x_t^1, \ldots, x_t^n)$, pertanto, in coordinate:

$$(\mathcal{L}_X|_p Y)^i := \left.\frac{d}{dt}\right|_{t=0} \sum_{j=1}^n \left.\frac{\partial x'^i}{\partial x^j}\right|_{(x_{-t}^1(p), \ldots, x_{-t}^n(p))} Y^j((x_{-t}^1(p), \ldots, x_{-t}^n(p))) \,.$$

Il calcolo esplicito fornisce allora:

$$\begin{aligned}(\mathcal{L}_X|_p Y)^i &:= \left.\frac{d}{dt}\right|_{t=0} \sum_{j=1}^n \left.\frac{\partial x'^i}{\partial x^j}\right|_{(x_{-t}^1(p), \ldots, x_{-t}^n(p))} Y^j((x^1(p), \ldots, x^n(p))) \;+\; \sum_{j=1}^n \delta_j^i \left.\frac{d}{dt}\right|_{t=0} Y^j((x_{-t}^1(p), \ldots, x_{-t}^n(p))) \\ &= \left.\left(\sum_{j=1}^n \frac{\partial X^i}{\partial x^j} Y^j \;-\; \sum_{k=1}^n X^k \frac{\partial Y^i}{\partial x^k} \right)\right|_{(x^1(p), \ldots, x^n(p))} \,,\end{aligned}$$

dove abbiamo usato il fatto che:

$$\left.\frac{d}{dt}\right|_{t=0} x'^i(x_t^1(p), \ldots, x_t^n(p)) = X^i(x^1(p), \ldots, x^n(p)) \,.$$

In definitiva abbiamo trovato che:

$$\mathcal{L}_X Y = [Y, X] \,, \tag{A.14}$$

dove il **commutatore** o **parentesi di Lie** $[X,Y]$ dei campi vettoriali X e Y di classe C^2 è definito come il campo vettoriale di classe C^1 che in coordinate locali si esprime come:

$$[Y,X](p) = \left(\sum_{j=1}^{n} \frac{\partial X^i}{\partial x^j} Y^j \; - \; X^j \frac{\partial Y^i}{\partial x^j} \right) \Bigg|_{(x^1(p),\ldots,x^n(p))} \frac{\partial}{\partial x^i}\Bigg|_p . \tag{A.15}$$

Intrinsecamente $[X,Y]$ è l'unico campo vettoriale, visto come operatore differenziale, che soddisfa:

$$[Y,X](f) = Y(X(f)) - X(Y(f)) \quad \text{per ogni funzione } f \in C^1(M).$$

Per computo diretto si possono verificare le seguenti proprietà del commutatore visto come funzione che associa un campo vettoriale $[X,Y]$ di classe C^{k-1} a coppie di campi vettoriali X e Y di classe C^k sulla varietà di classe C^r, $r \geq k$. Se X,Y,Z sono campi vettoriali di classe C^k sulla varietà di classe C^r, $r \geq k$:

(i) **antisimmetria**: $[X,Y] = -[Y,X]$,

(ii) $\mathbb{R}$-**linearità** : $[aX + bZ, Y] = a[X,Y] + b[Z,Y]$ per ogni coppia $a,b \in \mathbb{R}$,

(iii) **proprietà di Jacobi**: $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$.

Per la proprietà di Jacobi si è assunto che i tre campi coinvolti siano di classe C^2 e lo 0 a secondo membro indica il campo vettoriale che è ovunque nullo su M. Vale infine, se $\cdot$ indica il prodotto punto per punto:

(iv) $[X, f \cdot Y] = f \cdot [X,Y] + X(f) \cdot Y$ per ogni funzione $f \in C^1(M)$.

A.1.8 Immersione di spazi tangenti per sottovarietà embedded.

Se $N \subset M$ è una sottovarietà differenziabile embedded della varietà differenziabile M (entrambe di classe C^k per qualche $k > 0$) e $n < m$ indicano rispettivamente le dimensioni di N ed M, i punti $p \in N$ sono contemporaneamente punti di M attraverso la funzione $\imath$ che identifica N con un sottoinsieme di M. La stessa cosa vale per i vettori di T_pN che possono essere visti, in modo naturale, come vettori in T_pM usando il pushforward $d\imath_p$ come ora mostriamo in dettaglio. Scegliamo una carta locale attorno a p, (U,ϕ) in N ed una analoga carta locale (V,ψ) attorno a p in M. Supponiamo che $\phi : U \ni q \mapsto (y^1(q),\ldots,y^n(q))$ mentre $\psi : V \ni q \mapsto (x^1(q),\ldots,x^m(q))$. La funzione identità $\imath : N \to M$ che identifica N come varietà con N come sottoinsieme di M è differenziabile e

$$\psi \circ \imath \circ \phi^{-1} : (y^1,\ldots,y^n) \mapsto (x^1(y^1,\ldots y^n),\ldots,x^m(y^1,\ldots y^n)) .$$

Possiamo allora passare al pushforward $d\imath_p : T_pN \to T_{f(p)}M$, che come detto precedentemente, nelle coordinate dette è individuato da:

$$d\imath_p : \sum_{i=1}^{n} v^i \frac{\partial}{\partial y^i}\Bigg|_p \mapsto \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \frac{\partial x^k}{\partial y^i}\Bigg|_p v^i \right) \frac{\partial}{\partial x^k}\Bigg|_p , \tag{A.16}$$

Sappiamo che l'applicazione $d\imath_P$ non dipende dalla scelta delle carte locali attorno a p, (U,ϕ) e (V,ψ): se fossimo partiti da altre carte locali avremmo ottenuto la stessa funzione $d\imath_p$. In

virtù di tale fatto proviamo che $d\imath_p$ è iniettiva.
Scegliendo coordinate in M adattate ad N, cioé una carta (V, ψ) con $\psi : V \ni q \mapsto (x^1(q), \ldots, x^m(q))$ in modo tale che $N \cap U$ corrisponda ai punti di coordinate $x^{n+1} = \cdots = x^m = 0$, le prime n coordinate $y^1 = x^1, \ldots y^n = x^n$ definiscono una carta locale su N. Questo è vero per la definizione stessa di sottovarietà embedded. Con questa scelta di coordinate locali in N e M, l'espressione esplicita di $d\imath_p$ è banale

$$d\imath_p : \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial y^i}\right|_p \mapsto \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_p .$$

È evidente in questa rappresentazione che l'applicazione $d\imath_p$ è iniettiva. Pertanto l'applicazione $d\imath_p$ identifica T_pN con un sottospazio di T_pM.
Per ogni scelta di carte locali attorno a p, in N ed M rispettivamente, i vettori di base $\frac{\partial}{\partial y^i}|_p \in T_pN$ si identificano quindi, tramite $d\imath_p$, con vettori di T_pM. In questo senso un po' impropriamente si può scrivere:

$$\left.\frac{\partial}{\partial y^i}\right|_p = \sum_{k=1}^{m} \left.\frac{\partial x^k}{\partial y^i}\right|_p \left.\frac{\partial}{\partial x^k}\right|_p . \tag{A.17}$$

A.1.9 Fibrato tangente e cotangente, varietà fibrate e sezioni.

Gli esempi più semplici di varietà fibrate sono il *fibrato tangente* ed il *fibrato cotangente* associati alla varietà differenziabile M. Prima di dare la definizione generale definiamo questi due oggetti. Consideriamo la varietà differenziabile M di dimensione n e classe C^k con $k \geq 2$ e consideriamo l'insieme ad essa associato:

$$TM := \{(p, v) \mid p \in M,\ v \in T_pM\} .$$

È possibile dotare tale insieme di una struttura naturale di varietà differenziabile di dimensione $2n$ e classe $k - 1$; la varietà differenziabile ottenuta in questo modo (che si indica ancora con TM) si chiama **(varietà) fibrato tangente** di M. La struttura di varietà differenziabile è l'unica che ammette come atlante il seguente insieme di carte locali indotte dalla struttura differenziabile di M. Per ogni una carta locale della struttura differenziabile $\mathcal{A}_M$ di M, (U, ϕ), con $\psi : U \ni p \mapsto (x^1(p), \ldots, x^n(p)) \in \mathbb{R}^n$, definiamo

$$TU := \{(p, v) \in TM \mid p \in U\}$$

e definiamo l'applicazione iniettiva

$$T\psi : TU \ni (p, v) \mapsto (x^1(p), \ldots, x^n(p), v^1, \ldots, v^n) \in \mathbb{R}^{2n} , \quad \text{dove } v = \textstyle\sum_k v^k \left.\frac{\partial}{\partial x^k}\right|_p .$$

Dotiamo infine TM della topologia generata dalle controimmagini degli insiemi aperti di $\mathbb{R}^{2n}$ secondo le funzioni $T\psi$. Si osservi che tale topologia rende TM spazio di Hausdorff, a base numerabile, localmente omeomorfo a $\mathbb{R}^{2n}$ (omeomorfismi locali sono proprio le funzioni $T\psi$). È

facile verificare che l'insieme delle coppie $T\mathcal{A}_M := \{(TU, T\psi) \mid (U, \psi) \in \mathcal{A}_M\}$ è un atlante su TM di classe C^{k-1}. Spieghiamo l'origine del $k-1$. Se $(TU, T\psi)$ e $(TV, T\phi)$ sono carte locali nell'atlante detto e $T\psi : (p, v) \mapsto (x^1, \ldots, x^n, \dot{x}^1, \ldots \dot{x}^n)$ mentre $T\phi : (p, v) \mapsto (y^1, \ldots, y^n, \dot{y}^1, \ldots \dot{y}^n)$, su $TU \cap TV$ (ammesso che non sia vuoto) valgono relazioni di forma (con ovvie notazioni):

$$y^i = y^i(x^1, \ldots, x^n) , \tag{A.18}$$

$$\dot{y}^i = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j} \dot{x}^j . \tag{A.19}$$

La presenza della matrice jacobiana nella seconda equazione abbassa di 1 la classe di differenziabilità di TM rispetto a M.
La struttura differenziabile $\mathcal{A}_{TM}$ generata dall'atlante $T\mathcal{A}_M$ rende TM una varietà differenziabile di dimensione n e classe C^{k-1} detta **(varietà) fibrato tangente** di M.
La funzione *suriettiva* $\Pi : TM \ni (p, v) \mapsto p \in M$ risulta essere di classe C^{k-1} e si chiama **proiezione canonica**. Si osservi che se (U, ψ) è una carta locale su M, risulta che $TU = \Pi^{-1}(U)$. La varietà M si dice **base** del fibrato tangente. Per $p \in M$, lo spazio tangente $T_pM = \Pi^{-1}(p)$, che risulta essere una sottovarietà embedded di TM, si dice **fibra** di TM nel punto $p \in M$.

Il **fibrato cotangente** T^*M viene definito in modo del tutto analogo assegnando una struttura naturale di varietà differenziabile di dimensione $2n$ e classe C^{k-1} sull'insieme:

$$T^*M := \{(p, \omega_p) \mid p \in M, \, \omega_p \in T_p^*M\} .$$

Per ogni una carta locale della struttura differenziabile $\mathcal{A}_M$ di M, (U, ϕ), con $\psi : U \ni p \mapsto (x^1(p), \ldots, x^n(p)) \in \mathbb{R}^n$, definiamo

$$T^*U := \{(p, \omega_p) \in T^*M \mid p \in U\}$$

e definiamo l'applicazione iniettiva

$$T^*\psi : T^*U \ni (p, \omega_p) \mapsto (x^1(p), \ldots, x^n(p), (\omega_p)_1, \ldots, (\omega_p)_n) \in \mathbb{R}^{2n} , \quad \text{dove } \omega_p = \textstyle\sum_k (\omega_p)_k \, dx^k\big|_p .$$

L'atlante su T^*M contenente tutte le carte $(T^*U, T^*\psi)$ con $(U, \psi) \in \mathcal{A}_M$, definisce la struttura di varietà differenziabile di dimensione $2n$ e classe C^{k-1} su T^*M. Consideriamo la coppia di carte locali $(T^*U, T^*\psi)$ e $(T^*V, T^*\phi)$ su T^*M dove $T^*\psi : (p, v) \mapsto (x^1, \ldots, x^n, \tilde{x}_1, \ldots \tilde{x}_n)$ e dove $T^*\phi : (p, v) \mapsto (y^1, \ldots, y^n, \tilde{y}_1, \ldots \tilde{y}_n)$. Su $T^*U \cap T^*V$ (ammesso che non sia vuoto) valgono relazioni di forma (con ovvie notazioni):

$$y^i = y^i(x^1, \ldots, x^n) , \tag{A.20}$$

$$\tilde{y}_i = \sum_{j=1}^{n} \frac{\partial x^j}{\partial y^i} \tilde{x}_j . \tag{A.21}$$

La funzione *suriettiva* $\Pi : T^*M \ni (p, v) \mapsto p \in M$ risulta essere di classe C^{k-1} e si chiama **proiezione canonica**. Si osservi che se (U, ψ) è una carta locale su M, risulta che $T^*U =$

$\Pi^{-1}(U)$. La varietà M si dice **base** del fibrato cotangente. Per $p \in M$, lo spazio cotangente $T_p^*M = \Pi^{-1}(p)$, che risulta essere una sottovarietà embedded di T^*M, si dice **fibra** di T^*M nel punto $p \in M$.

Osservazioni A.6. Riferiamoci a TM anche se quanto diremo vale banalmente anche per T^*M. Prima di tutto notiamo che tutte le fibre T_pM sono diffeomorfe tra di loro essendo tutte spazi vettoriali della stessa dimensione e dato che gli isomorfismi di spazi vettoriali sono diffeomorfismi rispetto alla struttura differenziale delle fibre indotta da quella di TM. Tuttavia ci sono infiniti diffeomorfismi per ogni coppia T_pM, T_qM: nessuno di questi diffeomorfismi è più naturale degli altri. Quindi le fibre sono tutte diffeomorfe ma in modo non canonico e sono di conseguenza diffeomorfe in modo non canonico a $\mathbb{R}^n$ che indicheremo con F e penseremo come la "fibra astratta". Se fissiamo $p \in M$, c'è un intorno U di p e un diffeomorfismo $f_U : U \times F \to \Pi^{-1}(U)$ che soddisfa $\Pi(f_U(q, v)) = q$ per ogni $q \in U$ e $v \in F$. L'intorno U può essere scelto come il dominio della carta locale (U, ψ) di M ed il diffeomorfismo f_U come la funzione $(T\psi)^{-1}$ definita precedentemente. Abbiamo provato che TM è *localmente diffeomorfo* a $M \times F$.

La definizione di varietà fibrata si ottiene astraendo dalle precedenti costruzioni e tenendo conto dell'osservazione fatta sopra.

Definizione A.12. **(Varietà fibrata e sezioni).** Una **varietà fibrata** E (detta anche semplicemente **(spazio) fibrato**) è individuata da: (i) una varietà differenziabile, indicata ancora con E, (ii) una seconda varietà differenziabile M detta **base**, (iii) un'applicazione differenziabile suriettiva $\Pi : E \to M$ detta **proiezione canonica** e (iiv) una terza varietà differenziabile F detta **fibra standard**. Devono essere infine soddisfatte le seguenti due condizioni:
(a) per ogni $p \in M$ l'insieme $F_p := \Pi^{-1}(p)$, detto **fibra in** p, deve essere sottovarietà embedded di E che risulta anche essere diffeomorfa (in modo non canonico in generale) a F;
(b) E deve essere **localmente diffeomorfo** a $M \times F$ nel senso che segue: per ogni $p \in M$ deve esistere un suo intorno aperto U e un diffeomorfismo $f_U : U \times F \to \Pi^{-1}(U)$ con $\Pi(f_U(x, y)) = x$ per ogni $x \in U$ e $y \in F$.
Una **sezione** di E è un'applicazione differenziabile $s : M \to E$ con $\Pi(s(x)) = x$ per ogni $x \in M$.
$\diamondsuit$

N.B. Nella definizione data si suppone che gli ordini di differenziabilità delle funzioni e delle varietà coinvolte siano assegnati in modo coerente.

A.2 Elementi di Geometria Affine.

A.2.1 Spazi affini.

Lo spazio tridimensionale della fisica classica, ma anche lo spaziotempo quadridimensionale della fisica relativistica speciale, sono *spazi affini*. Tali spazi ammettono particolari trasformazioni, chiamate *traslazioni* che corrispondono, quando lo spazio affine viene dotato di significato fisico,

alle operazioni fisiche di traslazione dei corpi materiali. L'esistenza delle traslazioni è connessa alla proprietà di *omogeneità* dello spazio di quiete di un riferimento inerziale ed è il punto di partenza per definire la nozione di impulso e di energia nelle formulazioni avanzate della meccanica. Torneremo più avanti su questi argomenti. Dal punto di vista puramente matematico, gli spazi affini sono gli spazi della geometria di Euclide[2].
Richiamiamo brevemente la definizione e le principali caratteristiche degli spazi affini e degli spazi euclidei. Tali nozioni dovrebbero già essere note dai corsi di geometria elementare.

Definizione A.13. Uno **spazio affine (reale) di dimensione (finita)** n è un insieme $\mathbb{A}^n$, i cui elementi sono detti **punti**, dotato di alcune strutture che descriviamo di seguito.
(1) Uno spazio vettoriale reale n-dimensionale V, detto **spazio delle traslazioni** o **spazio dei vettori liberi**.
(2) Un'applicazione $\mathbb{A}^n \times \mathbb{A}^n \ni (P, Q) \mapsto P - Q \in V$ che gode delle due seguenti proprietà:
(i) per ogni coppia di elementi $Q \in \mathbb{A}^n$, $\mathbf{v} \in V$ c'è un *unico* punto $P \in \mathbb{A}^n$ tale che $P - Q = \mathbf{v}$;
(ii) $P - Q + Q - R = P - R$ per ogni terna $P, Q, R \in \mathbb{A}^n$.
Se $Q \in \mathbb{A}^n$ e $\mathbf{v} \in V$, $Q + v \in \mathbb{A}^n$ indica l'unico punto P in $\mathbb{A}^n$ tale che $P - Q = v$. Una **retta** in $\mathbb{A}^n$ di **origine** P e **vettore tangente** $\mathbf{u}$ è la funzione $\mathbb{R} \ni t \mapsto P + t\mathbf{u} \in \mathbb{A}^n$. Un **segmento** di retta si ottiene restringendo t ad un intervallo (diverso da un punto). $\diamondsuit$

N.B. *Gli spazi affini considerati in queste dispense sono esclusivamente reali e con dimensione finita.*

Esercizi A.3.
1. Provare che, per ogni $P \in \mathbb{A}^n$, $P - P = \mathbf{0}$ vettore nullo di V.
2. Provare che, se $Q \in \mathbb{A}^n$ e $\mathbf{u}, \mathbf{v} \in V$ allora:

$$(Q + \mathbf{u}) + \mathbf{v} = Q + (\mathbf{u} + \mathbf{v}) \,. \tag{A.22}$$

3. Provare che, se $Q, P \in \mathbb{A}^n$ allora:

$$P - Q = -(Q - P) \,. \tag{A.23}$$

4. Provare che, se $P, Q \in \mathbb{A}^n$ e $\mathbf{u} \in V$ allora:

$$P - Q = (P + \mathbf{u}) - (Q + \mathbf{u}) \,. \tag{A.24}$$

[2]Talvolta si trova scritto che la geometria di Euclide (in n dimensioni) è la geometria in $\mathbb{R}^n$ e, per esempio, si parla di $\mathbb{R}^2$ come del *piano euclideo*. Questo punto di vista non è corretto (anche dal punto di vista fisico), perché $\mathbb{R}^n$ ha una struttura che *non* è invariante per traslazioni, per esempio l'origine $(0, \ldots, 0)$ è un punto privilegiato (che non ha alcun corrispondente fisico!). Ignorare la parte della struttura di $\mathbb{R}^n$ che non invariante per traslazioni significa appunto vedere $\mathbb{R}^n$ come uno spazio affine. Si legga a tal proposito la "definizione" di spazio affine data a p.13 del fondamentale testo di Arnold di Metodi Matematici della Meccanica Classica [Arnold92].

Ogni spazio affine $\mathbb{A}^n$ ammette una struttura naturale di varietà differenziabile (di classe C^∞) indotta da una classe di sistemi di coordinate globali naturali, tra di loro compatibili, detti *sistemi di coordinate cartesiane*. Un tale sistema di coordinate si costruisce come segue. Si fissi un punto $O \in \mathbb{A}^n$, detto **origine** delle coordinate, ed una base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ dello spazio delle traslazioni V, detta **sistema di assi** delle coordinate. Variando $P \in \mathbb{A}^n$ le componenti, $((P-O)^1, \ldots, (P-O)^n)$, di ogni vettore $P-O$ rispetto alla base scelta definiscono una funzione biettiva $f : \mathbb{A}^n \to \mathbb{R}^n$ che permette di identificare i punti di $\mathbb{E}^n$ con i punti di $\mathbb{R}^n$. Questa corrispondenza tra punti, P, e n-ple, $((P-O)^1, \ldots, (P-O)^n)$, è biettiva. È iniettiva per la richiesta (i) nella definizione A.13 ed è suriettiva perché il dominio di $(P, O) \mapsto P - O \in V$ coincide, per definizione, con tutto $\mathbb{A}^n \times \mathbb{A}^n$ ed ora O è tenuto fisso.

Definizione A.14. Nello spazio affine $\mathbb{A}^n$ con spazio delle traslazioni V si fissi un punto $O \in \mathbb{E}^n$ ed una base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ di V. La carta globale $(\mathbb{A}^n, f)$, dove f associa a $P \in \mathbb{A}^n$ la n-pla di componenti di $P - O$, rispetto alla base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, è detta **sistema di coordinate cartesiane con origine** O **e assi** $\mathbf{e}_1, \ldots, \mathbf{e}_n$. I sistemi di coordinate (locali) non cartesiane sono detti sistemi di coordinate **curvilinee**. $\diamondsuit$

La topologia euclidea di $\mathbb{R}^n$ induce tramite f una topologia su $\mathbb{A}^n$ (definendo gli aperti di $\mathbb{A}^n$ come le controimmagini degli aperti di $\mathbb{R}^n$) che lo rende spazio topologico omeomorfo a $\mathbb{R}^n$ e quindi di Hausdorff a base numerabile. Si verifica facilmente che la topologia definita in tal modo non dipende dalla scelta di O e nemmeno dalla scelta della base in V. Ulteriormente la funzione f definisce da sola un atlante C^∞ su $\mathbb{A}^n$ e dota tale spazio di una struttura differenziabile generata da tale atlante. In base agli esercizi A.4.**1** e **2** si ha che, passando ad un altro sistema di coordinate cartesiane, $(\mathbb{E}^n, g)$, le funzioni $f \circ g^{-1}$ e $g \circ f^{-1}$ sono funzioni lineari non omogenee e quindi infinitamente differenziabili. Concludiamo che tutte le carte globali individuate da sistemi di coordinate cartesiane sono a due a due compatibili e che inducono, di conseguenza, la stessa struttura differenziabile di classe C^∞. Possiamo dire che tale struttura differenziabile è generata dalla sola struttura di spazio affine senza scelte arbitrarie, in questo senso è *naturale*. Considerando uno spazio affine $\mathbb{A}^n$ come una varietà differenziabile una domanda che sorge spontanea è quale relazione intercorra tra lo spazio tangente ad ogni punto $T_P\mathbb{A}^n$, $P \in \mathbb{A}^n$ e lo spazio delle traslazioni V. Si ha il seguente teorema.

Teorema A.3. *Sia $\mathbb{A}^n$ uno spazio affine con spazio delle traslazioni V. Si fissi $P \in \mathbb{A}^n$ e si consideri $T_P\mathbb{A}^n$. Vale quanto segue.*
(a) *Se $(\mathbb{A}^n, f)$ è un sistema di coordinate cartesiane di origine O e con assi $\mathbf{e}_1, \cdots, \mathbf{e}_n \in V$, indicate con $x^1, \ldots, x^n$ le funzioni coordinate della carta globale detta, l'applicazione:*

$$\chi_P : V \ni \sum_{i=1}^n v^i \, \mathbf{e}_i \mapsto \sum_{i=1}^n v^i \left.\frac{\partial}{\partial x^i}\right|_P \in T_P\mathbb{A}^n \,, \tag{A.25}$$

è un isomorfismo di spazi vettoriali.

(b) *Se χ'_P è un analogo isomorfismo ottenuto partendo da un'altro sistema di cartesiane, vale*

$$\chi_P = \chi'_P \,,$$

e in questo senso l'isomorfismo χ_P è naturale.

Dimostrazione. (a) Dato che sia V che $T_p\mathbb{A}$ hanno dimensione finita n, l'applicazione lineare χ_P definita in (A.25) tramite l'identificazione delle componenti di un vettore rispetto a due basi nei rispettivi spazi, è un isomorfismo. (b) Per l'esercizio A.4.**1** Se $(\mathbb{A}^n, f)$ è un sistema di coordinate cartesiane con coordinate $x^1, \cdots, x^n$ e $(\mathbb{A}^n, g)$ un altro sistema di coordinate cartesiane, con coordinate ${x'}^1, \cdots, {x'}^n$, di origine O' e assi $\mathbf{e'}_1, \ldots, e'_n$, in modo che valga

$$\mathbf{e}_i = \sum_j B^j{}_i \mathbf{e'}_j \,, \tag{A.26}$$

allora la funzione $g \circ f^{-1}$ è espressa, in coordinate, dalle relazioni:

$$x'^j = \sum_{i=1}^n B^j{}_i (x^i + b^i), \tag{A.27}$$

dove $(O - O') = \sum_i b^i e_i$. Definiamo allora:

$$\chi'_P : V \ni \sum_{j=1}^n v'^j \, \mathbf{e}'_j \mapsto \sum_{i=j}^n v'^j \left.\frac{\partial}{\partial x'^j}\right|_P \in T_P\mathbb{A}^n \,, \tag{A.28}$$

Come conseguenza di (A.27) si ha che.

$$\left.\frac{\partial}{\partial x^i}\right|_P = \sum_j B^j{}_i \left.\frac{\partial}{\partial x'^j}\right|_P \,. \tag{A.29}$$

Da (A.26) e (A.29) notiamo che la matrice di passaggio tra le basi $\{\mathbf{e}'_j\}$ e $\{\mathbf{e}_i\}$ nei primi membri di (A.25) e (A.28) è la stessa che si ha tra le basi $\{\left.\frac{\partial}{\partial x'^j}\right|_P\}$ e$\{\left.\frac{\partial}{\partial x^i}\right|_P\}$ nei secondi membri. Questo comporta immediatamente che $\chi_P = \chi'_P$. $\square$

Osservazioni A.7.
(1) L'identificazione dovuta a χ_P identifica vettori dello spazio V con vettori dello spazio tangente $T_P\mathbb{A}^n$. I primi non sono associati in alcun punto di $\mathbb{A}^n$ e, nella letteratura geometrica tradizionale, sono chiamati *vettori liberi*, mentre i secondi sono associati al punto $P \in \mathbb{A}^n$ e, nella letteratura geometrica tradizionale, sono chiamati *vettori applicati*. Questa distinzione ha senso solo negli spazi affini. Nelle varietà differenziabili generiche esistono solo i vettori applicati.
(2) In base al teorema provato, se $(\mathbb{A}^n, f)$ è un sistema di coordinate cartesiane di origine O e

con assi $\mathbf{e}_1, \cdots, \mathbf{e}_n \in V$, indicate con $x^1, \ldots, x^n$ le funzioni coordinate della carta globale detta sussiste l'identificazione canonica (dovuta all'esistenza dell'isomorfismo χ_P):

$$\mathbf{e}_i \equiv \left.\frac{\partial}{\partial x^i}\right|_P . \tag{A.30}$$

In particolare, il vettore tangente ad una curva $P = P(t)$ parametrizzata in coordinate cartesiane come $x^i = x^i(t)$, che in geometria differenziale è dato da

$$\mathbf{t}(t) = \sum_i \frac{dx^i}{dt} \left.\frac{\partial}{\partial x^i}\right|_{P(t)} ,$$

è riscrivibile come:

$$\mathbf{t}(t) = \sum_i \frac{dx^i}{dt}\, \mathbf{e}_i .$$

In realtà quando si usa la prima notazione, si pensa in vettore tangente come vettore applicato, quando si usa la seconda lo si pensa come vettore libero.

Le coordinate cartesiane sono importanti anche perché consentono di rappresentare in maniera semplice le cosiddette *trasformazioni affini*. Una trasformazione affine è una trasformazione che preserva la struttura di spazio affine. Formalmente si ha la seguente definizione.

Definizione A.15. Siano $\mathbb{A}_1^n$ e $\mathbb{A}_2^m$ spazi affini con spazi delle traslazioni V_1 e V_2 rispettivamente. $\psi : \mathbb{A}_1^n \to \mathbb{A}_2^m$ è detta **trasformazione affine** se valgono le condizioni:

(i) ψ è *invariante per traslazioni*, cioè

$$\psi(P + \mathbf{u}) - \psi(Q + \mathbf{u}) = \psi(P) - \psi(Q) , \text{ per ogni } P, Q \in \mathbb{A}_1^n \text{ e } \mathbf{u} \in V_1 ;$$

(ii) la funzione $P - Q \mapsto \psi(P) - \psi(Q)$ definisce una trasformazione lineare $V_1 \to V_2$, indicata con $d\psi : V_1 \to V_2$.

Esercizi A.4.

1. Sia $(\mathbb{A}^n, f)$ un sistema di coordinate cartesiane, come nella definizione A.14, con coordinate $x^1, \cdots, x^n$ e $(\mathbb{A}^n, g)$ un altro sistema di coordinate cartesiane, con coordinate $x'^1, \cdots, x'^n$, di origine O' e assi $\mathbf{e}'_1, \ldots, e'_n$, in modo che valga

$$\mathbf{e}_i = \sum_j B^j{}_i \mathbf{e}'_j .$$

Provare che la funzione $g \circ f^{-1}$ è espressa, in coordinate, dalle relazioni:

$$x'^j = \sum_{i=1}^n B^j{}_i (x^i + b^i), \tag{A.31}$$

dove $(O - O') = \sum_i b^i e_i$.

2. In riferimento all'esercizio precedente, provare che la funzione $f \circ g^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, in coordinate, è espressa da:

$$x^i = \sum_{j=1}^{n} (B^{-1})^i{}_j x'^j - b^i \,. \tag{A.32}$$

3. Provare che se $\psi : \mathbb{A}_1^n \to \mathbb{A}_2^m$ è affine, allora, per ogni scelta di sistemi di coordinate cartesiane nei due rispettivi spazi, ψ ha la forma $(i = 1, 2, \ldots, n)$

$$x_2^i = \sum_{j=1}^{n} L^i{}_j x_1^j + c^i \,. \tag{A.33}$$

per opportuni coefficienti $L^i{}_j$ e c^i, dipendenti da ψ e dai sistemi di coordinate.
Dimostrare che, viceversa, $\psi : \mathbb{A}_1^n \to \mathbb{A}_2^m$ è affine se esistono due sistemi di coordinate cartesiane nei rispettivi spazi in cui ψ ha la forma (A.33) in coordinate.

4. Mostrare che le trasformazioni affini trasformano rette in rette. Ossia, se $\psi : \mathbb{A}_1^n \to \mathbb{A}_2^n$ è affine e $P(t) := P + t\mathbf{u}$, con $t \in \mathbb{R}$ è la retta in $\mathbb{A}_1^n$ di origine P è vettore tangente $\mathbf{v} \in V_1$, allora $\psi(P(t))$, al variare di $t \in \mathbb{R}$ definisce ancora una retta in $\mathbb{A}^m$.

Per ogni spazio affine $\mathbb{A}^n$, lo spazio vettoriale V delle traslazioni agisce come insieme di trasformazioni $\{T_\mathbf{v}\}_{\mathbf{v}\in V}$ su $\mathbb{A}^n$. La trasformazione $T_\mathbf{v} : \mathbb{A}^n \to \mathbb{A}^n$ di $\mathbf{v} \in V$ su $\mathbb{A}^n$ è definita in modo ovvio come $T_\mathbf{v} : P \mapsto P + \mathbf{v}$. Dal punto di vista fisico, quando $\mathbb{A}^3$ è pensato come l'usuale spazio fisico, le trasformazioni $T_\mathbf{v}$ sono le traslazioni rigide dei corpi fisici. Tornando alla struttura matematica, esplicitiamo alcune caratteristiche dell'azione di V sullo spazio affine.

(i) L'insieme $\{T_\mathbf{v}\}_{\mathbf{v}\in V}$ è banalmente un *gruppo* rispetto alla composizione di applicazioni valendo $T_\mathbf{u} T_\mathbf{v} = T_{\mathbf{u}+\mathbf{v}}$.
Dato che $T_\mathbf{u} T_v = T_{\mathbf{u}+\mathbf{v}}$ e $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$, il gruppo risulta anche essere *abeliano* cioé commutativo: $T_\mathbf{u} T_\mathbf{v} = T_\mathbf{u} T_\mathbf{v}$ per ogni coppia $\mathbf{u}, \mathbf{v} \in V$. Dato che l'applicazione $V \ni v \mapsto T_v$ è iniettiva (poiché $\mathbf{v} = \mathbf{u}$ se $T_\mathbf{u} = T_\mathbf{v}$), essa è un isomorfismo gruppale quando V è visto come gruppo commutativo rispetto alla somma di vettori.

(ii) Solo $\mathbf{v} = \mathbf{0}$ soddisfa che $T_\mathbf{v}(P) = P$ per ogni $P \in \mathbb{A}^n$, in altre parole, l'azione del gruppo delle traslazioni è *libera*.

(iii) Per ogni coppia $P, Q \in \mathbb{A}^n$ esiste una traslazione $T_\mathbf{v}$ tale che $T_\mathbf{v}(P) = Q$, in altre parole, l'azione del gruppo delle traslazioni è *transitiva*.
Il gruppo delle traslazioni acquista ulteriore interesse quando si potenzia la struttura di spazio affine con un prodotto scalare.

A.2.2 Spazi Euclidei.

Definizione A.16. Uno spazio affine $\mathbb{E}^n$ di dimensione n (finita) e dotato di un prodotto scalare (reale simmetrico) $\cdot$ nello spazio delle traslazioni V, è detto **spazio euclideo (reale) di**

dimensione n.
I sistemi di coordinate cartesiane associati a basi ortonormali rispetto a $\cdot$ sono detti sistemi di coordinate cartesiane **ortonormali**.
Una trasformazione affine tra due spazi euclidei che conserva le rispettive distanze è detta **isometria affine**. $\diamondsuit$

Ricordiamo la definizione di *spazio metrico*.

Definizione A.17. Uno **spazio metrico** è un insieme M dotato di una funzione $d : M \times M \to \mathbb{R}$ detta **distanza**, soddisfacente:
(i) $d(P, Q) = d(Q, P)$,
(ii) $d(P, Q) \geq 0$ dove $=$ vale se solo se $P = Q$,
(iii) $d(P, Q) \leq d(P, R) + d(R, Q)$ per $P, Q, R \in M$.
Una funzione $f : M \to M'$ con M, M' spazi metrici con distanze, d e d' rispettivamente è detta **isometria** se $d(P, Q) = d'(f(P), f(Q))$ per ogni coppia $P, Q \in M$. $\diamondsuit$

Ovviamente, per (ii), le isometrie sono sempre trasformazioni iniettive. Nel caso degli spazi euclidei, la presenza del prodotto scalare arricchisce ulteriormente la struttura di spazio affine aggiungendo una struttura di spazio metrico, quando la distanza tra punti di $\mathbb{E}^n$ è definita come la norma standard associata al prodotto scalare $\cdot$ valutata su $P - Q$:

$$d(P, Q) := \sqrt{P - Q \cdot P - Q}\,. \tag{A.34}$$

Quindi gli spazi euclidei sono naturalmente degli spazi metrici. Vale il seguente notevole teorema che identifica isometrie ed isometrie affini. La dimostrazione è data negli esercizi.

Teorema A.4. *Siano $\mathbb{E}^n_1$ e $\mathbb{E}^n_2$ spazi euclidei con la stessa dimensione n (finita). $f : \mathbb{E}^n_1 \to \mathbb{E}^n_2$ è un'isometria se e solo se è un'isometria affine.*

Osservazioni A.8.
(1) La distanza d su uno spazio euclideo (A.34) gode di alcune interessanti proprietà . In primo luogo essa è *invariante sotto l'azione delle traslazioni*:

$$d(P + \mathbf{u}, Q + \mathbf{u}) = d(P, Q)\,, \ \forall\ P, Q \in \mathbb{E}^n,\ \mathbf{u} \in V\,. \tag{A.35}$$

La verifica di tale proprietà è immediata dalla definizione di d e tenendo conto della proprietà $(P + \mathbf{u}) - (Q + \mathbf{u}) = P - Q$.
Un'altra proprietà interessante della distanza di spazi euclidei è di natura completamente matematica: risulta per computo diretto che la funzione $P, Q \mapsto d(P, Q)^2$ è di classe C^∞ rispetto alla struttura differenziabile naturale di $\mathbb{E}^n \times \mathbb{E}^n$. Non lo è invece la funzione $P, Q \mapsto d(P, Q)$ che è continua, ma ammette una singolarità per $P = Q$.
(2) Come conseguenza dell'esercizio A.4.**1**, tenendo conto che le matrici di trasformazione

tra basi ortonormali sono le matrici ortogonali, la più generale legge di trasformazione tra le coordinate di due differenti sistemi di coordinate cartesiani ortonormali assume la forma

$$x'^j = \sum_{i=1}^{n} R^j{}_i (x^i + b^i), \tag{A.36}$$

dove i numeri reali b^i sono arbitrariamente fissati e i coefficienti $R^j{}_i$ definiscono individuano una qualsiasi fissata *matrice ortogonale* di dimensione n. Ricordiamo che le matrici ortogonali di ordine n sono le matrici reali $n \times n$, R, tali che $RR^t = I$ (ossia in componenti $\sum_k R^i{}_k R^j{}_k = \delta^{ij}$). Esse costituiscono un gruppo rispetto al prodotto matriciale righe per colonne.
(3) Si consideri un sistema di coordinate locali curvilinee $y^1, ..., y^n$ nello spazio euclideo $\mathbb{E}^n$ con prodotto scalare $(\cdot|\cdot)$. Possiamo associare a ciascuna coordinata y^i un versore $\mathbf{f}_i(P)$, tangente in P alla curva uscente da P che si ottiene facendo variare solo la coordinata y^i e mantenendo costanti le altre. Tenendo conto dell'identificazione canonica (A.30), vale

$$\mathbf{f}_i(P) = \frac{1}{\sqrt{\left(\frac{\partial}{\partial y^i}|_P \,\middle|\, \frac{\partial}{\partial y^i}|_P\right)}} \frac{\partial}{\partial y^i}|_P \tag{A.37}$$

Esercizi A.5.

1. Sia $\mathbb{E}^n$ uno spazio euclideo e d la sua distanza. Fissato un punto $O \in \mathbb{E}^n$, si identifichino (biunivocamente) i vettori dello spazio V delle traslazioni di $\mathbb{E}^n$ con i punti di $\mathbb{E}^n$ tramite la corrispondenza $\mathbf{u} \mapsto O + u$. Provare che il prodotto scalare $\cdot$ su V si può scrivere in termini di d come:

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2} d\left(O, O + (\mathbf{u} + \mathbf{v})\right)^2 + \frac{1}{2} d\left(O, O + (\mathbf{u} - \mathbf{v})\right)^2 \ . \tag{A.38}$$

2. Siano $\mathbb{E}^n_1$ e $\mathbb{E}^n_2$, due spazi euclidei con la stessa dimensione n, con distanze d_1 e d_2 rispettivamente e prodotti scalari $(\cdot|\cdot)_1$ e $(\cdot|\cdot)_2$ rispettivamente. Sia $\phi : \mathbb{E}^n_1 \to \mathbb{E}^m_2$ una trasformazione affine. Provare che $d\phi$ conserva il prodotto scalare tra vettori (e quindi anche l'angolo tra vettori), ossia, per ogni fissato $Q \in \mathbb{E}^3_1$:

$$(\phi(P) - \phi(Q)|\phi(P') - \phi(Q))_2 = (P - Q|P' - Q)_1 \ , \ \forall \ P, P' \in \mathbb{E}^3_1 \ . \tag{A.39}$$

se e solo se ϕ conserva le distanze, cioé

$$d_2(\phi(P), \phi(Q)) = d_1(P, Q) \ , \ \forall \ P, Q \in \mathbb{E}^3_1 \ . \tag{A.40}$$

3. Siano $\mathbb{E}^n_1$ e $\mathbb{E}^n_2$ due spazi euclidei con la stessa dimensione e con distanze d_1 e d_2 rispettivamente. Mostrare che la trasformazione $\phi : \mathbb{E}^n_1 \to \mathbb{E}^n_2$ è un isometria se e solo se, per una scelta (e quindi ogni scelta) di coordinate cartesiane ortonormali in $\mathbb{E}^n_1$ e $\mathbb{E}^n_2$, è rappresentata nella forma $(i = 1, \ldots, n)$

$$x^i_2 = \sum_{j=1}^{n} R^i{}_j \, x^i_1 + b^i \ . \tag{A.41}$$

essendo la matrice di coefficienti $R^i{}_j$ una matrice ortogonale $n \times n$ reale.

4. Considerando le isometrie tra spazi euclidei, mostrare che:

(i) le isometrie tra due spazi euclidei con la stessa dimensione sono diffeomorfismi rispetto alla struttura differenziabile naturale degli spazi affini;

(ii) le funzioni inverse di isometrie sono ancora isometrie;

(iii) la composizione di due isometrie è ancora un'isometria;

(iv) se $\phi : \mathbb{E}^n_1 \to \mathbb{E}^n_2$ è un isometria, allora $d\phi : V_1 \to V_2$ è un isomorfismo tra spazi vettoriali che conserva il prodotto scalare.

5. Mostrare che le isometrie $\phi : \mathbb{E}^n \to \mathbb{E}^n$ costituiscono un gruppo che è sottogruppo dei diffeomorfismi dallo spazio euclideo $\mathbb{E}^n$ in se stesso.

6.* Dimostrare il teorema A.4.

Appendice B

Soluzioni e/o suggerimenti per risolvere gli esercizi proposti.

B.1 Esercizi del Capitolo 1.

Esercizi 1.1

1.1.**1**. *Soluzione.* Nella situazione in esame, per dimostrare che $\sigma_{\mathscr{I}} : \mathbb{V}^4 \ni p \mapsto (T, x^1, x^2, x^3) \in \mathbb{R}^4$ definisce una carta compatibile con la struttura differenziabile di $\mathbb{V}^4$, per l'esercizio A.1.**5**, è sufficiente esibire, nell'intorno di ogni punto di $\mathbb{V}^4$, una carta locale della struttura differenziabile di $\mathbb{V}^4$, con coordinate locali y^1, y^2, y^3, y^4 tale che la matrice jacobiana della trasformazione $x^i = x^i(y^1, y^2, y^3, y^4)$ abbia determinante ovunque non nullo. Consideriamo $p \in \mathbb{V}^4$ e $\Sigma_t \ni p$. Dato che Σ_t è sottovarietà embedded di $\mathbb{V}^4$, c'è una carta locale (U, ϕ) con $\mathbb{V}^4 \supset U \ni p$ e $\phi : q \mapsto (z^0(q), z^1(q), z^2(q), z^3(q))$ e tale che $U \cap \Sigma_t$ è determinato dalla condizione $z^0 = 0$. Infine $\phi \restriction_{U \cap \Sigma_t}: q \mapsto (z^1(q), z^2(q), z^3(q))$ definisce una carta locale su $U \cap \Sigma_t$. In coordinate z^i la funzione tempo assoluto T deve soddisfare $\partial T/\partial z^0 \neq 0$ (in quanto $\partial T/\partial z^i = 0$ se $i = 1, 2, 3$ per costruzione e T è non singolare). Il determinante della matrice jacobiana, valutato in p, della trasformazione $T = T(z^1, z^2, z^3)$, $z^i = z^i$ (per $i = 1, 2, 3$), vale proprio $\partial T/\partial z^0 \neq 0$. Per il teorema A.1, restringendo il dominio U attorno a p ad un più piccolo intorno U', la funzione $(z^0, z^1, z^2, z^3) \mapsto (T(z^1, z^2, z^3), z^1, z^2, z^3)$ è differenziabile biettiva con inversa differenziabile. In definitiva, per ogni $p \in \mathbb{V}^4$, c'è una carta locale (U', ψ) della struttura differenziabile di $\mathbb{V}^4$, con $\mathbb{V}^4 \supset U' \ni p$ e tale che $\psi : q \mapsto (y^0(q), y^1(q), y^2(q), y^3(q))$ con $(y^0(q), y^1(q), y^2(q), y^3(q)) = (T(q), z^1(q), z^2(q), z^3(q))$. Ovviamente $\Sigma_t \cap U'$ è determinato dalla condizione $y^0 = t$ e y^1, y^2, y^3 sono coordinate locali su Σ_t. Le coordinate x^1, x^2, x^3 sono coordinate globali su Σ_t compatibili con la sua struttura differenziabile in quanto sono definite tramite $\Pi_{\mathscr{I}} \restriction_{\Sigma_t}: \Sigma_t \to \mathbb{E}^3_{\mathscr{I}}$ che è un diffeomorfismo essendo un'isometria tra spazi euclidei. La matrice jacobiana della trasformazione $x^i = x^i(y^1, y^2, y^3)$, con $i = 1, 2, 3$, deve quindi avere determinante J non nullo. Per computo diretto, si verifica che il determinante della matrice jacobiana della trasformazione $T = y^0, x^i = x^i(y^1, y^2, y^3)$, con $i = 1, 2, 3$, ha ancora determinante $J \neq 0$. Questo è quanto volevamo provare. Le prove di (a) e (b) sono allora ovvie in base alle corrispondenti

definizioni.

1.1.**2**. *Soluzione.* Fissiamo un evento $p \in \mathbb{V}^4$ e consideriamone le coordinate secondo $\sigma_{\mathscr{I}}$ e $\sigma_{\mathscr{I}'}$. La prima coordinata $t(p)$ in $(t(p), x^1(p), x^2(p), x^3(p))$ corrisponde al valore di $T(p)$. Cambiando riferimento, $t'(p)$ deve coincidere con $T(p)$ a meno di una costante additiva c arbitraria nella definizione del tempo assoluto T. Questo dimostra l'equazione (1.4). Consideriamo poi un evento $p \in \Sigma_t$. Le coordinate $(x'^1(p), x'^2(p), x'^3(p))$ sono connesse alla coordinate $(x^1(p), x^2(p), x^3(p))$ dalla trasformazione $\Pi_{\mathscr{I}'} \restriction_{\Sigma_t} \circ (\Pi_{\mathscr{I}} \restriction_{\Sigma_t})^{-1}$ che è una isometria in quanto composizione di isometrie (vedi l'esercizioA.5.**4**) per la definizione 1.4. Possiamo applicare l'esercizio A.5.**3** ottenendo proprio (1.4). La differenziabilità delle funzioni $R^i{}_j(t)$ e $c^j(t)$ è conseguenza immediata della condizione (ii) in definizione 1.4 scegliendo opportunamente il punto P che compare in tale condizione: per esempio, scegliendo $P \in \mathbb{E}^3_{\mathscr{I}}$ nell'origine delle coordinate cartesiane non primate, segue subito da (1.4) che le funzioni $c^j(t)$ sono differenziabili, descrivendo nelle coordinate primate la linea di universo di P.

1.1.**3**. *Soluzione.* (i) Rappresentiamo lo spaziotempo come $\mathbb{R}^4$ tramite le coordinate t, x^1, x^2, x^3. Consideriamo quindi la classe di curve etichettate con le terne di numeri $(x'^1, x'^2, x'^3) \in \mathbb{R}^3$ e parametrizzate nel parametro $t' \in \mathbb{R}$ tramite le equazioni:

$$x^j(t') = \sum_{i=1}^{3} (R(t'-c)^{-1})^j{}_i (x^i - c^i(t'-c)) \ , \ \ j = 1, 2, 3 \ . \tag{B.1}$$

Si verifica immediatamente che la classe di queste linee di universo soddisfa la definizione di riferimento 1.4 e, per costruzione, t', x'^1, x'^2, x'^3 sono coordinate cartesiane ortonormali solidali con tale riferimento $\mathscr{I}'$. Per costruzione (1.4) sono le trasformazioni tra le coordinate cartesiane solidali con $\mathscr{I}$ e quelle solidali con $\mathscr{I}'$ definito sopra.
(ii) È evidente che se le trasformazioni (1.4) sono trasformazioni di coordinate interne allo stesso riferimento $\mathscr{I}$, essendo $\Pi_{\mathscr{I}}$ la stessa nei due casi, non può comparire la dipendenza temporale nelle formule (1.4) e $c = 0$ dato che l'origine del tempo è fissata. Se viceversa risulta che le funzioni che appaiono in (1.4) non dipendono dal tempo e $c = 0$, ciò significa che le curve in (B.1) sono le stesse linee di universo del riferimento $\mathscr{I}$ etichettate con le terne $(x'^1, x'^2, x'^3) \in \mathbb{R}^3$, per cui $\mathscr{I} = \mathscr{I}'$.

Esercizi 1.2

1.2.1. *Soluzione.* Per prima cosa esprimiamo i versori delle coordinate polari in funzione di quelli delle coordinate cartesiane e viceversa. Vale

$$\mathbf{e}_r = \cos\phi\, \mathbf{e}_x + \sin\phi\, \mathbf{e}_y \ , \tag{B.2}$$

$$\mathbf{e}_\varphi = -\sin\phi\, \mathbf{e}_x + \cos\phi\, \mathbf{e}_y \ , \tag{B.3}$$

con inversa

$$\mathbf{e}_x = \cos\phi\, \mathbf{e}_r - \sin\phi\, \mathbf{e}_\varphi \ , \tag{B.4}$$

$$\mathbf{e}_y = \sin\phi\, \mathbf{e}_r + \cos\phi\, \mathbf{e}_\varphi \ . \tag{B.5}$$

Scelta un'origine O, avremo che, per definizione di $\mathbf{e}_r$

$$\mathbf{x}(t) = P(t) - O = x(t)\,\mathbf{e}_x + y(t)\,\mathbf{e}_y = r(t)\,\mathbf{e}_r(t)\,, \tag{B.6}$$

dove abbiamo esplicitato la dipendenza dal tempo di $\mathbf{e}_r$ tenendo conto che r e ϕ dipendono dal tempo quando il punto si muove. La derivata temporale di $P(t) - O$ si deve eseguire tenendo conto di tale dipendenza. Nel seguito indicheremo come di consuetudine (risalente a Newton), la derivata temporale con un punto sopra la variabile derivata. Abbiamo che

$$\mathbf{v}(t) = \dot{\mathbf{x}}(t) = \dot{r}(t)\,\mathbf{e}_r(t) + r(t)\,\dot{\mathbf{e}}_r(t)\,. \tag{B.7}$$

Assumendo la dipendenza temporale di $\mathbf{e}_r$ e $\mathbf{e}_\varphi$ in (B.2)-(B.3) e derivando in t e quindi usando (B.4)-(B.5), si ottiene

$$\dot{\mathbf{e}}_r = \dot{\phi}\,\mathbf{e}_\varphi\,, \tag{B.8}$$
$$\dot{\mathbf{e}}_\varphi = -\dot{\phi}\,\mathbf{e}_r\,, \tag{B.9}$$

e quindi, per esempio,

$$\ddot{\mathbf{e}}_r = \ddot{\phi}\,\mathbf{e}_\varphi - \dot{\phi}^2\,\mathbf{e}_r\,,$$
$$\ddot{\mathbf{e}}_\varphi = -\dot{\phi}^2\,\mathbf{e}_\varphi - \ddot{\phi}\,\mathbf{e}_r\,.$$

Inserendo (B.8)-(B.9) in (B.7) si trova

$$\mathbf{v} = \dot{r}\,\mathbf{e}_r + r\dot{\phi}\,\mathbf{e}_\varphi\,.$$

Derivando nel tempo ed usando (B.8)-(B.9), si trova l'espressione per l'accelerazione

$$\mathbf{a} = (\ddot{r} - r\dot{\phi}^2)\,\mathbf{e}_r + (r\ddot{\phi} + 2\dot{r}\dot{\phi})\,\mathbf{e}_\varphi\,.$$

1.2.2. *Soluzione.* Si procede come per l'esercizio precedente. Abbiamo che

$$\mathbf{x}(t) = P(t) - O = x(t)\,\mathbf{e}_x + y(t)\,\mathbf{e}_y + z(t)\,\mathbf{e}_z = r(t)\,\mathbf{e}_r(t)\,. \tag{B.10}$$

Inoltre

$$\mathbf{e}_r = \sin\theta\cos\phi\,\mathbf{e}_x + \sin\theta\sin\phi\,\mathbf{e}_y + \cos\theta\,\mathbf{e}_z\,, \tag{B.11}$$
$$\mathbf{e}_\theta = \cos\theta\cos\phi\,\mathbf{e}_x + \cos\theta\sin\phi\,\mathbf{e}_y - \sin\theta\,\mathbf{e}_z\,, \tag{B.12}$$
$$\mathbf{e}_\varphi = -\sin\phi\,\mathbf{e}_x + \cos\phi\,\mathbf{e}_y\,. \tag{B.13}$$

Si osservi che la terna è ortonormale, per cui dati due versori il terzo si ottiene con il prodotto vettoriale dei primi due facendo attenzione al segno. Al fine invertire la trasformazione di sopra, si vede subito da un disegno che

$$\mathbf{e}_z = \cos\theta\,\mathbf{e}_r - \sin\theta\,\mathbf{e}_\theta\,.$$

Risulta comodo introdurre il versore radiale uscente nel piano $z = 0$:

$$n := \sin\theta\, \mathbf{e}_r + \cos\theta\, \mathbf{e}_\theta \ .$$

Per ricavare $\mathbf{e}_x$ e $\mathbf{e}_y$ si procede usando le stesse formule che per le coordinate polari piane e rimpiazzando $\mathbf{e}_r$ con n:

$$\begin{aligned} \mathbf{e}_x &= \cos\phi\, n - \sin\phi\, \mathbf{e}_\varphi \ , \\ \mathbf{e}_y &= \sin\phi\, n + \cos\phi\, \mathbf{e}_\varphi \ . \end{aligned}$$

Mettendo tutto insieme otteniamo:

$$\mathbf{e}_x = \sin\theta\cos\phi\, \mathbf{e}_r - \cos\theta\cos\phi\, \mathbf{e}_\theta - \sin\phi\, \mathbf{e}_\varphi \ , \tag{B.14}$$

$$\mathbf{e}_y = \sin\theta\sin\phi\, \mathbf{e}_r + \cos\theta\sin\phi\, \mathbf{e}_\theta + \cos\phi\, \mathbf{e}_\varphi \ , \tag{B.15}$$

$$\mathbf{e}_z = \cos\theta\, \mathbf{e}_r - \sin\theta\, \mathbf{e}_\theta \ . \tag{B.16}$$

Lo stesso risultato si ottiene notando che entrambe le terne sono ortonormali e pertanto sono connesse da una trasformazione ortogonale. Questo significa che la matrice che rappresenta (B.14)-(B.16) è la trasposta della matrice che rappresenta (B.11)-(B.13).
Derivando rispetto a t le (B.11)-(B.13) e quindi usando (B.14)-(B.16) nel risultato, si ottiene dopo un laborioso calcolo:

$$\dot{\mathbf{e}}_r = \dot\theta\, \mathbf{e}_\theta + \dot\phi\sin\theta\, \mathbf{e}_\varphi \ , \tag{B.17}$$

$$\dot{\mathbf{e}}_\theta = -\dot\theta\, \mathbf{e}_r + \dot\phi\cos\theta\, \mathbf{e}_\varphi \ , \tag{B.18}$$

$$\dot{\mathbf{e}}_\varphi = -\dot\phi\sin\theta\, \mathbf{e}_r - \dot\phi\cos\theta\, \mathbf{e}_\theta \ . \tag{B.19}$$

Derivando in t (B.10) ed usando (B.17) si ottiene infine

$$\mathbf{v} = \dot r\, \mathbf{e}_r + r\dot\theta\, \mathbf{e}_\theta + r\dot\phi\sin\theta\, \mathbf{e}_\varphi \ .$$

Il calcolo dell'accelerazione produce, derivando l'espressione della velocità ed usando (B.17)-(B.19):

$$\mathbf{a} = (\ddot r - r\dot\theta^2 - r\dot\phi^2\sin^2\theta)\, \mathbf{e}_r + (r\ddot\theta + 2\dot r\dot\theta - r\dot\phi^2\sin\theta\cos\theta)\, \mathbf{e}_\theta + (r\ddot\phi\sin\theta + 2\dot r\dot\phi\sin\theta + 2r\dot\phi\dot\theta\cos\theta)\, \mathbf{e}_\varphi \ .$$

1.2.3. *Traccia di soluzione.* In realtà l'esercizio è essenzialmente già stato risolto come esercizio 1.2.1, visto che l'asse z non introduce nessuna difficoltà essendo $\mathbf{e}_z$ indipendente da t e valendo:

$$\mathbf{x}(t) = P(t) - O = x(t)\, \mathbf{e}_x + y(t)\, \mathbf{e}_y + z(t)\, \mathbf{e}_z = r(t)\, \mathbf{e}_r(t) + z(t)\, \mathbf{e}_z \ .$$

1.2.4. La spiegazione si basa sul fatto che, in coordinate cartesiane ortonormali, la connessione di Levi-Civita ha coefficienti di connessione nulli per cui la (1.13) fornisce automaticamente l'accelerazione. Si conclude che l'accelerazione del punto materiale altro non è che la derivata covariante seconda lungo la traiettoria del moto:

$$\mathbf{a} = \nabla_{\dot P}\dot P \ .$$

L'espressione, in coordinate del tutto arbitrarie, dell'identità di sopra è proprio la (1.13). Ovviamente, quindi, il risultato vale non solo in coordinate cilindriche o sferiche, ma in coordinate curvilinee arbitrarie.

Esercizi 1.2

1.3.**1**. *Soluzione.* Prendiamo un sistema di coordinate cartesiane ortonormali in cui Π risulti essere il piano $z = 0$. In tal caso le equazioni della curva sono date da, per ipotesi, $x = x(s)$, $y = y(s)$, $z = 0$. Di conseguenza, $\mathbf{t}(s) = dx/ds\ \mathbf{e}_x + dy/ds\ \mathbf{e}_y$. Per ipotesi $\mathbf{n}$ esiste (Γ non è riparametrizzabile come un segmento di retta in ogni sottointervallo di (a, b)) e, per costruzione, $\mathbf{n}$ appartiene ancora al piano Π essendo scritto come combinazione lineare di $\mathbf{e}_x$ e $\mathbf{e}_y$. Di conseguenza $\mathbf{b}(s) = \mathbf{t}(t) \wedge \mathbf{n}(s) = \mathbf{e}_z$ che è perpendicolare a Π ed è costante.

1.3.**2**. *Soluzione.* Sia $P = P(s)$ la curva con $s \in (a, b) \ni c$. Consideriamo la funzione di s, $f(s) := \mathbf{b} \cdot (P(s) - P(c))$. Essendo $\mathbf{b}$ costante in s, la derivata di essa vale $f'(s) = \mathbf{b} \cdot \mathbf{t}(s) = 0$. Quindi a funzione è costante. In realtà è sempre nulla essendo $f(c) = 0$. Concludiamo che $\mathbf{b} \cdot (P(s) - P(c)) = 0$ per ogni $s \in (a, b)$, ma l'insieme $\{P \in \mathbb{E}^3 \,|\, \mathbf{b} \cdot (P - P(c))\} = 0$ è l'equazione del piano passante per $P(c)$ e normale a $\mathbf{b} \neq 0$. Concludiamo che Γ appartiene a tale piano.

1.3.**3**. *Soluzione.* Prima di tutto notiamo che le ultime identità in ciascuna delle tre equazioni si ottengono subito usando la definizione di vettore di Darboux ed eseguendo i prodotti vettoriali e ricordando che $\mathbf{t}, \mathbf{n}, \mathbf{b}$ è una base ortonormale destrorsa. Concentriamoci allora sulle prime identità delle tre equazioni. La (1.29) è la definizione di versore $\mathbf{n}$ e non deve essere provata. Proviamo (1.27). Dato che $\mathbf{t}, \mathbf{n}, \mathbf{b}$ è una base ortonormale, varrà

$$d\mathbf{n}/ds = (d\mathbf{n}/ds \cdot \mathbf{t})\mathbf{t} + (d\mathbf{n}/ds \cdot \mathbf{n})\mathbf{n} + (d\mathbf{n}/ds \cdot \mathbf{b})\mathbf{b}\,.$$

Calcoliamo le componenti separatamente.

$$d\mathbf{n}/ds \cdot \mathbf{t} = d/ds(\mathbf{n} \cdot \mathbf{t}) - \mathbf{n} \cdot d\mathbf{t}/ds = 0 - \rho^{-1}(\mathbf{n} \cdot \mathbf{n}) = -\rho^{-1}\,.$$

Per la componente su $\mathbf{n}$ si ha:

$$d\mathbf{n}/ds \cdot \mathbf{n} = (1/2)d/ds\mathbf{n} \cdot \mathbf{n} = (1/2)d/ds1 = 0\,.$$

Per la componente su $\mathbf{b}$ si ha:

$$d\mathbf{n}/ds \cdot \mathbf{b} = d\mathbf{n}/ds \cdot (\mathbf{t} \wedge \mathbf{b}) =: -\tau^{-1} \quad \text{per definizione.}$$

Mettendo insieme si ottiene (1.27). Passiamo alla (1.28). Dato che $\mathbf{t}, \mathbf{n}, \mathbf{b}$ è una base ortonormale, varrà

$$d\mathbf{b}/ds = (d\mathbf{b}/ds \cdot \mathbf{t})\mathbf{t} + (d\mathbf{b}/ds \cdot \mathbf{n})\mathbf{n} + (d\mathbf{b}/ds \cdot \mathbf{b})\mathbf{b}\,.$$

Calcoliamo le componenti separatamente.

$$d\mathbf{b}/ds \cdot \mathbf{t} = d/ds(\mathbf{b} \cdot \mathbf{t}) - \mathbf{b} \cdot d\mathbf{t}/ds = 0 - \rho^{-1}(\mathbf{b} \cdot \mathbf{n}) = 0\,.$$

Per la componente su $\mathbf{b}$ si ha:

$$d\mathbf{b}/ds \cdot \mathbf{b} = (1/2)d/ds(\mathbf{b} \cdot \mathbf{b}) = (1/2)d/ds1 = 0\,.$$

Per la componente su $\mathbf{n}$ si ha:

$$d\mathbf{b}/ds \cdot \mathbf{n} = 0 - \mathbf{b} \cdot d\mathbf{n}/ds = -\mathbf{b} \cdot (-\mathbf{t}/\rho - \mathbf{b}/\tau) = (\mathbf{b} \cdot \mathbf{b})/\tau = 1/\tau \, .$$

Mettendo insieme si ottiene (1.28).

Esercizi 1.4

1.4.**1**.*Soluzione.* Esprimiamo i versori della base di $\mathscr{I}$ in funzione di quelli della base di $\hat{\mathscr{I}}$. Per costruzione $\mathbf{e}_3 \cdot \hat{\mathbf{e}}_3 = \cos\theta$, $\mathbf{e}_3 \cdot \hat{\mathbf{e}}_1 = \sin\theta \, \sin\phi$ e $\mathbf{e}_3 \cdot \hat{\mathbf{e}}_2 = -\sin\theta \, \cos\psi$, quindi

$$\mathbf{e}_3 = \sin\theta \, \sin\psi \, \hat{\mathbf{e}}_1 - \sin\theta \, \cos\psi \, \hat{\mathbf{e}}_2 + \cos\theta \, \hat{\mathbf{e}}_3 \, .$$

Passiamo a $\mathbf{e}_1$. Prima di tutto decomponiamolo sulla base $\mathbf{N}$, $\mathbf{e}_3, \mathbf{N}'$ dove $\mathbf{N}' := \mathbf{e}_3 \wedge \mathbf{N}$. Per costruzione $\mathbf{e}_1 = \cos\phi \, \mathbf{N} + \sin\phi \, \mathbf{N}'$ e pertanto, essendo $\mathbf{N} = \cos\psi \, \hat{\mathbf{e}}_1 + \sin\psi \, \hat{\mathbf{e}}_2$, vale

$$\mathbf{e}_1 = \cos\phi \, (\cos\psi \, \hat{\mathbf{e}}_1 + \sin\psi \, \hat{\mathbf{e}}_2) + \sin\phi \, \mathbf{e}_3 \wedge (\cos\psi \, \hat{\mathbf{e}}_1 + \sin\psi \, \hat{\mathbf{e}}_2)$$

Ossia:

$$\mathbf{e}_1 = \cos\phi(\cos\psi \, \hat{\mathbf{e}}_1 + \sin\psi \, \hat{\mathbf{e}}_2) + \sin\phi(\sin\theta \, \sin\psi \, \hat{\mathbf{e}}_1 - \sin\theta \, \cos\psi \hat{\mathbf{e}}_2 + \cos\theta \, \hat{\mathbf{e}}_3) \wedge (\cos\psi \, \hat{\mathbf{e}}_1 + \sin\psi \, \hat{\mathbf{e}}_2) \, ,$$

che, calcolato esplicitamente fornisce:

$$\mathbf{e}_1 = (\cos\phi \cos\psi - \sin\phi \sin\psi \cos\theta) \, \hat{\mathbf{e}}_1 + (\cos\phi \sin\psi + \sin\phi \cos\theta \cos\psi) \, \hat{\mathbf{e}}_2 + \sin\phi \sin\theta \, \hat{\mathbf{e}}_3 \, .$$

L'espressione di $\mathbf{e}_2$ si ottiene come prodotto vettoriale $\mathbf{e}_3 \wedge \mathbf{e}_1$ e produce

$$\mathbf{e}_2 = -(\sin\phi \cos\psi + \cos\phi \sin\psi \cos\theta) \, \hat{\mathbf{e}}_1 + (-\sin\phi \sin\psi + \cos\phi \cos\theta \cos\psi) \, \hat{\mathbf{e}}_2 + \cos\phi \sin\theta \, \hat{\mathbf{e}}_3$$

1.4.**2**. *Traccia di soluzione.* Dato che le matrici di trasformazione sono ortogonali, è sufficiente prendere la matrice trasposta dalla matrice di trasformazione dell'esercizio precedente.

1.4.**3**. *Traccia di soluzione.* Usando la legge di decomposizione di $\boldsymbol{\omega}$, $\boldsymbol{\omega}_{\hat{\mathscr{I}}}|_{\mathscr{I}}$, si ottiene come somma di 3 vettori $\boldsymbol{\omega}$. Sia $\mathscr{I}_2$ il sistema di riferimento con terna solidale data da $\hat{\mathbf{e}}_3, \mathbf{N}$, $\hat{\mathbf{e}}_3 \wedge \mathbf{N}$ e $\mathscr{I}_3$ quello con terna solidale data da $\mathbf{e}_3, \mathbf{N}$, $\mathbf{e}_3 \wedge \mathbf{N}$. Abbiamo allora che:

$$\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = \boldsymbol{\omega}_{\mathscr{I}}|_{\mathscr{I}_3} + \boldsymbol{\omega}_{\mathscr{I}_3}|_{\mathscr{I}_2} + \boldsymbol{\omega}_{\mathscr{I}_2}|_{\hat{\mathscr{I}}} \, ,$$

ossia, per definizione di angoli di Eulero:

$$\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} = \dot{\phi} \, \mathbf{e}_3 + \dot{\theta} \mathbf{N} + \dot{\psi} \, \hat{\mathbf{e}}_3 \, ,$$

dove $\mathbf{N} = \cos\psi \, \hat{\mathbf{e}}_1 + \sin\psi \, \hat{\mathbf{e}}_2$. Esprimendo i versori $\hat{\mathbf{e}}_i$ in funzione dei versori $\mathbf{e}_k$ usando le formule dell'esercizio precedente, si ricava subito la formula finale per $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}}$.

B.2 Esercizi del Capitolo 2.

Esercizi 2.1

2.1.**1**. *Traccia di soluzione.* Comporre le trasformazioni di Galileo associate a $(c_2, \vec{c_2}, \vec{v_2}, R_2)$ e $(c_1, \vec{c_1}, \vec{v_1}, R_1)$, nell'ordine detto, e ricavare gli analoghi coefficienti per la trasformazione di Galileo ottenuta.

2.1.**2**. *Traccia di soluzione.* Bisogna verificare che la legge di composizione detta sia associativa e che $(0, \vec{0}, \vec{0}, I)$

$$(c, \vec{c}, \vec{v}, R)^{-1} = \left(-c,\ -R^{-1}(\vec{c} + c\vec{v}),\ -R^{-1}\vec{v},\ R^{-1}\right)\ .$$

siano l'elemento neutro ed inverso rispetto a tale legge di composizione.

2.1.**3**. *Traccia di soluzione.* Si identifichino i vettori $(t, \vec{x}) \in \mathbb{R}^4$ con i vettori di $\mathbb{R}^5$ della forma $(1, t, \vec{x})$. Si associ quindi a ciascun elemento $G := (c, \vec{c}, \vec{v}, R)$ del gruppo di Galileo la matrice

$$\Omega_G := \left[\begin{array}{c|c|c} 1 & 0 & \vec{0} \\ \hline c & 1 & \vec{0} \\ \hline \vec{c} & \vec{v} & R \end{array}\right]\ .$$

Si verifichi che l'azione di Ω_G sul vettore (pensato come vettore colonna) $(1, t, \vec{x})$, Ω_G produca l'azione nota del elemento G del gruppo di Galileo su t e $\vec{x}$, lasciando immutato il primo numero 1. Si verifichi che la corrispondenza $G \mapsto \Omega_G$ sia iniettiva. Si verifichi Ω_G sia non singolare e che la legge di composizione di elementi G, G' del gruppo di Galileo corrisponda al prodotto matriciale delle matrici $\Omega_G, \Omega_{G'}$ rispettivamente associate e che l'elemento neutro del gruppo di Galileo sia associato alla matrice identità .

Esercizi 2.2

2.2.**1**. *Soluzione.* Siano $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ assi in quiete con la terra uscenti dal punto O, sia $\hat{O}$ solidale con il sedile e si supponga che gli assi solidali con il sedile $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$, $\hat{\mathbf{e}}_z$ siano uscenti da $\hat{O}$. Gli assi $\mathbf{e}_z$, $\hat{\mathbf{e}}_z$ si suppongono entrambi verticali e la ruota si assume trovarsi nel piano $y = \hat{y} = 0$. Si osservi che, per costruzione gli assi $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_z$ rimangono paralleli agli assi $\mathbf{e}_x$, $\mathbf{e}_z$ durante il moto del sedile. Pertanto $\boldsymbol{\omega}_{\hat{\mathscr{I}}}|_{\mathscr{I}} = 0$. Tuttavia $\hat{\mathscr{I}}$ non è inerziale. Supponiamo che sia $\alpha \neq 0$ la velocità angolare (costante) della ruota, posizionando O al centro della grande ruota, la trasformazione di coordinate vale (assumiamo $\hat{t} = t$): $x = R\cos(\alpha t) + \hat{x}$, $y = \hat{y}$, $z = R\sin(\alpha t) + \hat{z}$. Questa trasformazione non appartiene al gruppo di Galileo per cui $\hat{\mathscr{I}}$ non può essere inerziale se lo è $\mathscr{I}$.

Esercizi 2.3

2.3.**1**. *Traccia di soluzione.* Il secondo principio della dinamica per il sistema P, Q si scrive

$$m\mathbf{a}_P = -mg\,\mathbf{e}_z - \kappa(P - Q) + \boldsymbol{\phi}_P$$

$$m\mathbf{a}_Q = -mg\,\mathbf{e}_z - \kappa(Q - P) + \boldsymbol{\phi}_Q$$

Parametrizando Γ con l'ascissa curvilinea, determinando P con l'ascissa s_P e Q con l'ascissa s_Q, ed usando le formule dell'esempio **1** in esempi 1.1, le equazioni di sopra si riscrivono:

$$m\mathbf{a}_P = -mg\,\mathbf{e}_z - \kappa\left(\cos\left(\frac{s_P}{\sqrt{2}}\right) - \cos\left(\frac{s_Q}{\sqrt{2}}\right)\right)\mathbf{e}_x - \kappa\left(\sin\left(\frac{s_P}{\sqrt{2}}\right) - \sin\left(\frac{s_Q}{\sqrt{2}}\right)\right)\mathbf{e}_y$$

$$-\kappa\frac{s_P - s_Q}{\sqrt{2}}\,\mathbf{e}_z + \boldsymbol{\phi}_P$$

e

$$m\mathbf{a}_Q = -mg\,\mathbf{e}_z - \kappa\left(\cos\left(\frac{s_Q}{\sqrt{2}}\right) - \cos\left(\frac{s_P}{\sqrt{2}}\right)\right)\mathbf{e}_x - \kappa\left(\sin\left(\frac{s_Q}{\sqrt{2}}\right) - \sin\left(\frac{s_P}{\sqrt{2}}\right)\right)\mathbf{e}_y$$

$$-\kappa\frac{s_Q - s_P}{\sqrt{2}}\,\mathbf{e}_z + \boldsymbol{\phi}_Q\,.$$

A questo punto, usando le formule dell'esempio **1** in esempi 1.1, proiettando entrambe le equazioni lungo i vettori tangenti a Γ in P e Q, rispettivamente $\mathbf{t}_P$ e $\mathbf{t}_P$, e tenendo conto di (1.20) si trova il sistema di equazioni pure di movimento ($\boldsymbol{\phi}_P$ e $\boldsymbol{\phi}_Q$ sono normali a Γ per ipotesi e quindi non appaiono):

$$m\frac{d^2 s_P}{dt^2} = -\frac{mg}{\sqrt{2}} + \frac{\kappa}{\sqrt{2}}\sin\left(\frac{s_P}{\sqrt{2}}\right)\left[\cos\left(\frac{s_P}{\sqrt{2}}\right) - \cos\left(\frac{s_Q}{\sqrt{2}}\right)\right]$$

$$-\frac{\kappa}{\sqrt{2}}\cos\left(\frac{s_P}{\sqrt{2}}\right)\left[\sin\left(\frac{s_P}{\sqrt{2}}\right) - \sin\left(\frac{s_Q}{\sqrt{2}}\right)\right] - \frac{\kappa}{\sqrt{2}}(s_P - s_Q)$$

e

$$m\frac{d^2 s_Q}{dt^2} = -\frac{mg}{\sqrt{2}} + \frac{\kappa}{\sqrt{2}}\sin\left(\frac{s_Q}{\sqrt{2}}\right)\left[\cos\left(\frac{s_Q}{\sqrt{2}}\right) - \cos\left(\frac{s_P}{\sqrt{2}}\right)\right]$$

$$-\frac{\kappa}{\sqrt{2}}\cos\left(\frac{s_Q}{\sqrt{2}}\right)\left[\sin\left(\frac{s_Q}{\sqrt{2}}\right) - \sin\left(\frac{s_P}{\sqrt{2}}\right)\right] - \frac{\kappa}{\sqrt{2}}(s_Q - s_P)$$

2.3.**2**. *Traccia di soluzione.* Si calcoli la derivata temporale del secondo membro di E e si tenga conto dell'equazione (2.50). Il significato di E è quello dell'energia meccanica del punto materiale nel riferimento non inerziale $\mathscr{I}'$.

B.3 Esercizi del Capitolo 3.

Esercizi 3.1

3.1.**1**. *Suggerimento.* È sufficiente provare l'asserto per $k = \pm 1$ e poi iterare la dimostrazione k volte. Si applichi il teorema di esistenza ed unicità in forma globale usando il fatto che ($k = 1$)

$$\left.\frac{d\mathbf{x}'}{dt}\right|_{t=s-T} = \left.\frac{d\mathbf{x}}{dt}\right|_{t=s}$$

e che $\mathbf{f}$ è periodica in t di periodo T.

3.1.**2**. *Traccia di Soluzione.* Basta usare l'esercizio precedente per $k = 1$, notando che la funzione $\mathbf{f}$, essendo costante in t è periodica di qualsiasi periodo T in tale variabile.

3.1.**3**. *Soluzione.* Ignoriamo già dal principio il moto nella direzione z, che è di quiete a causa del vincolo. La reazione vincolare del piano vale quindi $mg\,\mathbf{e}_z$. Il problema si riduce ad un problema bidimensionale con equazioni:

$$\begin{aligned} \frac{d^2x}{dt^2} &= m^{-1}\kappa x\,, \\ \frac{d^2y}{dt^2} &= m^{-1}\kappa y\,. \end{aligned}$$

Si tratta di un sistema del secondo ordine in forma normale. Riducendolo al prim'ordine si ha che:

$$\begin{aligned} \frac{d^2v_x}{dt^2} &= m^{-1}\kappa x\,, \\ \frac{d^2v_y}{dt^2} &= m^{-1}\kappa y\,, \\ \frac{d^2x}{dt^2} &= v_x\,, \\ \frac{d^2y}{dt^2} &= v_y\,. \end{aligned} \tag{B.20}$$

È chiaro che il secondo membro è una funzione di classe C^1 e pertanto possiamo applicare i teoremi di esistenza ed unicità . Tornando alla forma del secondo ordine del sistema, notiamo che le due equazioni sono disaccoppiate, per cui le possiamo risolvere separatamente. Pensiamo dapprima le soluzioni $x = x(t)$ e $y = y(t)$ come soluzioni complesse, alla fine ci ridurremo al caso reale. Il polinomio caratteristico di entrambe è : $\chi^2 - \frac{\kappa}{m} = 0$. Le soluzioni sono:

$$b_\pm = \pm\sqrt{\frac{\kappa}{m}}\,, \quad \text{se } \kappa > 0\,,$$

oppure

$$b_\pm = \pm i\sqrt{\frac{-\kappa}{m}}\,, \quad \text{se } \kappa < 0.$$

Nel caso $\kappa > 0$, la legge oraria è dunque del tipo:

$$x(t) = c_1 e^{\sqrt{\frac{\kappa}{m}}t} + c_2 e^{-\sqrt{\frac{\kappa}{m}}t}\,, \tag{B.21}$$

$$y(t) = d_1 e^{\sqrt{\frac{\kappa}{m}}t} + d_2 e^{-\sqrt{\frac{\kappa}{m}}t}\,, \tag{B.22}$$

dove c_1, c_2, d_1, d_2 sono costanti arbitrarie in $\mathbb{R}$. Esse vengono fissate una volta assegnate condizioni iniziali.

Nel caso $\kappa < 0$ la discussione è leggermente più complessa. Lavorando in $\mathbb{C}$, le soluzioni sarebbero:

$$\begin{aligned} x(t) &= c_1 e^{i\sqrt{\frac{\kappa}{m}}t} + c_2 e^{-i\sqrt{\frac{\kappa}{m}}t} \,, \\ y(t) &= d_1 e^{i\sqrt{\frac{\kappa}{m}}t} + d_2 e^{-i\sqrt{\frac{\kappa}{m}}t} \,. \end{aligned}$$

In generale, queste soluzioni non sono accettabili, in quanto x e y risultano essere complessi. Ricordiamo le formule di Eulero: $e^{i\alpha} = \cos\alpha + i\sin\alpha$. Da ciò si ha che le soluzioni di sopra si possono riscrivere come:

$$x(t) = (c_1 + c_2)\cos\left(\sqrt{\frac{\kappa}{m}}t\right) + i(c_1 - c_2)\sin\left(\sqrt{\frac{\kappa}{m}}t\right) \,, \tag{B.23}$$

$$y(t) = (d_1 + d_2)\cos\left(\sqrt{\frac{\kappa}{m}}t\right) + i(d_1 - d_2)\sin\left(\sqrt{\frac{\kappa}{m}}t\right) \,. \tag{B.24}$$

Tutte le soluzioni reali si ottengono facendo variare le costanti $(c_1 + c_2), (d_1 + d_2)$ in $\mathbb{R}$ e le costanti $(c_1 - c_2), (d_1 - d_2)$ in $i\mathbb{R}$ (insieme degli immaginari). In definitiva, la soluzione generale ha la forma:

$$x(t) = A_1\cos\left(\sqrt{\frac{\kappa}{m}}t\right) + A_2\sin\left(\sqrt{\frac{\kappa}{m}}t\right) \,, \tag{B.25}$$

$$y(t) = B_1\cos\left(\sqrt{\frac{\kappa}{m}}t\right) + B_2\sin\left(\sqrt{\frac{\kappa}{m}}t\right) \,. \tag{B.26}$$

con $A_1, A_2, B_1, B_2 \in \mathbb{R}$.
Per concludere ricaviamo la legge oraria, nel caso $\kappa < 0$, con condizioni iniziali $\mathbf{x}(0) = \mathbf{0}$ e $\mathbf{v}(0) = v(\,\mathbf{e}_x - \,\mathbf{e}_y)$. La prima condizione, inserita in (B.25)-(B.26), fornisce: $0 = A_1$ e $0 = B_1$. Derivando (B.25)-(B.26), tenendo conto di quanto appena ottenuto, si ha:

$$v_x(0) = \sqrt{\frac{\kappa}{m}}A_2 \,, \tag{B.27}$$

$$v_y(0) = \sqrt{\frac{\kappa}{m}}B_2 \,. \tag{B.28}$$

Dovendo essere $\mathbf{v}(0) = v(\,\mathbf{e}_x - \,\mathbf{e}_y)$, ricaviamo che $v_x(0) = v$ e $v_y(0) = -v$, per cui $A_2 = v\sqrt{\frac{m}{\kappa}}$ e $B_2 = -v\sqrt{\frac{m}{\kappa}}$. La legge oraria richiesta è dunque:

$$\mathbf{x}(t) = v\sqrt{\frac{m}{\kappa}}\sin\left(\sqrt{\frac{\kappa}{m}}t\right)(\,\mathbf{e}_x - \,\mathbf{e}_y) \,, \quad \text{per } t \in \mathbb{R}.$$

3.1.**4**. *Soluzione.* Procedendo come per l'esercizio precedente, si ha il sistema:

$$\begin{aligned} \frac{d^2x}{dt^2} &= m^{-1}\kappa x - m^{-1}\beta\frac{dy}{dt} \,, \\ \frac{d^2y}{dt^2} &= m^{-1}\kappa y + m^{-1}\beta\frac{dx}{dt} \,. \end{aligned}$$

È chiaro che siamo nelle ipotesi del teorema di esistenza ed unicità . Le equazioni, a differenza dell'esercizio precedente non sono disaccoppiate. Conviene procedere nel modo seguente. Definiamo la variabile complessa $z = x + iy$ e riscriviamo le equazioni di sopra come:

$$\begin{aligned} \frac{d^2x}{dt^2} &= m^{-1}\kappa x + m^{-1}\beta i \frac{diy}{dt} \,, \\ \frac{d^2iy}{dt^2} &= m^{-1}\kappa iy + m^{-1}\beta i \frac{dx}{dt} \,. \end{aligned}$$

Sommando membro a membro otteniamo una sola equazione complessa che contiene entrambe le equazioni reali:

$$\frac{d^2z}{dt^2} - i\frac{\beta}{m}\frac{dz}{dt} - \frac{\kappa}{m}z = 0 \,. \tag{B.29}$$

Dobbiamo ricordare, alla fine, che le soluzioni in x e y si otterranno prendendo la parte reale ed immaginaria della soluzione z.
Il polinomio caratteristico è :

$$\chi^2 - i\frac{\beta}{m}\chi - \frac{\kappa}{m} = 0 \,,$$

ed ha soluzioni:

$$b_\pm = i\frac{\beta}{2m} \pm \frac{1}{2}\sqrt{\frac{4\kappa}{m} - \frac{\beta^2}{m^2}} \,.$$

Esaminiamo subito il caso di soluzioni coincidenti, cioé $\frac{4\kappa}{m} = \frac{\beta^2}{m^2}$. In questo caso, la soluzione generale ha la forma, con $a, b \in \mathbb{C}$ costanti arbitrarie:

$$z(t) = (a + bt)e^{i\frac{\beta}{2m}t} \,.$$

Usando la formula di Eulero si ha ancora che:

$$z(t) = (a + bt)\left(\cos\left(\frac{\beta}{2m}t\right) + i\sin\left(\frac{\beta}{2m}t\right)\right) \,.$$

Decomponendo $a = A + iB$ e $b = C + iD$ con $A, B, C, D \in \mathbb{R}$, e prendendo la parte reale ed immaginaria otteniamo le espressioni per la soluzione generale reale:

$$x(t) = (A + Ct)\cos\left(\frac{\beta}{2m}t\right) - (B + Dt)\sin\left(\frac{\beta}{2m}t\right)$$

e

$$y(t) = (B + Dt)\cos\left(\frac{\beta}{2m}t\right) + (A + Ct)\sin\left(\frac{\beta}{2m}t\right) \,.$$

Nel caso $\frac{4\kappa}{m} \neq \frac{\beta^2}{m^2}$, la soluzione generale, in $\mathbb{C}$, dell'equazione (B.29) è data quindi da:

$$z(t) = e^{i\frac{\beta}{2m}t}\left(C_+ e^{\frac{1}{2}\sqrt{\frac{4\kappa}{m} - \frac{\beta^2}{m^2}}t} + C_- e^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m} - \frac{\beta^2}{m^2}}t}\right) \,,$$

per $C_\pm \in \mathbb{C}$ scelte arbitrariamente. Possiamo scrivere la stessa soluzione come

$$z(t) = \cos\left(\frac{\beta}{2m}t\right)\left(C_+ e^{\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t} + C_- e^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t}\right) + i\sin\left(\frac{\beta}{2m}t\right)\left(C_+ e^{\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t} + C_- e^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t}\right).$$

Ci sono due casi da considerare.
(i) *Caso* $\frac{4\kappa}{m} > \frac{\beta^2}{m^2}$. In questo caso le funzioni esponenziali scritte sopra sono tutte reali. Decomponendo $C_+ = a + ib$ e $C_- = c + id$ in parte reale ed immaginaria, in modo da avere solo costanti reali, si trova immediatamente:

$$x(t) = \cos\left(\frac{\beta}{2m}t\right)\left(ae^{\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t} + ce^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t}\right) - \sin\left(\frac{\beta}{2m}t\right)\left(be^{\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t} + de^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t}\right),$$

$$y(t) = \cos\left(\frac{\beta}{2m}t\right)\left(be^{\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t} + de^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t}\right) + \sin\left(\frac{\beta}{2m}t\right)\left(ae^{\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t} + ce^{-\frac{1}{2}\sqrt{\frac{4\kappa}{m}-\frac{\beta^2}{m^2}}t}\right),$$

dove le costanti a, b, c, d sono fissate arbitrariamente in $\mathbb{R}$.
(i) *Caso* $\frac{4\kappa}{m} < \frac{\beta^2}{m^2}$. In questo caso gli esponenti delle funzioni sono immaginari:

$$z(t) = \cos\left(\frac{\beta}{2m}t\right)\left(C_+ e^{i\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t} + C_- e^{-i\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t}\right) + i\sin\left(\frac{\beta}{2m}t\right)\left(C_+ e^{i\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t} + C_- e^{-i\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t}\right).$$

Usando la formula di Eulero e decomponendo nuovamente $C_+ = a + ib$ e $C_- = c + id$ in parte reale ed immaginaria, in modo da avere solo costanti reali, si ottiene infine, prendendo la parte reale di z, l'espressione esplicita della componente x della soluzione:

$$x(t) = (a+c)\cos\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right) + (c-a)\sin\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)$$

$$+(-b+d)\cos\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right) - (b+d)\sin\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right).$$

Prendendo invece la parte immaginaria di z si ha:

$$y(t) = (d+b)\cos\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right) + (d-b)\sin\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)$$

$$+(a-c)\cos\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right) + (a+c)\sin\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right).$$

Cambiando nuovamente nome alle costanti reali abbiamo alla fine, nelle costanti reali A, B, C, D:

$$x(t) = A\cos\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right) + B\sin\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)$$

$$+C\cos\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)+D\sin\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)\ ,$$

e

$$y(t)=-D\cos\left(\frac{\beta}{2m}t\right)\cos\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)+C\sin\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)$$

$$-B\cos\left(\frac{\beta}{2m}t\right)\sin\left(\frac{1}{2}\sqrt{\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)+A\sin\left(\frac{\beta}{2m}t\right)\cos\left(\sqrt{\frac{1}{2}\frac{\beta^2}{m^2}-\frac{4\kappa}{m}}t\right)\ .$$

3.1.**5**. *Suggerimento.* Procedere come nell'esercizio 3, dato che le due equazioni differenziali per la componente x e la componente y risultano disaccoppiate.

B.4 Esercizi del Capitolo 4.

Esercizi 4.1

4.1.**1**. *Soluzione.* Dalla definizione di $\mathbf{\Gamma}_O|_{\mathscr{I}}$ ed usando le relazioni (1.68) (per O' in quiete in $\mathscr{I}'$) si ha:

$$\mathbf{\Gamma}_O|_{\mathscr{I}}(t)=\ \mathbf{\Gamma}_O|_{\mathscr{I}'}(t)+\sum_{k=1}^{n}m_k(P_k(t)-O(t))\wedge\left(\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\wedge(P_k(t)-O'(t))\right)$$

$$+\sum_{k=1}^{n}m_k(P_k(t)-O(t))\wedge\mathbf{v}_{O'}|_{\mathscr{I}}(t)\ .$$

Dato che il punto O' in quiete con $\mathscr{I}'$ può essere scelto arbitrariamente, lo scegliamo tale che, all'istante t coincida con $O(t)$ (ne dobbiamo scegliere uno diverso ad ogni istante, ma questo non è proibito). In questo modo otteniamo

$$\mathbf{\Gamma}_O|_{\mathscr{I}}(t)=\ \mathbf{\Gamma}_O|_{\mathscr{I}'}(t)+\sum_{k=1}^{n}m_k(P_k(t)-O(t))\wedge(\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}}(t)\wedge(P_k(t)-O(t)))\ .$$

4.1.**2**. *Soluzione.* Dalla definizione di $\mathbf{\Gamma}_O|_{\mathscr{I}}$:

$$\mathbf{\Gamma}_O|_{\mathscr{I}}=\sum_{k=1}^{n}m_k(P_k-O')\wedge\mathbf{v}_{P_k}|_{\mathscr{I}}+\sum_{k=1}^{n}m_k(O'-O)\wedge\mathbf{v}_{P_k}|_{\mathscr{I}}=\ \mathbf{\Gamma}_{O'}|_{\mathscr{I}}+(O-O')\wedge\ \mathbf{P}|_{\mathscr{I}}\ ,$$

dove, nell'ultimo passaggio abbiamo usato (4.1).

Esercizi 4.2

4.2.**1**. *Soluzione.* Lavoriamo in coordinate polari sferiche centrate in O. Consideriamo una curva regolare chiusa contenuta in Ω, $\Gamma: r=r(s),\theta=\theta(s),\varphi=\varphi(s)$ con $s\in[0,1]$. In tal caso $\mathbf{x}(s)=r(s)\,\mathbf{e}_r(s)$, per cui:

$$\frac{d\mathbf{x}}{ds}=\frac{dr}{ds}\,\mathbf{e}_r(s)+r(s)\frac{d\theta}{ds}\mathbf{e}_\theta(s)+r(s)\sin(\theta(s))\frac{d\phi}{ds}\,\mathbf{e}_\varphi(s)\ .$$

Tenendo conto che $\mathbf{e}_r(s)$ è sempre perpendicolare a $\mathbf{e}_\theta(s)$ e $\mathbf{e}_\varphi(s)$ per costruzione, si ha immediatamente che, se $\mathbf{F} = F(r)\,\mathbf{e}_r$:

$$\oint_\Gamma \mathbf{F}(r)\cdot d\mathbf{x} = \int_0^1 F(r(s))\,\mathbf{e}_r\cdot\frac{d\mathbf{x}}{ds}ds = \int_0^1 F(r(s))\frac{dr}{ds}ds = \int_{r(0)}^{r(1)}\mathbf{F}(r)dr = 0\,,$$

dato che, per ipotesi, $r(0) = r(1)$. L'arbitrarietà della curva chiusa Γ ed (a) del teorema 4.4 implicano che $\mathbf{F}$ è conservativa.

4.2.**2**. *Soluzione.* Per prima cosa notiamo che il momento angolare $\mathbf{\Gamma}_O|_{\mathscr{I}} = (P-O)\wedge\mathbf{v}_P|_{\mathscr{I}}$ soddisfa, sul moto: $d\,\mathbf{\Gamma}_O|_{\mathscr{I}}/dt = (P-O)\wedge\mathbf{F} = \mathbf{0}$, dato che $\mathbf{F}$ è parallela a $P-O$ per ipotesi. Quindi il momento angolare si conserva nel tempo. Se $\mathbf{\Gamma}_O|_{\mathscr{I}} = 0$ significa che, in coordinate cartesiane solidali con $\mathscr{I}$ e centrate in O, $\mathbf{x} = (x,y,z)$, deve essere

$$\frac{d\mathbf{x}}{dt} = \lambda(t)\mathbf{x}(t)\,,$$

dove $\lambda(t)$ è una funzione continua, e quindi integrabile, eccetto al più per i tempi in cui $\mathbf{x}(t_0) = \mathbf{0}$. Passando in componenti, quest'equazione differenziale si integra immediatamente producendo:

$$\mathbf{x}(t) = e^{\int_0^t \lambda(\tau)d\tau}\mathbf{x}(0)\,.$$

Il moto avviene dunque lungo una direzione fissata inizialmente e quindi in un piano, che è sicuramente perpendicolare al momento angolare essendo quest'ultimo il vettore nullo. Nel caso $\mathbf{\Gamma}_O|_{\mathscr{I}} \neq 0$, la condizione:

$$\mathbf{\Gamma}_O|_{\mathscr{I}} = (P(t)-O)\wedge\mathbf{v}_P|_{\mathscr{I}}(t)$$

dice, in particolare, che $P(t)-O$ è sempre perpendicolare al vettore costante non nullo $\mathbf{\Gamma}_O|_{\mathscr{I}}$ e quindi $P(t)$ appartiene al piano passante per O e perpendicolare a tale vettore.

Esercizi 4.3

4.3.**1**. *Soluzione.* (i) Nel riferimento inerziale $\mathscr{I}$, il secondo principio delle dinamica si scrive:

$$m\mathbf{a}_P|_{\mathscr{I}} = -mg\,\mathbf{e}_z - k(P-H) + \boldsymbol{\phi}\,.$$

Esprimendo l'accelerazione $\mathbf{a}_P|_{\mathscr{I}}$ in funzione di $\mathbf{a}_P|_{\mathscr{I}'}$ tramite le (1.70), tenendo conto che la velocità angolare è costante, otteniamo che il secondo principio della dinamica, nel riferimento $\mathscr{I}'$, si scrive:

$$m\mathbf{a}_P|_{\mathscr{I}'} + 2m\omega\,\mathbf{e}_z\wedge\mathbf{v}_P|_{\mathscr{I}'} + m\omega^2\,\mathbf{e}_z\wedge(\,\mathbf{e}_z\wedge(P-O)) = -mg\,\mathbf{e}_z - k(P-H) + \boldsymbol{\phi}\,,$$

ossia:

$$m\mathbf{a}_P|_{\mathscr{I}'} = -2m\omega\,\mathbf{e}_z\wedge\mathbf{v}_P|_{\mathscr{I}'} - m\omega^2\,\mathbf{e}_z\wedge(\,\mathbf{e}_z\wedge(P-O)) - mg\,\mathbf{e}_z - k(P-H) + \boldsymbol{\phi}\,.$$

Poniamo $\mathbf{x} := P - O = x\mathbf{e}_x$. In tal caso: $\mathbf{v}_P|_{\mathscr{I}'} = \dot{x}\mathbf{e}_x$, dove il punto indica la derivata temporale, e $P - H = x\,\mathbf{e}_x - h\,\mathbf{e}_z$. Tenendo infine conto che il vincolo è liscio, per cui $\boldsymbol{\phi} = \phi^y\,\mathbf{e}_y + \phi^z\,\mathbf{e}_z$, l'equazione di sopra si riscrive come:

$$m\ddot{x}\,\mathbf{e}_x = (m\omega^2 - \kappa)x\,\mathbf{e}_x + (\phi^y - 2m\omega\dot{x})\,\mathbf{e}_y + (\kappa h + \phi^z - mg)\,\mathbf{e}_z\,, \tag{B.30}$$

da cui ricaviamo: l'equazione pura di movimento

$$\frac{d^2x}{dt^2} = \left(\omega^2 - \frac{\kappa}{m}\right)x\,, \tag{B.31}$$

e due equazioni che determinano la reazione vincolare una volta noto il moto

$$\phi^y(t) = 2m\omega\frac{dx}{dt}\,, \tag{B.32}$$

$$\phi^z = mg - \kappa h\,. \tag{B.33}$$

La forma della soluzione generale dell'equazione lineare e omogenea a coefficienti costanti (B.31), dipende dal valore del rapporto $m\omega^2/\kappa$. Abbiamo 3 casi.
(1) $m\omega^2/\kappa < 1$. In tal caso le soluzioni del polinomio caratteristico sono immaginarie pure e pertanto:

$$x(t) = A\sin\left(\sqrt{\frac{\kappa}{m} - \omega^2}\,t\right) + B\cos\left(\sqrt{\frac{\kappa}{m} - \omega^2}\,t\right)\,,$$

con $A, B \in \mathbb{R}$ costanti arbitrarie.
(2) $m\omega^2/\kappa = 1$. In tal caso le soluzioni del polinomio caratteristico sono coincidenti e valgono 0. Pertanto:

$$x(t) = Ae^{0t} + Bte^{0t} = A + Bt\,,$$

con $A, B \in \mathbb{R}$ costanti arbitrarie.
(3) $m\omega^2/\kappa > 1$. In tal caso le soluzioni del polinomio caratteristico sono reali e pertanto:

$$x(t) = A\exp\left(\sqrt{\omega^2 - \frac{\kappa}{m}}\,t\right) + B\exp\left(\sqrt{\omega^2 - \frac{\kappa}{m}}\,t\right)\,,$$

con $A, B \in \mathbb{R}$ costanti arbitrarie.
(ii) Consideriamo il caso (1). La condizione iniziale $\mathbf{v}_P|_{\mathscr{I}'}(0) = \mathbf{0}$ fornisce immediatamente $A = 0$. Imponendo $P(0) - O = x_0\,\mathbf{e}_x$ con $x_0 > 0$ si trova allora che: $B = x_0$. Pertanto, il moto con queste condizioni iniziali è

$$x(t) = x_0\cos\left(\sqrt{\frac{\kappa}{m} - \omega^2}\,t\right)\,, \quad \text{per ogni } t \in \mathbb{R}.$$

La reazione vincolare al variare del tempo, usando (B.32)-(B.33), risulta essere

$$\boldsymbol{\phi} = -2m\omega x_0\sqrt{\frac{\kappa}{m} - \omega^2}\sin\left(\sqrt{\frac{\kappa}{m} - \omega^2}\,t\right)\,\mathbf{e}_y + (mg - \kappa h)\,\mathbf{e}_z\,.$$

(iii) e (iv) La seconda equazione cardinale fornisce nel caso in esame

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} [(P(t)-O)\wedge m\mathbf{v}_P|_{\mathscr{I}}(t)] = (P(t)-O)\wedge(-mg\,\mathbf{e}_z - \kappa(P(t)-H)+\boldsymbol{\phi}(t))\ ,$$

cioé

$$\left.\frac{d}{dt}\right|_{\mathscr{I}} [(P(t)-O)\wedge m\mathbf{v}_P|_{\mathscr{I}}(t)] = x(t)\mathbf{e}_x(t)\wedge(-mg\,\mathbf{e}_z - \kappa x(t)\,\mathbf{e}_x(t) + h\kappa\,\mathbf{e}_z + \phi^y(t)\,\mathbf{e}_y(t) + \phi^z\,\mathbf{e}_z)\,.$$

La componente del secondo membro lungo l'asse z è , usando (B.32):

$$x(t)\phi^y(t) = 2m\omega x(t)\frac{dx}{dt} = m\omega\frac{d}{dt}x(t)^2\ .$$

Concludiamo che il momento angolare l'ungo l'asse z non si conserva eccetto nel caso in cui il moto non soddisfi $x(t)^2 = c$ con c costante, per ogni $t\in\mathbb{R}$, cioé $x(t)=\pm c$ per ogni $t\in\mathbb{R}$. Questo significa, dato che i moti sono funzioni continue, che gli unici moti ammissibili sono $x(t)=x_0$ per ogni $t\in\mathbb{R}$.
Nel caso (1): $m\omega^2/\kappa<1$, questo è possibile solo se $A=B=0$. Pertanto deve essere in tal caso $x(0)=0$ e $dx/dt(0)=0$.
Nel caso (2): $m\omega^2/\kappa=1$, questo è possibile solo se $A=0$. Pertanto deve essere in tal caso $x(0)=x_0$ arbitrario e $dx/dt(0)=0$.
Nel caso (3): $m\omega^2/\kappa>1$, questo è possibile solo se $A=B=0$. Pertanto deve essere in tal caso $x(0)=0$ e $dx/dt(0)=0$.

4.3.**2**. *Soluzione.* (i) Consideriamo in teorema delle forze vive nel riferimento $\mathscr{I}'$. Esso si esprime come

$$\frac{d\mathscr{T}|_{\mathscr{I}'}}{dt} = \mathbf{v}_P|_{\mathscr{I}'}\cdot\mathbf{F}\ ,$$

dove $\mathbf{F}$ include tutte le forze vere e inerziali che sono viste agire su P nel riferimento $\mathscr{I}'$ $\mathbf{F}$ è quindi pari al secondo membro di (B.30). In termini espliciti si ottiene:

$$\frac{d\mathscr{T}|_{\mathscr{I}'}}{dt} = \frac{dx}{dt}\,\mathbf{e}_x\cdot((m\omega^2-\kappa)x\,\mathbf{e}_x + (\phi^y - 2m\omega\dot{x})\,\mathbf{e}_y + (\kappa h + \phi^z - mg)\,\mathbf{e}_z)\,.$$

Notare che, in particolare, la forza di Coriolis ($2m\omega\dot{x}\,\mathbf{e}_y$) e la reazione vincolare non dissipano potenza, dando contributo nullo al secondo membro dell'equazione di bilancio scritta sopra. In definitiva si ha:

$$\frac{d\mathscr{T}|_{\mathscr{I}'}}{dt} = (m\omega^2-\kappa)\frac{dx(t)}{dt}x(t)\,.$$

che può scriversi come:

$$\frac{d\mathscr{T}|_{\mathscr{I}'}}{dt} = -\frac{d}{dt}\left(\kappa\frac{x^2}{2} - m\omega^2\frac{x^2}{2}\right)\ .$$

In definitiva, sui moti è conservata la quantità :

$$\mathscr{E}|_{\mathscr{I}'} := \mathscr{T}|_{\mathscr{I}'} + \kappa\frac{x^2}{2} - m\omega^2\frac{x^2}{2}\ .$$

Il secondo addendo a secondo membro è l'ordinaria energia potenziale della molla, mentre l'ultimo termine è un'energia potenziale attribuibile alla forza centrifuga, infatti:

$$-\nabla\left(-m\omega^2\frac{x^2}{2}\right) = (m\omega^2)x\,\mathbf{e}_x = -m\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (\boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P(t)-O)))$$

è proprio la forza centrifuga che appare in $\mathscr{I}'$.
(ii) Esaminiamo la situazione energetica nel riferimento $\mathscr{I}$. Nel riferimento $\mathscr{I}$ agiscono la forza della molla, la reazione vincolare e la forza di gravità . Le forze conservative sono quelle della molla e quella di gravità . L'energia potenziale gravitazionale mgz è però costante a causa del vincolo a cui è sottoposto il punto materiale. In definitiva, il teorema di bilancio dell'energia meccanica, definita come $\mathscr{E}|_{\mathscr{I}} = (1/2)m\mathbf{v}_P|_{\mathscr{I}} + (1/2)\kappa(P-H)^2 + mgz$, dice che deve essere:

$$\frac{d\mathscr{E}|_{\mathscr{I}}}{dt} = \mathbf{v}_P|_{\mathscr{I}} \cdot (\phi^y\,\mathbf{e}_y(t) + \phi^z\,\mathbf{e}_z)\,,$$

dove $\phi^y(t) = 2m\omega\dot{x}(t)$ per la (B.32), mentre $\mathbf{v}_P|_{\mathscr{I}} \cdot \phi^z\,\mathbf{e}_z = 0$ dato che il moto avviene nel piano $z = 0$. Sappiamo anche che:

$$\mathbf{v}_P|_{\mathscr{I}} = \mathbf{v}_P|_{\mathscr{I}'} + \boldsymbol{\omega}_{\mathscr{I}'}|_{\mathscr{I}} \wedge (P(t)-O) = \dot{x}(t)\,\mathbf{e}_x(t) + \omega\,\mathbf{e}_z \wedge x(t)\,\mathbf{e}_x(t) = \dot{x}(t)\,\mathbf{e}_x(t) + \omega x(t)\,\mathbf{e}_y(t)\,.$$

In definitiva:

$$\frac{d\mathscr{E}|_{\mathscr{I}}}{dt} = (\dot{x}(t)\,\mathbf{e}_x(t) + \omega x(t)\,\mathbf{e}_y(t)) \cdot 2m\omega\dot{x}(t)\,\mathbf{e}_y(t) = m\omega^2\frac{dx(t)^2}{dt}\,.$$

Otteniamo la stessa situazione già discussa nel precedente esercizio: $\mathscr{E}|_{\mathscr{I}}$ si conserva solo sui moti che soddisfano $x(t) =$ costante per ogni $t \in \mathbb{R}$. Le condizioni iniziali che forniscono tali moti sono state studiate nel precedente esercizio.

4.3.**3**. *Soluzione.* Lavoriamo in coordinate cilindriche r, φ, z adattate alla superficie cilindrica in modo ovvio. Ricordiamo che allora

$$P - O = z(P)\,\mathbf{e}_z + R\,\mathbf{e}_r(P)\,,\ \ \mathbf{v}_P = \dot{z}(P)\,\mathbf{e}_z + R\dot{\varphi}\,\mathbf{e}_\varphi(P)\,,$$

$$\mathbf{a}_P = -R\dot{\varphi}^2(P)\,\mathbf{e}_r(P) + R\ddot{\varphi}(P)\,\mathbf{e}_\varphi(P) + \ddot{z}\,\mathbf{e}_z$$

Il secondo principio della dinamica, per il punto P, è scritto come:

$$-MR\dot{\varphi}^2(P)\,\mathbf{e}_r(P) + MR\ddot{\varphi}(P)\,\mathbf{e}_\varphi(P) + M\ddot{z}\,\mathbf{e}_z = -Mg\,\mathbf{e}_z - \kappa(P-Q) - \gamma(P-O) + \phi^r(P)\,\mathbf{e}_r(P)\,,$$

che possiamo riscrivere

$$-MR\dot{\varphi}^2(P)\mathbf{e}_r(P) + MR\ddot{\varphi}(P)\mathbf{e}_\varphi(P) + M\ddot{z}(P)\mathbf{e}_z = -Mg\mathbf{e}_z - (\kappa+\gamma)(P-O) + \kappa(Q-O) + \phi^r(P)\mathbf{e}_r(P). \tag{B.34}$$

Per il punto Q avremo similmente:

$$-mR\dot{\varphi}^2(Q)\mathbf{e}_r(Q) + mR\ddot{\varphi}(Q)\mathbf{e}_\varphi(Q) + m\ddot{z}(Q)\mathbf{e}_z = -mg\mathbf{e}_z - (\kappa+\gamma)(Q-O) + \kappa(P-O) + \phi^r(Q)\mathbf{e}_r(Q). \tag{B.35}$$

Per aver tutto nella base associata alle coordinate cilindriche, teniamo conto che vale:

$$\mathbf{e}_x = \cos\varphi\,\mathbf{e}_r - \sin\varphi\,\mathbf{e}_\varphi \quad \text{e} \quad \mathbf{e}_y = \sin\varphi\,\mathbf{e}_r + \cos\varphi\,\mathbf{e}_\varphi\,.$$

Pertanto banalmente $P - O = z(P)\,\mathbf{e}_z + R\,\mathbf{e}_r(P)$, mentre

$$Q - O = z(Q)\,\mathbf{e}_z + R\,\mathbf{e}_r(Q) = z(Q)\,\mathbf{e}_z + R\cos\varphi(Q)\,\mathbf{e}_x + R\sin\varphi(Q)\,\mathbf{e}_y =$$

$$z(Q)\mathbf{e}_z + R\cos\varphi(Q)(\cos\varphi(P)\mathbf{e}_r(P) - \sin\varphi(P)\mathbf{e}_\varphi(P)) + R\sin\varphi(Q)(\sin\varphi(P)\mathbf{e}_r(P) + \cos\varphi(P)\mathbf{e}_\varphi(P)).$$

Cioé :

$$\begin{aligned} Q - O = z(Q)\,\mathbf{e}_z &+ R(\cos\varphi(Q)\cos\varphi(P) + \sin\varphi(Q)\sin\varphi(P))\,\mathbf{e}_r(P) \\ &+ R(\sin\varphi(Q)\cos\varphi(P) - \cos\varphi(Q)\sin\varphi(P))\,\mathbf{e}_\varphi(P)\,, \end{aligned}$$

ossia:

$$Q - O = z(Q)\,\mathbf{e}_z + R\cos\left(\varphi(Q) - \varphi(P)\right)\,\mathbf{e}_r(P) + R\sin\left(\varphi(Q) - \varphi(P)\right)\,\mathbf{e}_\varphi(P)$$

Inserendo in (B.34), raccogliendo i singoli fattori dei tre versori $\mathbf{e}_r(P)$, $\mathbf{e}_\varphi(P)$, $\mathbf{e}_z$ otteniamo le 3 equazioni, cambiando leggermente le notazioni:

$$\begin{aligned} MR\ddot{\varphi}_P &= \kappa R\sin\left(\varphi_Q - \varphi_P\right)\,, && \text{(B.36)} \\ M\ddot{z}_P &= -Mg - (\kappa+\gamma)z_P + \kappa z_Q\,, && \text{(B.37)} \\ MR\dot{\varphi}_P^2 &= (\gamma+\kappa)R - \kappa R\cos\left(\varphi_Q - \varphi_P\right) - \phi_P^r\,. && \text{(B.38)} \end{aligned}$$

Nello stesso modo, si ottengono (B.35) le equazioni per Q:

$$\begin{aligned} mR\ddot{\varphi}_Q &= -\kappa R\sin\left(\varphi_Q - \varphi_P\right)\,, && \text{(B.39)} \\ m\ddot{z}_Q &= -Mg - (\kappa+\gamma)z_Q + \kappa z_P\,, && \text{(B.40)} \\ mR\dot{\varphi}_Q^2 &= (\gamma+\kappa)R - \kappa R\cos\left(\varphi_Q - \varphi_P\right) - \phi_Q^r\,. && \text{(B.41)} \end{aligned}$$

Abbiamo allora il sistema delle equazioni pure di movimento

$$\begin{aligned} \ddot{\varphi}_P &= \frac{\kappa}{M}\sin\left(\varphi_Q - \varphi_P\right)\,, && \text{(B.42)} \\ \ddot{\varphi}_Q &= -\frac{\kappa}{m}\sin\left(\varphi_Q - \varphi_P\right)\,, && \text{(B.43)} \\ \ddot{z}_P &= -g - \frac{(\kappa+\gamma)}{M}z_P + \frac{\kappa}{M}z_Q\,, && \text{(B.44)} \\ \ddot{z}_Q &= -g - \frac{(\kappa+\gamma)}{m}z_Q + \frac{\kappa}{m}z_P\,, && \text{(B.45)} \end{aligned}$$

e le equazioni che determinano le reazioni vincolari su ogni moto:

$$\begin{aligned} \phi_P^r &= (\gamma+\kappa)R - \kappa R\cos\left(\varphi_Q - \varphi_P\right) - MR\dot{\varphi}_P^2\,. && \text{(B.46)} \\ \phi_Q^r &= (\gamma+\kappa)R - \kappa R\cos\left(\varphi_Q - \varphi_P\right) - mR\dot{\varphi}_Q^2\,. && \text{(B.47)} \end{aligned}$$

Quindi se a $t = 0$, $\varphi_P = 0$, $\varphi_Q = \pi/2$ e $\dot{\varphi}_P = \dot{\varphi}_Q = 0$, indipendentemente dai valori di z_P e z_Q le reazioni vincolari valgono $\boldsymbol{\phi}_P = (\gamma + \kappa)R\,\mathbf{e}_x$ e $\boldsymbol{\phi}_Q = (\gamma + \kappa)R\,\mathbf{e}_y$.
Per concludere proviamo che la quantità $M\dot{\varphi}_P + m\dot{\varphi}_Q$ si conserva sui moti. La prova si può ottenere direttamente dalle equazioni (B.42) e (B.43), moltiplicando ambo i membri delle prima equazione per M, quelli della seconda per m e sommando membro a membro le identità ottenute in tal modo. In tal caso si ricava che $M\dot{\varphi}_P + m\dot{\varphi}_Q$ ha derivata temporale nulla sulle soluzioni delle equazioni del moto, ed è pertanto una costante del moto. Una via alternativa è quella di notare che $MR\dot{\varphi}_P + mR\dot{\varphi}_Q$ altro non è che la componente lungo l'asse z del momento angolare totale del sistema rispetto al polo O e che, d'altra parte, il momento totale delle forze esterne applicate sul sistema non ha componenti lungo l'asse z. Di conseguenza, la seconda equazione cardinale assicura la conservazione di $R(M\dot{\varphi}_P + m\dot{\varphi}_Q)$.

B.5 Esercizi del Capitolo 5.

Esercizi 5.1

5.1.**1**. *Soluzione.* Le equazioni differenziali che determinano il moto sono

$$\begin{aligned} \frac{d^2x}{dt^2} &= -\frac{2\kappa}{m}x^2\,, \\ \frac{d^2y}{dt^2} &= -\frac{2\kappa}{m}y^2\,, \\ \frac{d^2z}{dt^2} &= \frac{4\kappa}{m}z^3\,. \end{aligned} \tag{B.48}$$

Consideriamo la soluzione massimale con condizione iniziale $\mathbf{x}(0) = (0, 0, z_0)$ e $\mathbf{v}(0) = (0, 0, 0)$ dove $z_0 > 0$. Sappiamo che tale soluzione esiste ed è unica in quanto ricadiamo nelle ipotesi del teorema di esistenza ed unicità . Le prime due equazioni scritte sopra ammettono come unica soluzione, rispettivamente, $x(t) = 0$ e $y(t) = 0$ per ogni $t \in \mathbb{R}$. La (B.48) non è integrabile in forma esplicita. Tuttavia si possono ottenere informazioni, sufficienti per dimostrare quanto voluto, usando un opportuno integrale primo. Moltiplicando per dz/dt ambo membri di (B.48) e raccogliendo a primo membro, con qualche banale passaggio si ottiene:

$$\frac{d}{dz}\left[\left(\frac{dz(t)}{dt}\right)^2 - 2\frac{\kappa}{m}z(t)^4\right] = 0\,.$$

Di conseguenza

$$2\frac{\kappa}{m}z_0^4 = \left(\frac{dz(t)}{dt}\right)^2 - 2\frac{\kappa}{m}z(t)^4$$

da cui:

$$\frac{dz(t)}{dt} = \pm\sqrt{2\frac{\kappa}{m}\left[z(t)^4 - z_0^4\right]}\,.$$

Quest'equazione può essere integrata[1] in:

$$t(z) = \pm \int_{z_0}^{z} \frac{d\zeta}{\sqrt{2\frac{\kappa}{m}\left[\zeta^4 - z_0^2\right]}} \,. \tag{B.49}$$

Fissato il segno, dato che l'integrando è strettamente positivo, la funzione può essere invertita ottenendo due funzioni monotone: $z = z_\pm(t)$. z_+ è definita per $t \geq 0$, mentre z_- è definita per $t < 0$ e z corre, in entrambi i casi, in $[z_0, +\infty)$. La soluzione massimale, che sappiamo esistere ed essere unica, deve coincidere, a tratti, con le funzioni $z_\pm$. Essa dunque è definita come: $z := z_-(t)$ se $t < 0$ e $z = z_+(t)$ se $t \geq 0$. Non è completa in quanto l'integrale a secondo membro in (B.49) converge a qualche valore $\omega \in \mathbb{R}$ per $z \to +\infty$, e qualche valore $\alpha \in \mathbb{R}$ per $z \to -\infty$. Questo è dovuto al fatto che, per $z \to \pm\infty$:

$$\frac{1}{\sqrt{z^4 - z_0^4}} = O(z^{-2}) \,.$$

Consideriamo una palla aperta B_ϵ in $\mathbb{R}^3 \times \mathbb{R}^3$ centrata nell'origine, $\mathbf{x} = \mathbf{0}$, $\mathbf{v} = \mathbf{0}$, e di raggio $\epsilon > 0$. Non è possibile trovare alcun intorno aperto U dell'origine di $\mathbb{R}^3 \times \mathbb{R}^3$ tale che, tutte le condizioni iniziali V producano soluzioni massimali con orbite contenute in B_ϵ. Possiamo infatti scegliere come condizione iniziale $\mathbf{x}(0) = (0, 0, z_0)$ e $\mathbf{v}(0) = (0, 0, 0)$ dove $z_0 > 0$ è sufficientemente piccolo da ricadere in V. La soluzione massimale associata fuoriesce da B_ϵ in un tempo finito, dato che la componente $z(t)$ della soluzione è crescente e ammette limite $+\infty$ per $t \to \omega$. La stessa cosa accade considerando l'andamento a $t \to \alpha$: $z(t) \to -\infty$. Concludiamo che la configurazione di equilibrio data dall'origine di $\mathbb{R}^3$ non è stabile nel passato e nel futuro.

B.6 Esercizi del Capitolo 6.

Esercizi 6.1

6.1.**1**. *Soluzione.* Notiamo preventivamente che il sistema $\mathscr{I}$ è inerziale per cui dobbiamo considerare solo forze vere, inoltre i vincoli sono olonomi ed ideali: possiamo applicare il formalismo lagrangiano.

(a). Usando le coordinate libere ϕ_P e ϕ_Q, le posizione dei punti P e Q sono determinate, nel riferimento $\mathscr{I}$, dai vettori posizione:

$$\mathbf{x}_P = R\cos\phi_P\,\mathbf{e}_x + R\sin\phi_P\,\mathbf{e}_y + R\phi_P\,\mathbf{e}_z \,, \tag{B.50}$$

$$\mathbf{x}_Q = R\cos\phi_Q\,\mathbf{e}_x + R\sin\phi_Q\,\mathbf{e}_y + R\phi_P\,\mathbf{e}_z \,. \tag{B.51}$$

[1] Si osservi che l'integrale converge per $z \to z_0$ in quanto:

$$\frac{1}{\sqrt{z^4 - z_0^4}} = \frac{1}{\sqrt{z - z_0}\sqrt{(z - z_0)(z^2 + z_0^2)}}$$

e la singolarità è dovuta solo al primo fattore al secondo membro che è integrabile.

Da cui ricaviamo le velocità (rispetto a $\mathscr{I}$):

$$\begin{aligned} \mathbf{v}_P &= \dot{\phi}_P \left(-R\sin\phi_P\,\mathbf{e}_x + R\cos\phi_P\,\mathbf{e}_y + R\,\mathbf{e}_z\right)\ , && \text{(B.52)} \\ \mathbf{v}_Q &= \dot{\phi}_Q \left(-R\sin\phi_Q\,\mathbf{e}_x + R\cos\phi_Q\,\mathbf{e}_y + R\,\mathbf{e}_z\right)\ . && \text{(B.53)} \end{aligned}$$

Dalle due ultime formule possiamo scrivere l'energia cinetica $\mathscr{T} = \frac{m}{2}\mathbf{v}_P^2 + \frac{m}{2}\mathbf{v}_Q^2$ riferita a $\mathscr{I}$, in funzione delle coordinate libere ϕ_P, ϕ_Q e delle associate coordinate puntate $\dot{\phi}_P, \dot{\phi}_Q$:

$$\mathscr{T} = \frac{mR^2}{2}\dot{\phi}_P^2\left(\sin^2\phi_P + \cos^2\phi_P + 1\right) + \frac{mR^2}{2}\dot{\phi}_Q^2\left(\sin^2\phi_Q + \cos^2\phi_Q + 1\right)\ .$$

Con ovvie semplificazioni si trova:

$$\mathscr{T} = mR^2\left(\dot{\phi}_P^2 + \dot{\phi}_Q^2\right)\ .$$

Consideriamo inizialmente il caso $\gamma = 0$. In questo caso tutte le forze attive sono conservative e di conseguenza il sistema ammette lagrangiana, $\mathscr{L}|_{\mathscr{I}} = \mathscr{L} = \mathscr{T} - \mathscr{U}$, dove $\mathscr{U} = \mathscr{U}|_{\mathscr{I}}$ è la somma delle energie potenziale di ogni forza conservativa agente sul sistema. In pratica abbiamo un contributo a $\mathscr{U}$ dovuto all'energia potenziale gravitazionale:

$$mgz_P + mgz_Q = mgR(\phi_P + \phi_Q)$$

ed un contributo dovuto all'energia potenziale della molla tra P e Q:

$$\frac{\kappa}{2}(P-Q)^2 = \frac{\kappa R^2}{2}\left[(\cos\phi_P - \cos\phi_Q)\,\mathbf{e}_x + (\sin\phi_P - \sin\phi_Q)\,\mathbf{e}_y + (\phi_P - \phi_Q)\,\mathbf{e}_z\right]^2\ .$$

Quest'ultimo contributo, eseguendo i quadrati, si semplifica in:

$$\frac{\kappa R^2}{2}\left[2(1 - \cos\phi_P\cos\phi_Q - \sin\phi_P\sin\phi_Q) + (\phi_P - \phi_Q)^2\right] = \frac{\kappa R^2}{2}\left[2(1 - \cos(\phi_P - \phi_Q)) + (\phi_P - \phi_Q)^2\right].$$

In definitiva:

$$\mathscr{U} = mgR(\phi_P + \phi_Q) + \kappa R^2\left[1 - \cos(\phi_P - \phi_Q) + \frac{(\phi_P - \phi_Q)^2}{2}\right]\ .$$

La lagrangiana del sistema, nel caso $\gamma = 0$, è data dunque da:

$$\mathscr{L} = mR^2\left(\dot{\phi}_P^2 + \dot{\phi}_Q^2\right) - mgR(\phi_P + \phi_Q) - \kappa R^2\left[1 - \cos(\phi_P - \phi_Q) + \frac{(\phi_P - \phi_Q)^2}{2}\right]\ . \qquad \text{(B.54)}$$

Nel caso $\gamma = 0$, la parte non banale del sistema di equazioni differenziali di E-L prende la forma:

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{q}^k} - \frac{\partial\mathscr{L}}{\partial q^k} = 0\,, \quad \text{dove } q^1 = \phi_P \text{ e } q^2 = \phi_Q.$$

La parte banale del sistema di equazioni è , ovviamente, quella che sancisce che, sul moto, le variabili puntate sono le derivate temporali di quelle non puntate. Nel caso in esame si trova allora il sistema in forma definitiva (del secondo ordine):

$$2mR^2\frac{d^2\phi_P}{dt^2} = -mgR - \kappa R^2\left[\sin(\phi_P-\phi_Q)+(\phi_P-\phi_Q)\right] , \tag{B.55}$$

$$2mR^2\frac{d^2\phi_Q}{dt^2} = -mgR + \kappa R^2\left[\sin(\phi_P-\phi_Q)+(\phi_P-\phi_Q)\right] . \tag{B.56}$$

Passiamo al caso in cui sia presente la forza viscosa agente sul solo punto P. Tale forza non ammette potenziale non essendo posizionale e quindi nemmeno conservativa. Pertanto dobbiamo usare la formulazione delle equazioni di E-L in cui appaiono esplicitamente le componenti lagrangiane delle forze attive:

$$\frac{d}{dt}\frac{\partial\mathscr{T}}{\partial\dot{q}^k} - \frac{\partial\mathscr{T}}{\partial q^k} = \mathscr{Q}_k , \quad \text{dove } q^1=\phi_P \text{ e } q^2=\phi_Q.$$

Tuttavia, tra le forze agenti sui due punti, solo la forza viscosa non ammette potenziale, per cui possiamo semplificare le equazioni di sopra come, usando la stessa lagrangiana di prima:

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{q}^k} - \frac{\partial\mathscr{L}}{\partial q^k} = \mathscr{Q}_k^{(visc)} , \quad \text{dove } q^1=\phi_P \text{ e } q^2=\phi_Q, \tag{B.57}$$

dove $\mathscr{Q}_k^{(visc)}$ sono le componenti lagrangiane della sola forza viscosa. Le forze conservative sono state inglobate a primo membro nel termine $-\mathscr{U}$ presente nella lagrangiana $\mathscr{L}$. Nelle nostre ipotesi abbiamo che

$$\mathscr{Q}_k^{(visc)} = \sum_{i=1}^{2}\frac{\partial\mathbf{x}_i}{\partial q^k}\cdot\mathbf{F}_i^{(visc)} = \frac{\partial\mathbf{x}_R}{\partial q^k}\cdot(-\gamma\mathbf{v}_P) .$$

In definitiva usando (B.50) e (B.50), otteniamo che nel calcolo delle componenti lagrangiane della forza viscosa, solo la componente con $k=1$ (cioé relativa a ϕ_P) è non nulla. Essa è data da:

$$\mathscr{Q}_1^{(visc)} = -\gamma\left[\frac{\partial}{\partial\phi_P}(R\cos\phi_P\,\mathbf{e}_x + R\sin\phi_P\,\mathbf{e}_y + R\phi_P\,\mathbf{e}_z)\right]\cdot\dot{\phi}_P\,(-R\sin\phi_P\,\mathbf{e}_x + R\cos\phi_P\,\mathbf{e}_y + R\,\mathbf{e}_z) .$$

Con qualche calcolo si ha l'espressione finale:

$$\mathscr{Q}_1^{(visc)} = -2\gamma R^2\dot{\phi}_P .$$

In questo caso le equazioni di E-L diventano quindi, esplicitando (B.57):

$$2mR^2\frac{d^2\phi_P}{dt^2} = mgR - \kappa R^2\left[\sin(\phi_P-\phi_Q)+(\phi_P-\phi_Q)\right] - 2\gamma R^2\frac{d\phi_P}{dt} , \tag{B.58}$$

$$2mR^2\frac{d^2\phi_Q}{dt^2} = mgR + \kappa R^2\left[\sin(\phi_P-\phi_Q)+(\phi_P-\phi_Q)\right] . \tag{B.59}$$

(b). Il cambiamento di variabili $\theta = \phi_P - \phi_Q$, $\tau = \phi_p + \phi_Q$ è biettivo e bidifferenziabile. Le formule inverse sono: $\phi_P = (\theta + \tau)/2$ e $\phi_Q = (\tau - \theta)/2$ da cui è evidente che, variando la coordinata τ e tenendo fissa la coordinata θ, si produce una rotazione rigida del sistema di punti di un angolo $\tau/2$ composta con una traslazione lungo l'asse z di $R\tau/2$.
Dalle leggi di trasformazione di coordinate scritte sopra si ha che: $\dot{\phi}_P = (\dot{\theta} + \dot{\tau})/2$ e $\dot{\phi}_Q = (\dot{\tau} - \dot{\theta})/2$. Sostituendo tutto nell'espressione (B.54) troviamo infine l'espressione della lagrangiana in funzione delle nuove variabili:

$$\mathscr{L}(\theta, \tau, \dot{\theta}, \dot{\tau}) = \frac{mR^2}{2}\left((\dot{\theta} + \dot{\tau})^2 + (\dot{\tau} - \dot{\theta})^2\right) + mgR\tau - \kappa R^2\left[1 - \cos\theta + \frac{\theta^2}{2}\right] ,$$

ossia

$$\mathscr{L}(\theta, \tau, \dot{\theta}, \dot{\tau}) = \frac{mR^2}{2}\left(\dot{\theta}^2 + \dot{\tau}^2\right) + mgR\tau - \kappa R^2\left[1 - \cos\theta + \frac{\theta^2}{2}\right] .$$

Nel caso $g = 0$ la lagrangiana si riduce a:

$$\mathscr{L}(\theta, \tau, \dot{\theta}, \dot{\tau}) = \frac{mR^2}{2}\left(\dot{\theta}^2 + \dot{\tau}^2\right) - \kappa R^2\left[1 - \cos\theta + \frac{\theta^2}{2}\right] .$$

In questo caso l'equazione di E-L per la coordinata τ, visto che $\mathscr{L}$ non dipende esplicitamente da τ risulta essere:

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{\tau}}\right) = 0 .$$

In altre parole, la grandezza:

$$I := \frac{\partial \mathscr{L}}{\partial \dot{\tau}} = mR^2 \frac{d\tau}{dt} ,$$

si conserva nel tempo sul moto (cioé sulle soluzioni delle equazioni di Eulero-Lagrange).
Tornando alle coordinate libere ϕ_P e ϕ_Q la grandezza L_z assume forma:

$$I := \frac{\partial \mathscr{L}}{\partial \dot{\tau}} = mR^2\left(\frac{d\phi_P}{dt} + \frac{d\phi_Q}{dt}\right) .$$

Consideriamo ora il momento angolare del sistema di punti nel riferimento $\mathscr{I}$ rispetto al polo O:

$$\mathbf{\Gamma}_O|_{\mathscr{I}} = (P - O) \wedge m\mathbf{v}_P + (Q - O) \wedge m\mathbf{v}_Q .$$

Il calcolo esplicito, usando (B.50), (B.51), (B.52) e (B.53), fornisce immediatamente che

$$\mathbf{e}_z \cdot \mathbf{\Gamma}_O|_{\mathscr{I}} = mR^2\left(\frac{d\phi_P}{dt} + \frac{d\phi_Q}{dt}\right) .$$

Concludiamo, di conseguenza, che la grandezza conservata I non è altro che la componente lungo l'asse z momento angolare del sistema di punti nel riferimento $\mathscr{I}$ rispetto al polo O.

6.1.**2**. *Soluzione*. Notiamo preventivamente che il sistema $\mathscr{I}$ è inerziale per cui dobbiamo considerare solo forze vere, inoltre i vincoli sono olonomi ed ideali: possiamo applicare il formalismo lagrangiano.

(a) Determiniamo la lagrangiana del sistema, che indicheremo con $\mathscr{L}$ sottointendendo il riferimento $\mathscr{I}$. Dato che tutte la forze attive sono conservative la lagrangiana avrà forma $\mathscr{L} = \mathscr{T} - \mathscr{U}$, dove $\mathscr{T}$ è l'energia cinetica riferita ad $\mathscr{I}$, mentre $\mathscr{U}$ è l'energia potenziale totale riferita allo stesso sistema di riferimento. Valutiamo per prima l'energia cinetica. Scriviamo a tal fine la posizione e la velocità dei due punti in funzione delle coordinate libere r_1, φ_1 e r_2, φ_1 e delle associate coordinate puntate. Tenendo conto che, data l'equazione del cono, la quota z_1 di P_1 coincide con il raggio r_1 che ha lo stesso punto in coordinate polari piane nel piano $z = 0$, si ha banalmente che

$$P_1 - O = r_1(\cos\varphi_1\, \mathbf{e}_x + \sin\varphi_1\, \mathbf{e}_y + \mathbf{e}_z)\,.$$

e similmente:

$$P_2 - O = r_1(\cos\varphi_2\, \mathbf{e}_x + \sin\varphi_2\, \mathbf{e}_y + \mathbf{e}_z)\,.$$

Ora lavoriamo solo con il punto P_1 . Assumendo che le coordinate libere r_1, φ_1 dipendano dal tempo, derivando nel tempo e sostituendo con le variabili puntate le derivate temporali delle coordinate libere, si ha:

$$\mathbf{v}_1 = \dot{r}_1(\cos\varphi_1\, \mathbf{e}_x + \sin\varphi_1\, \mathbf{e}_y + \mathbf{e}_z) + r_1\dot{\varphi}_1(-\sin\varphi_1\, \mathbf{e}_x + \cos\varphi_1\, \mathbf{e}_y)\,,$$

da cui, quadrando e tenendo conto dell'identità $\cos^2\varphi_1 + \sin^2\varphi_1 = 1$ troviamo:

$$\mathbf{v}_1^2 = 2\dot{r_1}^2 + r_1^2\dot{\varphi}_1^2\,.$$

Analogamente deve essere:

$$\mathbf{v}_2^2 = 2\dot{r_2}^2 + r_2^2\dot{\varphi}_2^2\,.$$

Mettendo tutto insieme, l'energia cinetica si scrive:

$$\mathscr{T} = \frac{m}{2}\left(2(\dot{r_1}^2 + \dot{r_2}^2) + r_1^2\dot{\varphi}_1^2 + r_2^2\dot{\varphi}_2^2\right)\,. \tag{B.60}$$

L'energia potenziale è data dalla somma delle energie potenziali gravitazionali $mgz_1 + mgz_2$ più dall'energia potenziale della molla: $\kappa(P_1 - R_2)^2/2$. In particolare, dall'espressione trovata sopra per $P_1 - O$ e $P_2 - O$ si ha:

$$(P_1 - P_2)^2 = (r_1\cos v\phi_1 - r_2\cos\varphi_2)^2 + (r_1\sin\varphi_1 - r_2\sin\varphi_2)^2 + (r_1 - r_2)^2\,,$$

da cui, sviluppando i quadrati e tenendo conto dell'identità : $\cos(a-b) = \cos a\cos b + \sin a\sin b$ si ricava:

$$(P_1 - P_2)^2 = r_1^2 + r_2^2 - 2r_1r_2\cos(\varphi_1 - \varphi_2).$$

In definitiva:

$$\mathscr{U}(P_1, P_2) = mg(r_1 + r_2) + \frac{\kappa}{2}\left(r_1^2 + r_2^2 - 2r_1r_2\cos(\varphi_1 - \varphi_2)\right)\,, \tag{B.61}$$

da cui, tenendo conto di (B.60):

$$\mathscr{L} = \frac{m}{2}\left(2(\dot{r_1}^2 + \dot{r_2}^2) + r_1^2\dot{\varphi}_1^2 + r_2^2\dot{\varphi}_2^2\right) - mg(r_1 + r_2) - \frac{\kappa}{2}\left(r_1^2 + r_2^2 - 2r_1r_2\cos(\varphi_1 - \varphi_2)\right)\,. \tag{B.62}$$

Con questa lagrangiana le equazioni del moto di Eulero-Lagrange risultano essere

$$
\begin{aligned}
2m\frac{d^2r_1}{dt^2} &= mr_1\left(\frac{d\varphi_1}{dt}\right)^2 - mg - k(r_1 - r_2\cos(\varphi_1-\varphi_2))\,, && \text{(B.63)}\\
2m\frac{d^2r_2}{dt^2} &= mr_2\left(\frac{d\varphi_2}{dt}\right)^2 - mg - k(r_2 - r_1\cos(\varphi_1-\varphi_2))\,, && \text{(B.64)}\\
m\frac{d}{dt}\left(r_1^2\frac{d\varphi_1}{dt}\right) &= -\kappa r_1 r_2\sin(\varphi_1-\varphi_2)\,, && \text{(B.65)}\\
m\frac{d}{dt}\left(r_2^2\frac{d\varphi_2}{dt}\right) &= -\kappa r_1 r_2\sin(\varphi_2-\varphi_1)\,. && \text{(B.66)}
\end{aligned}
$$

(b) Vale $\Phi := (\varphi_1-\varphi_2)/2$ e $\Theta = (\varphi_1+\varphi_2)/2$ da cui $\varphi_1 = \Phi+\Theta$, $\varphi_2 = \Theta-\Phi$, $\dot\varphi_1 = \dot\Phi+\dot\Theta$, $\dot\varphi_2 = \dot\Theta-\dot\Phi$. La lagrangiana (B.62) espressa nelle nuove variabili assume pertanto la forma:

$$
\begin{aligned}
\mathscr{L} = \frac{m}{2}\Big[2(\dot r_1{}^2+\dot r_2{}^2) + (r_1^2+r_2^2)\left(\dot\Phi^2+\dot\Theta^2\right) + 2(r_1^2-r_2^2)\dot\Phi\dot\Theta\Big]\\
-mg(r_1+r_2) - \frac{\kappa}{2}\left(r_1^2+r_2^2-2r_1r_2\cos 2\Phi\right)\,.
\end{aligned}
$$

Quindi la lagrangiana non dipende esplicitamente da Θ e pertanto la grandezza $\partial\mathscr{L}/\partial\dot\Theta$ si conserva sui moti del sistema in virtù delle equazioni di Eulero-Lagrange. Tornando nelle coordinate iniziali:

$$
\frac{\partial\mathscr{L}}{\partial\dot\Theta} = \frac{m}{2}2\Big[(\dot r_1{}^2+\dot r_2{}^2)\dot\Theta + (\dot r_1{}^2-\dot r_2{}^2)\dot\Phi\Big] = \frac{m}{2}\left[(\dot r_1{}^2+\dot r_2{}^2)(\dot\varphi_1+\dot\varphi_2) + (\dot r_1{}^2-\dot r_2{}^2)(\dot\varphi_1-\dot\varphi_2)\right]\,.
$$

Sviluppando l'espressione si ha infine, sul moto:

$$
\frac{\partial\mathscr{L}}{\partial\dot\Theta} = mr_1^2\frac{d\varphi_1}{dt} + mr_2^2\frac{d\varphi_2}{dt}\,.
$$

Con banali passaggi di calcolo vettoriale, esprimendo $P_i - O$ e le velocità $\mathbf{v}_i$ in funzione delle coordinate libere, come determinato all'inizio dell'esercizio, si verifica facilmente che:

$$
\frac{\partial\mathscr{L}}{\partial\dot\Theta} = \mathbf{e}_z\cdot\mathbf{\Gamma}_0|_{\mathscr{I}}\,,
$$

dove:

$$
\mathbf{\Gamma}_0|_{\mathscr{I}} = (P_1-0)\wedge m\mathbf{v}_1 + (P_2-0)\wedge m\mathbf{v}_2
$$

è il momento angolare totale del sistema nel riferimento $\mathscr{I}$ rispetto al polo O.

6.1.**3**. *Soluzione.* (a) Usiamo come coordinata libera la componente x del vettore $P-O$ nel riferimento non inerziale $\mathscr{I}$. Dato che nel riferimento inerziale non compaiono forze non conservative (esclusa ovviamente la reazione vincolare $\boldsymbol{\phi}$), è conveniente usare la lagrangiana rispetto al riferimento inerziale $\mathscr{L}|_{\hat{\mathscr{I}}} = \mathscr{T}|_{\hat{\mathscr{I}}} - \mathscr{U}|_{\hat{\mathscr{I}}}$. Ricaviamo l'energia cinetica $\mathscr{T}|_{\hat{\mathscr{I}}} = \frac{1}{2}m\mathbf{v}_P^2|_{\hat{\mathscr{I}}}$

espressa in funzione di x e $\dot{x}$. La velocità nel riferimento non inerziale $\mathscr{I}$ è data banalmente dalla derivata nel tempo di:

$$P - O = x(t)\,\mathbf{e}_x + \sinh x(t)\,\mathbf{e}_z\ ,$$

tenendo conto che i versori $\mathbf{e}_x$ *e* $\mathbf{e}_z$ *sono fissi in tale riferimento.* Abbiamo allora che:

$$\mathbf{v}_P|_{\mathscr{I}} = \dot{x}\,\mathbf{e}_x + \dot{x}\cosh x\,\mathbf{e}_z\ .$$

La velocità nel riferimento $\hat{\mathscr{I}}$ si ottiene allora come

$$\mathbf{v}_P|_{\hat{\mathscr{I}}} = \mathbf{v}_P|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge (P - O) = \dot{x}\,\mathbf{e}_x + \dot{x}\cosh x\,\mathbf{e}_z + \Omega\,\mathbf{e}_z \wedge (x\,\mathbf{e}_x + \sinh x\,\mathbf{e}_z)\ .$$

Svolgendo il calcolo otteniamo:

$$\mathbf{v}_P|_{\hat{\mathscr{I}}} = \dot{x}\,\mathbf{e}_x + \Omega x\,\mathbf{e}_y + \dot{x}\cosh x\,\mathbf{e}_z\ . \tag{B.67}$$

Otteniamo di conseguenza che:

$$\mathscr{T}|_{\hat{\mathscr{I}}} = \frac{m}{2}\dot{x}^2\left(1 + \cosh^2 x\right) + \frac{m}{2}\Omega^2 x^2\ . \tag{B.68}$$

Abbiamo due contributi all'energia potenziale $\mathscr{U}|_{\hat{\mathscr{I}}}$. Uno è dovuto alla forza di gravità : mgz e l'altro è dovuto alla forza della molla: $(\kappa/2)x^2$. In definitiva:

$$\mathscr{L}|_{\hat{\mathscr{I}}} = \frac{m}{2}\dot{x}^2\left(1 + \cosh^2 x\right) - mg\sinh x - \frac{1}{2}(\kappa - m\Omega^2)x^2\ . \tag{B.69}$$

Con questa lagrangiana abbiamo che

$$\frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{x}} = m\dot{x}(1 + \cosh^2 x)\ , \qquad \frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial x} = m\dot{x}^2\sinh x\cosh x + (m\Omega^2 - \kappa)x - mg\cosh x\ .$$

Per cui, le equazioni di Eulero-Lagrange si scrivono:

$$\frac{d}{dt}m\left(\frac{dx}{dt}(1 + \cosh^2 x)\right) = m\dot{x}^2\sinh x\cosh x + (m\Omega^2 - \kappa)x - mg\cosh x$$

che forniscono:

$$m\frac{d^2x}{dt^2}(1+\cosh^2 x) + 2m\left(\frac{dx}{dt}\right)^2\sinh x\cosh x = m\left(\frac{dx}{dt}\right)^2\sinh x\cosh x + (m\Omega^2 - \kappa)x - mg\cosh x\,.$$

La forma finale delle equazioni di Eulero-Lagrange è pertanto:

$$m\frac{d^2x}{dt^2} = \frac{(m\Omega^2 - \kappa)x - mg\cosh x - m\left(\frac{dx}{dt}\right)^2\sinh x\cosh x}{1 + \cosh^2 x} \tag{B.70}$$

(b) Direttamente dalla definizione di $\mathscr{H}$ abbiamo che:

$$\frac{d}{dt}\mathscr{H} = \frac{d\dot{x}}{dt}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial \dot{x}} + \dot{x}\frac{d}{dt}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial \dot{x}} - \frac{d}{dt}\mathscr{L}|_{\hat{\mathscr{I}}} \,.$$

Usando le equazioni di Eulero-Lagrange possiamo riscrivere l'identità di sopra come:

$$\frac{d}{dt}\mathscr{H} = \frac{d\dot{x}}{dt}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial \dot{x}} + \dot{x}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial x} - \frac{d}{dt}\mathscr{L}|_{\hat{\mathscr{I}}} \,.$$

Aggiungendo e togliendo $\partial \mathscr{L}|_{\hat{\mathscr{I}}}/\partial t$ si ottiene infine:

$$\frac{d}{dt}\mathscr{H} = \frac{d\dot{x}}{dt}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial \dot{x}} + \dot{x}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial x} + \frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial t} - \frac{d}{dt}\mathscr{L}|_{\hat{\mathscr{I}}} - \frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial t} \,.$$

D'altra parte, dato che $\mathscr{L}|_{\hat{\mathscr{I}}} = \mathscr{L}|_{\hat{\mathscr{I}}}(t, x, \dot{x})$ deve anche essere:

$$\frac{d}{dt}\mathscr{L}|_{\hat{\mathscr{I}}} = \frac{d\dot{x}}{dt}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial \dot{x}} + \dot{x}\frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial x} + \frac{\partial \mathscr{L}|_{\hat{\mathscr{I}}}}{\partial t} \,.$$

Inserendo questa identità nell'espressione trovata sopra otteniamo infine che, su ogni soluzione delle equazioni di Eulero-Lagrange vale l'identità notevole:

$$\frac{d}{dt}\mathscr{H}(t, x(t), \dot{x}(t)) = -\frac{\partial}{\partial t}\mathscr{L}|_{\hat{\mathscr{I}}}(t, x(t), \dot{x}(t)) \,.$$

Dato che, nel caso in esame da (B.69) è evidente che $\mathscr{L}|_{\hat{\mathscr{I}}}$ non dipende esplicitamente dal tempo, si ha

$$\frac{d}{dt}\mathscr{H}(t, x(t), \dot{x}(t)) = -\frac{\partial}{\partial t}\mathscr{L}|_{\hat{\mathscr{I}}}(t, x(t), \dot{x}(t)) = 0 \,,$$

e pertanto su ogni soluzione delle equazioni di Eulero-Lagrange $\mathscr{H}$ è una costante. Passiamo a stabilire il significato fisico di $\mathscr{H}$, che calcolato esplicitamente da (B.69) fornisce subito:

$$\mathscr{H}(x, \dot{x}) = \frac{m}{2}\dot{x}^2\left(1 + \cosh^2 x\right) + mg\sinh x + \frac{1}{2}(\kappa - m\Omega^2)x^2 \,. \tag{B.71}$$

Vogliamo dimostrare, per verifica diretta che $\mathscr{H}$ è l'energia meccanica totale valutata nel riferimento non inerziale $\mathscr{I}$. Nel riferimento inerziale $\hat{\mathscr{I}}$ non si conserva l'energia meccanica: per mantenere il sistema in rotazione bisogna fornire lavoro dall'esterno. Tale lavoro è dissipato dalla reazione vincolare $\boldsymbol{\phi}$. Si osservi infatti che la reazione vincolare $\boldsymbol{\phi}$ in generale compie lavoro nel riferimento $\hat{\mathscr{I}}$ dato che essa è normale alla curva, ma la direzione della curva non è quella della velocità del punto materiale, visto che la curva stessa è in movimento in $\hat{\mathscr{I}}$. Valutiamo la situazione nel riferimento non inerziale $\mathscr{I}$, tenendo conto della presenza delle due forze inerziali di Coriolis $-m\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge \mathbf{v}_P|_{\mathscr{I}}$ e quella centrifuga $-m\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge (\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge (P - O))$ (le altre forze inerziali sono assenti dato che $\boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}}$ è costante nel tempo). Nel riferimento non inerziale, la forza $\boldsymbol{\phi}$ non compie lavoro in quanto, per costruzione, è sempre perpendicolare alla velocità del punto materiale P in quel riferimento. Le altre forze agenti sono quella di gravità che

è conservativa, quella della molla che è conservativa, quella di Coriolis che non compie lavoro (dato che per definizione è perpendicolare alla velocità), infine c'è ancora la forza centrifuga che, come vedremo è conservativa. L'energia meccanica totale nel sistema di riferimento $\mathscr{I}$, data dalla somma dell'energia cinetica e delle tre forme di energia potenziale suddette è quindi una costante del moto. Mostriamo che coincide con $\mathscr{H}$. Scriviamo l'equazione del moto in forma Newtoniana nel riferimento $\mathscr{I}$. Se $\mathbf{x} = P - O = x\mathbf{e}_x + \sinh x \mathbf{e}_z$ e quindi $\mathbf{v}_P|_{\mathscr{I}} = \dot{x}\mathbf{e}_x + \dot{x}\cosh x \mathbf{e}_z$, abbiamo che:

$$m\,\frac{d^2}{dt^2}\bigg|_{\mathscr{I}}\mathbf{x} = -mg\mathbf{e}_z - \kappa x\mathbf{e}_x - m\Omega\mathbf{e}_z\wedge\left(\frac{dx}{dt}\,\mathbf{e}_x + \frac{dx}{dt}\cosh x\,\mathbf{e}_z\right) - m\Omega^2\mathbf{e}_z\wedge(\,\mathbf{e}_z\wedge(x\,\mathbf{e}_x + \sinh x\,\mathbf{e}_z)) + \boldsymbol{\phi}.$$

Sviluppando entrambi i membri troviamo:

$$m\frac{d^2}{dt^2}(x\,\mathbf{e}_x + \sinh x\,\mathbf{e}_z) = -mg\,\mathbf{e}_z - \kappa x\,\mathbf{e}_x + m\Omega^2 x\,\mathbf{e}_x - m\Omega\frac{dx}{dt}\,\mathbf{e}_y + \boldsymbol{\phi}\,. \tag{B.72}$$

Le due ultime forze non compiono lavoro essendo perpendicolari alla velocità (che è sempre diretta lungo $\mathbf{e}_x$). La forza centrifuga si ottiene derivando in x e cambiando segno all'energia potenziale centrifuga $-\frac{1}{2}m\Omega^2x^2$, che corrisponde a quella di una molla con costante elastica negativa. In base al teorema di conservazione dell'energia meccanica, si deve conservare nel tempo, su ogni moto, la somma dell'energia cinetica e dell'energia potenziale totale. L'energia cinetica in $\mathscr{I}$ vale

$$\frac{1}{2}m\mathbf{v}_P^2|_{\mathscr{I}} = \frac{m}{2}\left(\frac{dx}{dt}\right)^2\left(1 + \cosh^2 x\right)\,,$$

pertanto l'integrale primo suddetto ha la forma:

$$\frac{m}{2}\left(\frac{dx}{dt}\right)^2\left(1 + \cosh^2 x\right) + mg\sinh x + \frac{1}{2}(\kappa - m\Omega^2)x^2\,.$$

Questa è , su ogni moto, proprio l'espressione trovata per $\mathscr{H}$ (B.71).
(c) Direttamente da (B.72) abbiamo che:

$$\phi^x = m\frac{d^2x}{dt^2} + (\kappa - m\Omega^2)x\,.$$

Usando (B.70) per esplicitare md^2x/dt^2 in funzione di x e della sua derivata prima nel tempo, si trova su ogni moto:

$$\phi^x(x,\dot{x}) = (\kappa - m\Omega^2)x + \frac{(m\Omega^2 - \kappa)x - mg\cosh x - m\left(\frac{dx}{dt}\right)^2\sinh x\cosh x}{1 + \cosh^2 x}$$

cioè , in funzione di x e $\dot{x}$:

$$\phi^x(x,\dot{x}) = \frac{(\kappa - m\Omega^2)x\cosh^2 x - mg\cosh x - m\dot{x}^2\sinh x\cosh x}{1 + \cosh^2 x}\,. \tag{B.73}$$

(d) Dato che $\mathscr{H}$ si conserva nel tempo, per il moto con condizioni iniziali $x(0) = 0$ e $\dot{x}(0) = v > 0$ deve essere:

$$\mathscr{H}(0, v) = \mathscr{H}(x(t), \dot{x}(t)) \,, \quad \text{per ogni } t.$$

Di conseguenza, direttamente da (B.71) troviamo che deve essere

$$mv^2 = \frac{m}{2}\dot{x}(t)^2 \left(1 + \cosh^2 x(t)\right) + mg \sinh x(t) + \frac{1}{2}(\kappa - m\Omega^2)x(t)^2 \,,$$

da cui:

$$m\dot{x}(t)^2 = \frac{2mv^2 - 2mg \sinh x(t) + (m\Omega^2 - \kappa)x(t)^2}{1 + \cosh^2 x(t)} \,.$$

Sostituendo in (B.73) si ricava l'espressione richiesta dal quesito:

$$\phi^x(x(t)) = \frac{(\kappa - m\Omega^2)x(t)\cosh^2 x(t) - mg \cosh x(t)}{1 + \cosh^2 x(t)}$$

$$+ \frac{2mg \sinh x(t) - 2mv^2 + (\kappa - m\Omega^2)x(t)^2}{(1 + \cosh^2 x(t))^2} \sinh x(t) \cosh x(t)$$

6.1.**4**. *Soluzione*. (a) La Lagrangiana riferita al riferimento inerziale $\hat{\mathscr{I}}$, tenendo conto che tutte le forze attive sono conservative, ha la forma generale, con ovvie notazioni:

$$\mathscr{L} = \frac{m}{2}(\mathbf{v}_P^2 + \mathbf{v}_Q^2) - \mathscr{U} \,. \tag{B.74}$$

Le velocità $\mathbf{v}_P$ e $\mathbf{v}_Q$ sono riferite al sistema $\hat{\mathscr{I}}$. Dobbiamo scrivere tutte le grandezze che appaiono nella lagrangiana di sopra in funzione di $\phi, \dot{\phi}, \psi, \dot{\psi}$. Cominciamo con il notare che $Q - O = Q - P + P - O$ e ancora

$$P - O = L(\cos\phi\, \mathbf{e}_x + \sin\phi\, \mathbf{e}_z) \,, \quad Q - P = L(\cos\psi\, \mathbf{e}_x + \sin\psi\, \mathbf{e}_z) \,. \tag{B.75}$$

Di conseguenza

$$Q - O = L[(\cos\phi + \cos\psi)\, \mathbf{e}_x + (\sin\phi + \sin\psi)\, \mathbf{e}_z] \,. \tag{B.76}$$

Queste espressioni permettono di esplicitare l'espressione di

$$\mathscr{U} = mg(z_P + z_Q) + \frac{k}{2}(Q - O)^2$$

in funzione degli angoli, ottenendo:

$$\mathscr{U}(\phi, \psi) = mgL(2\sin\phi + \sin\psi) + \frac{kL^2}{2}\left((\cos\phi + \cos\psi)^2 + (\sin\phi + \sin\psi)^2\right)^2 \,,$$

da cui:

$$\mathscr{U}(\phi, \psi) = mgL(2\sin\phi + \sin\psi) + kL^2(1 + \cos(\phi - \psi)) \,. \tag{B.77}$$

Le espressioni calcolate sopra per $P-O$ e $Q-O$ permettono anche di calcolare le velocità di P e Q nel riferimento $\mathscr{I}$ in funzione di $\phi, \dot{\phi}, \psi, \dot{\psi}$. Le velocità nel riferimento $\hat{\mathscr{I}}$ si ottengono usando le solite leggi di trasformazione cinematica, tenendo conto del moto relativo dei due riferimenti. Assumendo che gli angoli varino su un moto del sistema e derivando nel tempo $P-O$ e $Q-O$ in (B.75) e (B.76), tenendo conto che i versori sono costanti nel tempo nel riferimento $\mathscr{I}$ otteniamo:

$$\begin{aligned} \mathbf{v}_P|_{\mathscr{I}} &= L(-\dot{\phi}\sin\phi\,\mathbf{e}_x + \dot{\phi}\cos\phi\,\mathbf{e}_z)\,, && \text{(B.78)} \\ \mathbf{v}_Q|_{\mathscr{I}} &= -L(\dot{\phi}\sin\phi + \dot{\psi}\sin\psi)\,\mathbf{e}_x + L(\dot{\phi}\cos\phi + \dot{\psi}\cos\psi)\,\mathbf{e}_z\,. && \text{(B.79)} \end{aligned}$$

La velocità di P e Q in $\hat{\mathscr{I}}$ si ottengono dalle formule:

$$\mathbf{v}_P = \mathbf{v}_P|_{\mathscr{I}} + \boldsymbol{\omega}_{\mathscr{I}}|_{\hat{\mathscr{I}}} \wedge (P-O) = \mathbf{v}_P|_{\mathscr{I}} + \Omega\,\mathbf{e}_z \wedge (P-O)\,, \quad \mathbf{v}_Q = \mathbf{v}_Q|_{\mathscr{I}} + \Omega\,\mathbf{e}_z \wedge (Q-O)\,.$$

Il calcolo esplicito, facendo uso di (B.75), (B.76), (B.78) e (B.79), produce facilmente:

$$\begin{aligned} \mathbf{v}_P &= -\dot{\phi}L\sin\phi\,\mathbf{e}_x + \Omega L\cos\phi\,\mathbf{e}_y + L\dot{\phi}\cos\phi\,\mathbf{e}_z\,, && \text{(B.80)} \\ \mathbf{v}_Q &= -L(\dot{\phi}\sin\phi + \dot{\psi}\sin\psi)\,\mathbf{e}_x + \Omega L(\cos\phi + \cos\psi)\,\mathbf{e}_y + L(\dot{\phi}\cos\phi + \dot{\psi}\cos\psi)\,\mathbf{e}_z\,. && \text{(B.81)} \end{aligned}$$

Sostituendo queste espressioni per le velocità nel secondo membro di (B.74) ed usando anche l'espressione (B.77) per l'energia potenziale, troviamo l'espressione finale della lagrangiana $\mathscr{L}$ in funzione delle coordinate scelte:

$$\begin{aligned} \mathscr{L} = \frac{mL^2}{2}\left(2\dot{\phi}^2 + \dot{\psi}^2 + 2\dot{\phi}\dot{\psi}\cos(\phi-\psi)\right) + \frac{mL^2\Omega^2}{2}\left[(\cos\phi + \cos\psi)^2 + \cos^2\phi\right] \\ -mgL(2\sin\phi + \sin\psi) - kL^2(1 + \cos(\phi-\psi))\,. \end{aligned} \tag{B.82}$$

Le equazioni di Eulero-Lagrange risultano essere di conseguenza:

$$\begin{aligned} 2mL^2\frac{d^2\phi}{dt^2} + mL^2\cos(\phi-\psi)\frac{d^2\psi}{dt^2} = -mL^2\left(\frac{d\psi}{dt}\right)^2\sin(\phi-\psi) - mL^2\Omega^2(2\cos\phi + \cos\psi)\sin\phi \\ +kL^2\sin(\phi-\psi) - 2mgL\cos\phi\,, \end{aligned} \tag{B.83}$$

insieme a

$$\begin{aligned} mL^2\frac{d^2\psi}{dt^2} + mL^2\cos(\psi-\phi)\frac{d^2\phi}{dt^2} = -mL^2\left(\frac{d\phi}{dt}\right)^2\sin(\psi-\phi) - mL^2\Omega^2(\cos\psi + \cos\phi)\sin\psi \\ +kL^2\sin(\psi-\phi) - mgL\cos\psi\,. \end{aligned} \tag{B.84}$$

(b) Le equazioni di Eulero Lagrange scritte sopra *non* sono in forma normale: a primo membro compare una combinazione lineare delle derivate di ordine massimo e non le derivate di ordine massimo separatamente su ogni riga. Il sistema di equazioni scritto sopra può essere trascritto come:

$$mL^2\begin{bmatrix} 2 & \cos(\phi-\psi) \\ \cos(\phi-\psi) & 1 \end{bmatrix}\begin{bmatrix} \frac{d^2\phi}{dt^2} \\ \frac{d^2\psi}{dt^2} \end{bmatrix}$$

$$= \begin{bmatrix} -mL^2\left(\frac{d\psi}{dt}\right)^2 \sin(\phi-\psi) - ML^2\Omega^2(2\cos\phi+\cos\psi)\sin\phi + kL^2\sin(\phi-\psi) - 2mgL\cos\phi \\ \\ -mL^2\left(\frac{d\phi}{dt}\right)^2 \sin(\psi-\phi) - ML^2\Omega^2(\cos\psi+\cos\phi)\sin\psi + kL^2\sin(\psi-\phi) - mgL\cos\psi \end{bmatrix}$$

Se la matrice quadrata a primo membro risulta invertibile, il sistema si può scrivere in forma normale moltiplicando ambo i membri per la matrice inversa. Il determinante della matrice detta è (a parte il fattore mL^2), $2-\cos^2(\phi-\psi)$ che è sempre strettamente positivo e pertanto la matrice può essere invertita. Il calcolo con la regola di Cramer fornisce:

$$\begin{bmatrix} 2 & \cos(\phi-\psi) \\ \cos(\phi-\psi) & 1 \end{bmatrix}^{-1} = (2-\cos^2(\phi-\psi))^{-1} \begin{bmatrix} 1 & -\cos(\phi-\psi) \\ -\cos(\phi-\psi) & 2 \end{bmatrix}$$

Il sistema di equazioni di Eulero-Lagrange si riscrive in forma normale come:

$$\begin{bmatrix} \frac{d^2\phi}{dt^2} \\ \\ \frac{d^2\psi}{dt^2} \end{bmatrix} = \frac{1}{mL^2(2-\cos^2(\phi-\psi))} \begin{bmatrix} 1 & -\cos(\phi-\psi) \\ -\cos(\phi-\psi) & 2 \end{bmatrix} \times$$

$$\times \begin{bmatrix} -mL^2\left(\frac{d\psi}{dt}\right)^2 \sin(\phi-\psi) - mL^2\Omega^2(2\cos\phi+\cos\psi)\sin\phi + kL^2\sin(\phi-\psi) - 2mgL\cos\phi \\ \\ -mL^2\left(\frac{d\phi}{dt}\right)^2 \sin(\psi-\phi) - mL^2\Omega^2(\cos\psi+\cos\phi)\sin\psi + kL^2\sin(\psi-\phi) - mgL\cos\psi \end{bmatrix}$$

(c) Dato che la lagrangiana (B.82) non dipende esplicitamente dal tempo, di conseguenza, per il teorema 8.3 si conserva l'hamiltoniana:

$$\mathscr{H} = \dot\phi\frac{\partial\mathscr{L}}{\partial\dot\phi} + \dot\psi\frac{\partial\mathscr{L}}{\partial\dot\psi} - \mathscr{L}\ .$$

Il calcolo esplicito fornisce immediatamente:

$$\mathscr{H} = \frac{mL^2}{2}\left(2\dot\phi^2+\dot\psi^2+2\dot\phi\dot\psi\cos(\phi-\psi)\right) - \frac{mL^2\Omega^2}{2}\left[(\cos\phi+\cos\psi)^2+\cos^2\phi\right]$$
$$+mgL(2\sin\phi+\sin\psi) + kL^2(1+\cos(\phi-\psi))\ . \tag{B.85}$$

Il significato fisico si $\mathscr{H}$ è quello dell'energia meccanica del sistema fisico nel riferimento non inerziale $\mathscr{I}$. Infatti, dalle (B.78)-(B.79) si ricava immediatamente che l'energia cinetica del sistema dei due punti nel riferimento $\mathscr{I}$ è :

$$\mathscr{T}|_{\mathscr{I}} = \frac{mL^2}{2}\left(2\dot\phi^2+\dot\psi^2+2\dot\phi\dot\psi\cos(\phi-\psi)\right)\ .$$

Nel riferimento $\mathscr{I}$ agiscono sul sistema fisico le reazioni vincolari che non compiono complessivamente lavoro per l'idealità dei vincoli, le forze conservative di gravità e della molla che producono un'energia potenziale complessiva:

$$+mgL(2\sin\phi + \sin\psi) + kL^2(1 + \cos(\phi - \psi)) \ ,$$

le forze di Coriolis che non compiono lavoro in quanto perpendicolari ai vettori velocità , le forze centrifughe che sono perpendicolari ed uscenti dall'asse z ed hanno modulo pari a $m\Omega^2 d_P$ e $m\Omega^2 d_Q$ rispettivamente, dove d_P e d_Q sono le distanze dei punti indicati dall'asse z. Queste forze hanno la stessa struttura di forze di molle repulsive di costante $-m\Omega^2$ fissate all'asse z a punti P' e Q' che stanno sempre alla stessa quota del rispettivo altro estremo della molla P e Q. Le due forze centrifughe sono dunque forze conservative con energia potenziale totale pari a:

$$-\frac{m\Omega^2}{2}\left(d_P^2 + d_Q^2\right) = -\frac{mL^2\Omega^2}{2}\left[\cos^2\phi + (\cos\phi + \cos\psi)^2\right] \ .$$

Sommando tutti i contributi, l'energia meccanica totale nel riferimento $\mathscr{I}$ vale:

$$E|_{\mathscr{I}} = \frac{mL^2}{2}\left(2\dot\phi^2 + \dot\psi^2 + 2\dot\phi\dot\psi\cos(\phi - \psi)\right) - \frac{mL^2\Omega^2}{2}\left[(\cos\phi + \cos\psi)^2 + \cos^2\phi\right]$$

$$+mgL(2\sin\phi + \sin\psi) + kL^2(1 + \cos(\phi - \psi)) \ .$$

Pertanto essa coincide con $\mathscr{H}$ sui moti del sistema.
(d) Direttamente dal sistema di equazioni (B.83)-(B.84) si vede che le 4 curve costanti $(\phi(t), \psi(t)) = (\pm\pi/2, \pm\pi/2)$ soddisfano le equazioni del moto e le condizioni iniziali dette. Dato che le equazioni sono scrivibili in forma normale, con secondo membro di classe C^∞ vale il teorema di unicità delle soluzioni, quindi quelle trovate sono le uniche soluzioni con le condizioni iniziali imposte per ipotesi.

B.7 Esercizi del Capitolo 9.

Esercizi 9.1

9.1.**1**. *Soluzione.* (a) Per prima cosa notiamo che, direttamente dell'equazione di $\mathbb{T}^2$ pensando P come il punto materiale e $\phi = \phi(t), \theta = \theta(t)$, abbiamo che:

$$\mathbf{x} = P - O = (R + r\cos\theta)\cos\phi\,\mathbf{e}_x + (R + r\cos\theta)\sin\phi\,\mathbf{e}_y + r\sin\phi\,\mathbf{e}_z$$

e quindi:

$$\mathbf{v}_P = (R + r\cos\theta)\dot\phi(-\sin\phi\,\mathbf{e}_x + \cos\phi\,\mathbf{e}_y) + r\dot\theta\cos\theta\,\mathbf{e}_z - r\dot\theta\sin\theta\,(\cos\phi\,\mathbf{e}_x + \sin\phi\,\mathbf{e}_y) \ .$$

Ovvero:

$$\mathbf{v}_P = -\left[(R + r\cos\theta)\dot\phi\sin\phi + r\dot\theta\sin\theta\cos\phi\right]\mathbf{e}_x + \left[(R + r\cos\theta)\dot\phi\cos\phi - r\dot\theta\sin\theta\sin\phi\right]\mathbf{e}_y + r\dot\theta\cos\theta\mathbf{e}_z.$$

Eseguendo il quadrato del secondo membro e moltiplicando il risultato per $m/2$ troviamo l'energia cinetica in $\mathscr{I}$ che vale esplicitamente:

$$\mathscr{T} = \frac{m(R + r\cos\theta)^2}{2}\dot{\phi}^2 + \frac{mr^2}{2}\dot{\theta}^2 \,.$$

L'energia potenziale della molla è :

$$\frac{k}{2}(P - O)^2 = \frac{k}{2}r^2\sin^2\theta + \frac{k}{2}(R + r\cos\theta)^2 = \frac{k}{2}\left(R^2 + r^2 + 2rR\cos\theta\right) \,.$$

Possiamo ridefinire l'energia potenziale omettendo la costante additiva senza perdere informazioni:

$$\mathscr{U}(\theta, \phi) := krR\cos\theta$$

Nel caso $\gamma > 0$ la forza viscosa sarà descritta tramite le componenti lagrangiane:

$$\mathscr{Q}_\phi = \frac{\partial \mathbf{x}}{\partial \phi} \cdot (-\gamma \mathbf{v}_P) = -\gamma(R + r\cos\theta)^2\dot{\phi} \,,$$

e

$$\mathscr{Q}_\theta = \frac{\partial \mathbf{x}}{\partial \theta} \cdot (-\gamma \mathbf{v}_P) = -\gamma r^2\dot{\theta} \,.$$

Introdotta la lagrangiana parziale (che non tiene conto della forza viscosa!):

$$\mathscr{L} := \mathscr{T} - \mathscr{U} = \frac{m(R + r\cos\theta)^2}{2}\dot{\phi}^2 + \frac{mr^2}{2}\dot{\theta}^2 - krR\cos\theta \tag{B.86}$$

le due equazioni di Eulero-Lagrange non banali possono scriversi:

$$\begin{aligned} \frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\phi}} - \frac{\partial \mathscr{L}}{\partial \phi} &= \mathscr{Q}_\phi \,, \\ \frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\theta}} - \frac{\partial \mathscr{L}}{\partial \theta} &= \mathscr{Q}_\theta \,, \end{aligned}$$

che si riscrivono, esplicitamente:

$$m\frac{d}{dt}\left((R + r\cos\theta)^2\frac{d\phi}{dt}\right) = -\gamma(R + r\cos\theta)^2\frac{d\phi}{dt} \,, \tag{B.87}$$

$$mr^2\frac{d^2\theta}{dt^2} = -mr(R + r\cos\theta)\left(\frac{d\phi}{dt}\right)^2\sin\theta - \gamma r^2\frac{d\theta}{dt} + krR\sin\theta \,. \tag{B.88}$$

(b) Nel caso in cui $\gamma = 0$, la lagrangiana completa del sistema è data in (B.86). Si vede che tale lagrangiana ammette la coordinata ϕ come coordinata ciclica. In tal caso il momento coniugato associato è un integrale primo. Quindi abbiamo l'integrale primo:

$$p_\phi := \frac{\partial \mathscr{L}}{\partial \dot{\phi}} = m(R + r\cos\theta)^2\frac{d\phi}{dt} \,.$$

Dimostriamo che questa è la componente z del momento angolare rispetto al polo O in $\mathscr{I}$. Da

$$\mathbf{x} = P - O = (R + r\cos\theta)\cos\phi\,\mathbf{e}_x + (R + r\cos\theta)\sin\phi\,\mathbf{e}_y + r\sin\phi\,\mathbf{e}_z$$

e

$$\mathbf{v}_P = -\Big[(R + r\cos\theta)\dot\phi\sin\phi + r\dot\theta\sin\theta\cos\phi\Big]\,\mathbf{e}_x + \Big[(R + r\cos\theta)\dot\phi\cos\phi - r\dot\theta\sin\theta\sin\phi\Big]\,\mathbf{e}_y + r\dot\theta\cos\theta\mathbf{e}_z,$$

abbiamo facilmente che:

$$\Gamma_O|_{\mathscr{I}}\cdot\,\mathbf{e}_z = m\mathbf{x}\wedge\mathbf{v}_P\cdot\,\mathbf{e}_z = m(xv_{Py} - yv_{Px}) = m(R + r\cos\theta)^2\dot\phi\,,$$

cioé su ogni moto:

$$p_\phi = \Gamma_O|_{\mathscr{I}}\cdot\,\mathbf{e}_z = m(R + r\cos\theta)^2\dot\phi\,. \tag{B.89}$$

Per quanto riguarda l'energia meccanica totale, siamo nelle ipotesi del teorema di Jacobi dato che nelle coordinate usate (che sono adattate a $\mathscr{I}$), la lagrangiana non dipende esplicitamente dal tempo. Pertanto essa si conserva e coincide con l'hamiltoniana del sistema. Vale

$$\mathscr{H}(\phi,\theta,\dot\phi,\dot\theta) = \mathscr{T} + \mathscr{U} = \frac{m(R + r\cos\theta)^2}{2}\dot\phi^2 + \frac{mr^2}{2}\dot\theta^2 + krR\cos\theta\,. \tag{B.90}$$

(c) Dobbiamo scrivere la legge che connette le coordinate $\phi,\theta,p_\phi,p_\theta$ in funzione delle coordinate $\phi,\dot\phi,\theta,\dot\theta$, invertire questa trasformazione ed esplicitare l'hamiltoniana in (B.90) in funzione di $\phi,\theta,p_\phi,p_\theta$. Sappiamo che p_ϕ è dato in (B.89) mentre:

$$p_\theta = \frac{\partial\mathscr{L}}{\partial\dot\theta} = mr^2\dot\theta\,.$$

quindi:

$$\dot\theta = \frac{1}{mr^2}p_\theta\,, \tag{B.91}$$

$$\dot\phi = \frac{1}{m(R + r\cos\theta)^2}p_\phi\,. \tag{B.92}$$

Nelle variabili $\phi,\theta,p_\phi,p_\theta$, l'hamiltoniana (B.90) si riscrive:

$$\mathscr{H}(\phi,\theta,p_\phi,p_\theta) = \frac{p_\phi^2}{2m(R + r\cos\theta)^2} + \frac{p_\theta^2}{2mr^2} + krR\cos\theta\,.$$

Le equazioni di Hamilton sono allora:

$$\frac{dp_\phi}{dt} = \left(-\frac{\partial\mathscr{H}}{\partial\phi} =\right) 0\,, \tag{B.93}$$

$$\frac{d\phi}{dt} = \left(\frac{\partial\mathscr{H}}{\partial p_\phi} =\right) \frac{p_\phi}{m(R + r\cos\theta)^2}\,, \tag{B.94}$$

$$\frac{dp_\theta}{dt} = \left(-\frac{\partial\mathscr{H}}{\partial\theta} =\right) -\frac{p_\phi^2 r\sin\theta}{m(R + r\cos\theta)^3} + krR\sin\theta\,, \tag{B.95}$$

$$\frac{d\theta}{dt} = \left(\frac{\partial\mathscr{H}}{\partial p_\theta} =\right) \frac{p_\theta}{mr^2}\,. \tag{B.96}$$

(d) Consideriamo le equazioni di Eulero-Lagrange (B.87)- (B.88). Se inseriamo nella seconda equazione le condizioni iniziali il secondo membro si annulla. Si osservi allora che la funzione $\theta(t) = 0$ costantemente per ogni $t \in \mathbb{R}$ soddisfa sia le condizioni inziali in θ che la seconda equazione qualunque sia la soluzione dell'altra equazione. Dato che il sistema di equazioni è in forma normale con secondo membro regolare quella trovata è parte dell'unica soluzione. Con questa funzione $\theta = \theta(t)$, la prima equazione si riscrive:

$$m\frac{d}{dt}\left((R+r)^2\frac{d\phi}{dt}\right) = -\gamma(R+r)^2\frac{d\phi}{dt}\,,$$

ossia:

$$\frac{d^2\phi}{dt^2} + \frac{\gamma}{m}\frac{d\phi}{dt} = 0\,.$$

Il polinomio caratteristico associato è

$$\chi^2 + \frac{\gamma}{m}\chi = 0\,,$$

con soluzioni $\chi = 0, -\gamma/m$. La soluzione generale dell'equazione dunque è :

$$\phi(t) = A + Be^{-\gamma t/m}\,.$$

Le condizioni iniziali dicono che: $\phi(0) = A + B = 0$ da cui $A = -B$, e ancora $v = d/dt|_{t=0}\phi(t) = -\gamma Be^{-\gamma t/m}/m|_{t=0}$ da cui: $B = -mv/\gamma$. La soluzione (massimale) del problema, cioé il moto del sistema individuato dalle condizioni iniziali dette, è dunque:

$$\theta(t) = 0\,, \quad \phi(t) = \frac{mv}{\gamma}(1 - e^{-\gamma t/m})\,, \quad \text{per ogni } t \in \mathbb{R}.$$

9.1.**2**. *Soluzione.* (a) Useremo la lagrangiana $\mathscr{L}|_{\hat{\mathscr{I}}}$ relativa al riferimento inerziale $\hat{\mathscr{I}}$. Dobbiamo per prima cosa esprimere la velocità dei due punti rispetto al riferimento $\hat{\mathscr{I}}$ usando le coordinate libere x, θ, s. Usiamo un sistema di coordinate polari piane (r, θ) nel piano perpendicolare a $\mathbf{e}_x$ e passante per $O + x\,\mathbf{e}_x$. Allora

$$\mathbf{e}_r = \cos\theta\,\mathbf{e}_y + \sin\theta\,\mathbf{e}_z \qquad \mathbf{e}_\theta = -\sin\theta\,\mathbf{e}_y + \cos\theta\,\mathbf{e}_z\,.$$

Si tenga conto che $P - O = x\,\mathbf{e}_x + \cosh x\,\mathbf{e}_z$ mentre $Q - O = x\,\mathbf{e}_x + r\,\mathbf{e}_r$. Vale:

$$\mathbf{v}_P|_{\mathscr{I}} = \dot{x}\,\mathbf{e}_x + \dot{x}\sinh x\,\mathbf{e}_z\,, \quad \mathbf{v}_Q|_{\mathscr{I}} = \dot{x}\,\mathbf{e}_x + \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta\,. \tag{B.97}$$

Di conseguenza:

$$\mathbf{v}_P|_{\hat{\mathscr{I}}} = \mathbf{v}_P|_{\mathscr{I}} + \Omega\,\mathbf{e}_z \wedge (x\,\mathbf{e}_x + \cosh x\,\mathbf{e}_z) = \dot{x}\,\mathbf{e}_x + \Omega x\,\mathbf{e}_y + \dot{x}\sinh x\,\mathbf{e}_z\,,$$

mentre:

$$\mathbf{v}_Q|_{\hat{\mathscr{I}}} = \mathbf{v}_Q|_{\mathscr{I}} + \Omega\,\mathbf{e}_z \wedge (x\,\mathbf{e}_x + r\,\mathbf{e}_r) = \dot{x}\,\mathbf{e}_x + \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta + \Omega x\,\mathbf{e}_y + \Omega\,\mathbf{e}_z \wedge (r\cos\theta\,\mathbf{e}_y + r\sin\theta\,\mathbf{e}_z)$$

$= \dot{x}\,\mathbf{e}_x + \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta + \Omega x\,\mathbf{e}_y - \Omega r\cos\theta\,\mathbf{e}_x = \dot{x}\,\mathbf{e}_x + \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta + \Omega x(\cos\theta\,\mathbf{e}_r - \sin\theta\,\mathbf{e}_\theta) - \Omega r\cos\theta\,\mathbf{e}_x$.

Abbiamo trovato che:

$$\mathbf{v}_P|_{\hat{\mathscr{I}}} = \dot{x}\,\mathbf{e}_x + \Omega x\,\mathbf{e}_y + \dot{x}\sinh x\,\mathbf{e}_z\,, \tag{B.98}$$

$$\mathbf{v}_Q|_{\hat{\mathscr{I}}} = (\dot{x} - \Omega r\cos\theta)\,\mathbf{e}_x + (\dot{r} + \Omega x\cos\theta)\,\mathbf{e}_r + (r\dot{\theta} - \Omega x\sin\theta)\,\mathbf{e}_\theta\,. \tag{B.99}$$

Entrambe le velocità sono ora espresse rispetto a due terne ortonormali (diverse) di versori. L'energia cinetica in $\mathscr{T}_{\hat{\mathscr{I}}}$ vale allora, tenendo conto che $\cosh^2 x + \sinh^2 x = 1$,

$$\mathscr{T}|_{\hat{\mathscr{I}}} = \frac{m}{2}\left(\mathbf{v}_P|^2_{\hat{\mathscr{I}}} + \mathbf{v}_Q|^2_{\hat{\mathscr{I}}}\right)$$

$$= \frac{m}{2}\left(\dot{x}^2 + \Omega^2 x^2 + \dot{x}^2\sinh^2 x + (\dot{x} - \Omega r\cos\theta)^2 + (\dot{r} + \Omega x\cos\theta)^2 + (r\dot{\theta} - \Omega x\sin\theta)^2\right)\,.$$

Quindi:

$$\mathscr{T}|_{\hat{\mathscr{I}}} = \frac{m}{2}\left(\dot{x}^2(1+\cosh^2 x) + \dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{m}{2}\left(2\Omega^2 x^2 + \Omega^2 r^2\cos^2\theta\right)$$
$$-m\Omega\dot{\theta}rx\sin\theta + m(\dot{r}x - \dot{x}r)\Omega\cos\theta\,.$$

L'energia potenziale è data dalla somma dell'energia potenziale gravitazionale più quella della molla. Usando $P - Q = P - O - (Q - O) = x\,\mathbf{e}_x + \cosh x\,\mathbf{e}_z - (x\,\mathbf{e}_x + r\,\mathbf{e}_r) = \cosh x\,\mathbf{e}_z - r\,\mathbf{e}_r$ e $\mathbf{e}_r\cdot\,\mathbf{e}_z = \sin\theta$, troviamo:

$$\mathscr{U} = mg\,(z_P + z_Q) + \frac{k}{2}\,(P - Q)^2 = mg\,(\cosh x + r\sin\theta) + \frac{k}{2}\left(r^2 + \cosh^2 x - 2r\cosh x\sin\theta\right)\,.$$

La lagrangiana in $\hat{\mathscr{I}}$ ha quindi forma:

$$\mathscr{L}|_{\hat{\mathscr{I}}} = \frac{m}{2}\left(\dot{x}^2(1+\cosh^2 x) + \dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{m}{2}\left(2\Omega^2 x^2 + \Omega^2 r^2\cos^2\theta\right)$$

$$-m\Omega\dot{\theta}rx\sin\theta + m(\dot{r}x - \dot{x}r)\Omega\cos\theta - mg\,(\cosh x + r\sin\theta) - \frac{k}{2}\left(r^2 + \cosh^2 x - 2r\cosh x\sin\theta\right)\,.$$

Le tre equazioni di E-L risultano allora essere, in riferimento alle coordinate x, r, θ nell'ordine scritto:

$$m\frac{d}{dt}\left((1+\cosh^2 x)\frac{dx}{dt} - r\Omega\cos\theta\right) = m\left(\frac{dx}{dt}\right)^2\sinh x\cosh x + 2m\Omega^2 x - m\Omega\frac{d\theta}{dt}r\sin\theta + m\frac{dr}{dt}\Omega\cos\theta$$
$$-mg\sinh x - k\cosh x\sinh x + kr\sinh x\sin\theta\,,$$

$$m\frac{d}{dt}\left(\frac{dr}{dt} + x\Omega\cos\theta\right) = mr\left(\frac{d\theta}{dt}\right)^2 + m\Omega^2 r\cos^2\theta - m\Omega\frac{d\theta}{dt}x\sin\theta - m\frac{dx}{dt}\Omega\cos\theta - mg\sin\theta$$
$$-kr - k\cosh x\sin\theta\,,$$

$$m\frac{d}{dt}\left(r^2\frac{d\theta}{dt} - \Omega rx\sin\theta\right) = -m\Omega^2 r^2\cos\theta\sin\theta - m\frac{d\theta}{dt}\Omega rx\cos\theta - m\left(\frac{dr}{dt}x - \frac{dx}{dt}r\right)\Omega\sin\theta$$

$$-mgr\cos\theta + kr\cosh x\cos\theta\,.$$

(b) La forma espicita dell'hamiltoniana associata alla lagrangiana $\mathscr{L}|_{\hat{\mathscr{I}}}$ riferita a coordinate x, r, θ è :

$$\mathscr{H} = \dot{x}\frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{x}} + \dot{r}\frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{r}} + \dot{\theta}\frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{\theta}} - \mathscr{L}\,,$$

ossia, esplicitamente:

$$\mathscr{H} = \frac{m}{2}\left(\dot{x}^2(1+\cosh^2 x) + \dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{m\Omega^2}{2}\left(x^2 + (x^2 + r^2\cos^2\theta)\right) + mg\left(\cosh x + r\sin\theta\right)$$
$$+\frac{k}{2}\left(r^2 + \cosh^2 x - 2r\cosh x\sin\theta\right)\,. \tag{B.100}$$

Dato che $\mathscr{L}|_{\hat{\mathscr{I}}}$ non dipende esplicitamente dal tempo, il teorema di Jacobi assicura che $\mathscr{H}$ si conservi sulle soluzioni delle equazioni di E-L. Il significato di

$$\frac{m}{2}\left(\dot{x}^2(1+\cosh^2 x) + \dot{r}^2 + r^2\dot{\theta}^2\right)$$

è quello di energia cinetica nel riferimento $\mathscr{I}$ come si ricava direttamente dalle (B.97). Il terzo ed il quarto addendo a secondo membro in (B.100) rappresentano l'energia potenziale della gravità e della molla come abbiamo calcolato precedentemente. Il secondo addendo

$$-\frac{m\Omega^2}{2}\left(x^2 + (x^2 + r^2\cos^2\theta)\right)$$

si può invece scrivere come $-\frac{m\Omega^2}{2}d_P^2 - \frac{m\Omega^2}{2}d_Q^2$. $d_P = |x_P|$ è la distanza di P dall'asse z e d_Q è la distanza tra Q e l'asse z. d_Q è la lunghezza della proiezione nel piano $z = 0$ del vettore $Q - O = x\,\mathbf{e}_x + r\,\mathbf{e}_r = x\,\mathbf{e}_x + r\cos\theta\,\mathbf{e}_y + r\sin\theta\,\mathbf{e}_z$; pertanto il quadrato di d_Q è $x^2 + r^2\cos^2\theta$. Come sappiamo

$$-\frac{m\Omega^2}{2}d_P^2 - \frac{m\Omega^2}{2}d_Q^2$$

è l'energia potenziale della forza centrifuga. La forza di Coriolis non compie lavoro nel riferimento $\mathscr{I}$ perché è perpendicolare alla velocità . Le reazioni vincolari non compiono lavoro complessivo essendo i vincoli ideali ed indipendenti dal tempo nel riferimento $\mathscr{I}$. In definitiva il secondo memebro dell'espressione trovata sopra per $\mathscr{H}$ è proprio l'energia meccanica totale nel riferimento $\mathscr{I}$.

(c) Abbiamo esplicitamente:

$$p_x = \frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{x}} = m\dot{x}(1+\cosh^2 x) - mr\Omega\cos\theta\,, \tag{B.101}$$

$$p_r = \frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{r}} = m\dot{r} + mx\Omega\cos\theta\,, \tag{B.102}$$

$$p_\theta = \frac{\partial\mathscr{L}|_{\hat{\mathscr{I}}}}{\partial\dot{\theta}} = mr^2\dot{\theta} - m\Omega xr\sin\theta\,. \tag{B.103}$$

Le formule inverse risultano essere:

$$\dot{x} = \frac{p_x + mr\Omega\cos\theta}{m(1+\cosh^2 x)}\,, \tag{B.104}$$

$$\dot{r} = \frac{p_r - mx\Omega\cos\theta}{m}\,, \tag{B.105}$$

$$\dot{\theta} = \frac{p_\theta + m\Omega xr\sin\theta}{mr^2}\,. \tag{B.106}$$

Sostituendo in (B.100) si trova:

$$\mathscr{H} = \frac{(p_x + mr\Omega\cos\theta)^2}{2m(1+\cosh^2 x)} + \frac{(p_r - mx\Omega\cos\theta)^2}{2m} + \frac{(p_\theta + m\Omega xr\sin\theta)^2}{2mr^2} - \frac{m\Omega^2}{2}\left(2x^2 + r^2\cos^2\theta\right)$$

$$+mg\left(\cosh x + r\sin\theta\right) + \frac{k}{2}\left(r^2 + \cosh^2 x + 2r\cosh x\sin\theta\right)\,. \tag{B.107}$$

Le equazioni di Hamilton della forma: $\frac{dq^k}{dt} = \frac{\partial\mathscr{H}}{\partial p_k}$ forniscono immediatamente le equazioni:

$$\frac{dx}{dt} = \frac{p_x + mr\Omega\cos\theta}{m(1+\cosh^2 x)}\,,$$

$$\frac{dr}{dt} = \frac{p_r - mx\Omega\cos\theta}{m}\,,$$

$$\frac{d\theta}{dt} = \frac{p_\theta + m\Omega xr\sin\theta}{mr^2}\,.$$

Le rimanenti equazioni di Hamilton: $\frac{dp_k}{dt} = -\frac{\partial\mathscr{H}}{\partial q^k}$ sono allora:

$$\frac{dp_x}{dt} = \frac{(p_x + mr\Omega\cos\theta)^2}{m(1+\cosh^2 x)^2}\sinh x\cosh x + \Omega\cos\theta(p_r - mx\Omega\cos\theta) - \Omega r\sin\theta\frac{p_\theta + m\Omega xr\sin\theta}{r^2}$$

$$+2m\Omega^2 x - mg\sinh x - k\cosh x\sinh x + kr\sinh x\sin\theta\,,$$

$$\frac{dp_r}{dt} = \frac{(p_\theta + m\Omega xr\sin\theta)^2}{mr^3} - \Omega\cos\theta\frac{p_x + mr\Omega\cos\theta}{(1+\cosh^2 x)} - \Omega x\sin\theta\frac{p_\theta + m\Omega xr\sin\theta}{r^2}$$

$$+m\Omega^2 r\cos^2\theta - mg\sin\theta - kr - k\cosh x\sin\theta\,,$$

$$\frac{dp_\theta}{dt} = -\left(x(p_r - mx\Omega\cos\theta) - r\frac{p_x + mr\Omega\cos\theta}{1+\cosh^2 x}\right)\Omega\sin\theta - \Omega x\cos\theta\frac{p_\theta + m\Omega xr\sin\theta}{r}$$

$$-m\Omega^2 r^2\cos\theta\sin\theta - mgr\cos\theta + kr\cosh x\cos\theta\,.$$

B.8 Esercizi dell'Appendice A.

Esercizi A.1

A.1.**1**. *Traccia di soluzione.* Per costruire l'atlante $\{(U_i, \phi_i)\}_{i\in I}$ su $\mathbb{S}^2$, considerare sistemi di coordinate polari sferiche con asse polare z disposto in varie direzioni e usando, in luogo della coordinata radiale r, la nuova coordinata $r' := r - 1$.

A.1.**2**. *Traccia di soluzione.* La risposta ad entrambe le domande è negativa: N non è localmente omeomorfa a $\mathbb{R}^1$ (cosa succede attorno al punto $(0, 0)$?).

A.1.**3**. *Traccia di soluzione.* Il vertice del cono $(0, 0, 0)$ non ammette, nel suo intorno, sistemi di coordinate del tipo di quelli nella condizione (ii) della definizione A.4. Tuttavia una struttura differenziabile per C si ottiene dalla carta globale che proietta i punti di C sul piano $x^3 = 0$ e associa alle proiezioni le loro coordinate (x^1, x^2). Eliminando il vertice del cono, C^* risulta avere struttura di sovarietà embedded in $\mathbb{R}^3$ di dimensione 2 ed una carta globale si ottiene proprio dalla proiezione sulla base del cono, come detto sopra.

A.1.**4**. *Traccia di soluzione.* Usare il fatto che le carte locali dei due atlanti sono restrizioni di carte locali di M tra di loro compatibili.

A.1.**5**. *Soluzione.* Nelle ipotesi fatte, per il teorema A.1, per ogni punto p nell'aperto $\phi(U\cap V)$, c'è un intorno aperto $O \subset \phi(U\cap V)$ di p che è trasformato in un aperto $O' \subset \psi(U\cap V)$ da $\psi\circ\phi^{-1}$. Inoltre, su O', l'inversa di $(\psi\circ\phi^{-1})\restriction_U$, g è ben definita e di classe C^k. Dato che l'unica l'inversa di $\psi\circ\phi^{-1} : \phi(U\cap V) \to \psi(U\cap V)$ per ipotesi è $\phi\circ\psi^{-1} : \psi(U\cap V) \to \phi(U\cap V)$, g deve coincidere con la restrizione di quest'ultima a O'. Dato che, al variare di p, gli intorni O' ricoprono $\psi(U\cap V)$, concludiamo che la restrizione di $\phi\circ\psi^{-1} : \psi(U\cap V) \to \phi(U\cap V)$ ad un'intorno (O') di ogni suo punto nel dominio $\psi(U\cap V)$ è di classe C^k. Ma allora $\phi\circ\psi^{-1} : \psi(U\cap V) \to \phi(U\cap V)$ è di classe C^k.

Esercizi A.2

A.2.**1**. Si considerino le coordinate locali $x^1, \ldots, x^n$ della carta (U, ϕ) e le coordinate locali $y^1, \ldots, y^n$ della carta (V, ψ) con $U \cap V \neq \emptyset$. L'identità

$$id_{\phi(U\cap U)} = (\phi\circ\psi^{-1})\circ(\psi\circ\phi^{-1}) ,$$

usando le coordinate ed essendo $\psi\circ\phi^{-1}$ scrivibile come $y^i = y^i(x^1, \cdots, x^n)$ e $\phi\circ\psi^{-1}$ scrivibile come $x^i = x^i(y^1, \cdots, y^n)$, si riscrive:

$$x^i = x^i\left(y^1(x^1, \ldots, x^2), \ldots y^n(x^1, \ldots, x^2)\right) , \quad i = 1, \ldots, n . \tag{B.108}$$

Ammettendo che entrambe le funzioni $\psi\circ\phi^{-1}$ e $\phi\circ\psi^{-1}$ siano differenziabili (almeno C^1) come segue dalla richiesta di compatibilità delle due carte, ed applicando la regola di derivazione di funzioni composte in (B.108), si ha, derivando rispetto a x^k:

$$\delta^i_k = \sum_{j=1}^n \frac{\partial x^i}{\partial y^j}\frac{\partial y^j}{\partial x^k} .$$

Quest'equazione dice che il prodotto delle due matrici jacobiane quadrate che appaiono a secondo membro è la matrice identica. Quindi entrambe le matrici quadrate sono invertibili ed, in particolare, hanno determinante non nullo.

A.2.**2**. *Soluzione.* Passando dalle coordinate locali $x^1, \ldots, x^n$ della carta (U, ϕ) attorno a p alle coordinate locali $y^1, \ldots, y^n$ della carta (V, ψ) attorno a p in cui il requisito di non singolarità è soddisfatto, vale:

$$\frac{\partial f \circ \phi^{-1}}{\partial x^i}|_{\phi^{-1}(p)} = \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i}|_{\phi^{-1}(p)} \frac{\partial f \circ \psi^{-1}}{\partial y^j}|_{\psi^{-1}(p)} \,.$$

Dato che la matrice di coefficienti $\frac{\partial y^j}{\partial x^i}|_{\phi^{-1}(p)}$ è non singolare per l'esercizio precedente, se il vettore riga di componenti $\frac{\partial f \circ \psi^{-1}}{\partial y^j}|_{\psi^{-1}(p)}$ è non nullo, deve essere non nullo anche quello di coefficienti $\frac{\partial f \circ \phi^{-1}}{\partial x^i}|_{\phi^{-1}(p)}$.

A.2.**3**. *Traccia di soluzione.* Si generalizzi il ragionamento che ha portato alla soluzione dell'esercizio precedente.

Esercizi A.3

A.3.**1**. *Soluzione.* Sia $\mathbf{u} \in V$. Per (i) in definizione A.12 deve esistere $R \in \mathbb{A}^n$ tale che $R - P = u$. Usando (ii), $\mathbf{u} + (P - P) = (R - P) + (P - P) = R - P = \mathbf{u}$. Per l'unicità dell'elemento neutro additivo $\mathbf{0}$ in V, deve essere $P - P = \mathbf{0}$.

A.3.**2**. *Soluzione.* Usando (ii) in definizione A.12 si ha: $((Q + \mathbf{u}) + \mathbf{v}) - Q = [((Q + \mathbf{u}) + v) - (Q + \mathbf{u})] + [(Q + \mathbf{u}) - Q]$. Per definizione il primo ed il secondo addendo a secondo membro sono rispettivamente: $\mathbf{v}$ e $\mathbf{u}$. Abbiamo ottenuto che $((Q + \mathbf{u}) + \mathbf{v}) - Q = \mathbf{u} + \mathbf{v}$. Per definizione si ha allora che: $(Q + \mathbf{u}) + \mathbf{v} = Q + (\mathbf{u} + \mathbf{v})$.

A.3.**3**. *Soluzione.* Usando (ii) in definizione A.12 si ha che: $P - Q + Q - P = P - P = \mathbf{0}$ da cui la tesi.

A.3.**4**. *Soluzione.* Usando (ii) in definizione A.12 si ha:

$$(P + \mathbf{u}) - (Q + \mathbf{u}) = [(P + \mathbf{u}) - P] + [P - (Q + \mathbf{u})] = [(P + \mathbf{u}) - P] + [P - Q][Q - (Q + \mathbf{u})] \,.$$

Applicando la definizione di $P + \mathbf{u}$ e $Q + \mathbf{u}$ nell'ultimo membro, tenendo conto dell'esercizio precedente, troviamo:

$$(P + \mathbf{u}) - (Q + \mathbf{u}) = \mathbf{u} + [P - Q] - \mathbf{u} = P - Q \,.$$

Esercizi A.4

A.4.**1**. *Soluzione.* Per costruzione

$$\mathbf{e}_i = \sum_j B^j{}_i \mathbf{e}'_j \,,$$

dove la matrice B di elementi $B^j{}_i$ è non singolare. Se $\mathbf{v} \in V$ ed esso si decompone sulle due basi dette come $\mathbf{v} = \sum_i v^i \mathbf{e}_i = \sum_j v'^j \mathbf{e}'_j$, esplicitando $\mathbf{e}_i$ in funzione degli $\mathbf{e}'_j$ si ricava:

$$\sum_{i,j} v^i B^j{}_i \mathbf{e}'_j = \sum_j v'^j \mathbf{e}'_j \,.$$

Da cui

$$\sum_j \left(v'^j - \sum_{i=1}^n B^j{}_i v^i \right) \mathbf{e}'_j = 0 \, .$$

Sfruttando l'indipendenza lineare dei vettori $\mathbf{e}'_j$ si hanno infine le *relazioni di trasformazione* delle componenti di $\mathbf{v}$ tra le due basi considerate:

$$v'^j = \sum_{i=1}^n B^j{}_i v^i \, . \tag{B.109}$$

Da queste relazioni possiamo facilmente ricavare la legge di trasformazione tra i due differenti sistemi di coordinate cartesiane f e g. Consideriamo ora un punto $P \in \mathbb{A}^n$, possiamo scrivere, in virtù delle proprietà degli spazi affini:

$$P - O' = (P - O) + (O - O') \tag{B.110}$$

Le componenti di $P - O$, $x^1, \cdots, x^n$, sulla base $\{\mathbf{e}_i\}_{i=1,\ldots,n}$ sono le coordinate di P nella carta globale $(\mathbb{A}^n, f)$, mentre quelle di $P - O'$, $x'^1, \cdots, x'^n$, nella base $\{\mathbf{e}'_j\}_{j=1,\ldots,n}$ sono le coordinate di P nell'altra carta globale $(\mathbb{A}^n, g)$. In base alla (B.109) si ha immediatamente da (B.110):

$$(P - O')'^j = \sum_{i=1}^n B^j{}_i (P - O')^i = \sum_{i=1}^n B^j{}_i (P - O)^i + \sum_{i=1}^n B^j{}_i (O - O')^i \, ,$$

da cui

$$x'^j = \sum_{i=1}^n B^j{}_i (x^i + b^i),$$

dove $(O - O') = \sum_i b^i \mathbf{e}_i$.

A.4.**2**. *Soluzione.* $f \circ g^{-1}$ non è altro che la relazione che esprime le coordinate $x^1, \cdots, x^n$ in funzione delle coordinate $x'^1, \cdots, x'^n$. Essa si ottiene invertendo la funzione (A.31). Vale, per definizione di matrice inversa, $\sum_j (B^{-1})^k{}_j B^j{}_i = \delta^k{}_i$ dove $\delta^k{}_i$ è il solito *delta di Kroneker* definito da $\delta^k{}_i = 0$ se $k \neq 1$ e $\delta^k{}_k = 1$. Moltiplicando entrambi i membri di (A.31) per $(B^{-1})^k{}_j$ e sommando su j si ottiene:

$$\sum_j (B^{-1})^k{}_j x'^j = \sum_{i,j} (B^{-1})^k{}_j B^j{}_i (x^i + b^i) = \sum_{i,j} \delta^k{}_i (x^i + b^i) = x^k + b^k \, ,$$

da cui segue immediatamente (A.32).

A.4.**3**. Siano $O_1 \in \mathbb{A}^n$ e $O_2 \in \mathbb{A}^m$ due origini per sistemi di coordinate cartesiane e $\{\mathbf{e}_1, \cdots, e_n\} \subset V_1$ e $\{f_1, \cdots, f_m\} \subset V_2$ basi negli spazi delle traslazioni. Indichiamo con $x_1^1, \cdots, x_1^n$ e $x_2^1, \cdots, x_2^m$ le coordinate cartesiane rispettivamente associate in $\mathbb{A}^n$ e $\mathbb{A}^m$. Se $x_1^1, \cdots, x_1^n$ sono le coordinate di $P \in \mathbb{A}^n$, vale $P - O_1 = \sum_j x_1^j \mathbf{e}_j$ e quindi:

$$\psi(P) = \psi\left(O_1 + \sum_j x_1^j \mathbf{e}_j \right) = \left[\psi\left(O_1 + \sum_j x_1^j \mathbf{e}_j \right) - \psi(O_1) \right] + \psi(O_1) \, .$$

La funzione in parentesi quadre è una funzione lineare $d\psi(\sum_j x_1^j \mathbf{e}_j)$ per ipotesi, per cui:

$$\psi(P) = \sum_j x_1^j d\psi(\mathbf{e}_j) + \psi(O_1)\,.$$

Quindi

$$\psi(P) - O_2 = \sum_j x_1^j d\psi(\mathbf{e}_j) + (\psi(O_1) - O_2) = \sum_j x_1^j L(\mathbf{e}_j) + c\,,$$

dove $c := \psi(O) - O_2 \in V_2$. Possiamo decomporre tutti vettori che compaiono sopra sulla base $f_1, \cdots, f_m \subset V_2$:

$$\sum_i (\psi(P) - O_2)^i f_i = \sum_j x_1^j \sum_i L^i(\mathbf{e}_j) f_i + \sum_i c^i f_i\,.$$

Per definizione $x_2^i := (\psi(P) - O_2)^i$ e, posto $L^i{}_j := d\psi^i(\mathbf{e}_j)$, l'identità di sopra si riscrive:

$$\sum_i \left[x_2^i - \left(\sum_j L^i{}_j x_1^j + c^i \right) \right] f_i = 0\,.$$

L'indipendenza lineare dei vettori f_i prova la validità della (A.33). Il fatto che una trasformazione $\psi : \mathbb{A}^n \to \mathbb{A}^m$ sia affine se è della forma (A.33) quando rappresentata in coordinate cartesiane è di immediata verifica diretta.

A.4.**4**. Vale:

$$\psi(P(t)) = \psi(P(t)) - \psi(P) + \psi(P) = \psi(P + tu) - \psi(P) + \psi(P) = d\psi(tu) + \psi(P) = \psi(P) + t d\psi(u)\,.$$

Esercizi A.5

A.5.**1**. *Traccia di soluzione.* Sia $||\mathbf{u}|| := \sqrt{\mathbf{u} \cdot \mathbf{u}}$ è la norma associata al prodotto scalare. Dalle proprietà di simmetria e linearità del prodotto scalare segue che:

$$\mathbf{u} \pm \mathbf{v} \cdot \mathbf{u} \pm \mathbf{v} = (\mathbf{u} \cdot \mathbf{u})^2 + (\mathbf{v} \cdot \mathbf{v})^2 \pm 2(\mathbf{u} \cdot \mathbf{v})\,,$$

da cui

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2}||\mathbf{u} + \mathbf{v}||^2 + \frac{1}{2}||\mathbf{u} - \mathbf{v}||^2\,.$$

Dato che $d(O, O + \mathbf{w}) := ||\mathbf{w}||$ la tesi segue immediatamente.

A.5.**2**. *Soluzione.* Ovviamente la conservazione del prodotto scalare implica quella delle distanze, per cui dobbiamo provare solo che la conservazione delle distanze implica la conservazione del prodotto scalare. Dalle proprietà del prodotto scalare (reale simmetrico) vale

$$(\mathbf{u}|\mathbf{v})_1 = \frac{1}{2}[(\mathbf{u} + \mathbf{v}|\mathbf{u} + \mathbf{v})_1 - (\mathbf{u} - \mathbf{v}|\mathbf{u} - \mathbf{v})_1]\,,$$

per cui il prodotto scalare tra due vettori $P - Q$ e $P' - Q$, definendo $R \in \mathbb{E}_1^3$ t.c. $R - Q = (P - Q) + (P' - Q)$, vale

$$(P - Q|P' - Q)_1 = \frac{1}{2}[(R - Q|R - Q)_1 - (P - P'|P - P')_1] = \frac{1}{2}[d_1(R, Q)^2 - d_1(P, P')^2]\,.$$

Dato che ϕ è affine:

$$\phi(R) - \phi(Q) = d\psi(R - Q) = d\psi(P - Q) + d\psi(P' - Q) = (\phi(P) - \phi(Q)) + (\phi(P') - \phi(Q)) \,.$$

Come conseguenza del fatto che ϕ conserva le distanze, si ha che

$$(P - Q|P' - Q)_1 = \frac{1}{2}[d_1(R, Q)^2 - d_1(P, P')^2] = \frac{1}{2}[d_2(\phi(R), \phi(Q))^2 - d_2(\phi(P), \phi(P'))^2]$$

L'ultimo membro è proprio $(\phi(P) - \phi(Q)|\phi(P') - \phi(Q))_2$.

A.5.**3**. *Soluzione*. Se ϕ è espressa in coordinate dalla (A.41), allora, nelle basi ortonormali associate a tali distanze, la rappresentazione matriciale della trasformazione lineare $d\phi$ è semplicemente data dalla matrice R di coefficienti $R^i{}_j$. Essendo tale matrice ortogonale, la conservazione dei prodotti scalari segue immediatamente. Dall'esercizio precedente segue che ϕ conserva le distanze. Supponiamo viceversa che ϕ sia affine e conservi le distanze, quindi conserverà i prodotti scalari per l'esercizio precedente. In coordinate cartesiane ortonormali la trasformazione affine ϕ (esercizio A.4.**3**) sarà rappresentata da:

$$x_2^i = \sum_{j=1}^n R^i{}_j x_1^j + b^i \,,$$

da cui, la trasformazione lineare $d\phi : V_1 \to V_2$ si esprime come, nelle stesse basi usate per definire le coordinate cartesiane:

$$w_2^i = \sum_{j=1}^n R^i{}_j w_1^j$$

dove $\mathbf{w}_2 = d\phi(\mathbf{w}_1)$ con $\mathbf{w}_1 \in V_1$ e $\mathbf{w}_2 \in V_2$. Tenendo conto che le basi usate sono ortonormali, la conservazione dei prodotti scalari si esprime con la relazione

$$\sum_{i=1}^n v_2^i u_2^i = \sum_{j=1}^n v_1^j u_1^j \,.$$

Usando la rappresentazione di $d\phi$ data sopra, si ottengono le condizioni:

$$\sum_{i=1}^n v_1^k R^i{}_k R^i{}_j u_1^j = \sum_{j=1}^n v_1^j u_1^j$$

ossia

$$\sum_{i,j,k=1}^n v_1^k (R^i{}_k R^i{}_j - \delta_{kj}) u_1^j = 0 \,.$$

Scegliendo tutti i coefficienti v_1^k e u_1^j nulli eccetto una coppia di essi, e facendo variare tali coppie in tutti i modi possibili, si trovano le relazioni, valide per ogni scelta di k e j:

$$\sum_{i=1}^n R^i{}_k R^i{}_j = \delta_{kj}$$

che in forma matriciale si riscrivono

$$R^t R = I\,.$$

Si osservi che questa relazione equivale a $RR^t = I$ ($R^t R = I$ significa che R^t è un'inversa sinistra di R, ma, dato che la matrice R è quadrata, l'inversa sinistra coincide con l'inversa destra ed è unica, per cui vale anche $RR^t = I$. Il fatto che $RR^t = I$ implichi $R^t R = I$ si prova analogamente). Quindi R è una matrice ortogonale $n \times n$ reale.

A.5.**4**. *Soluzione.* Proviamo (i), (ii) e (iii). Se $\phi : \mathbb{E}_1^n \to \mathbb{E}_2^n$ conserva le distanze allora è affine per il teorema A.4, allora si esprime come (A.41) in coordinate cartesiane ortonormali. Dato che la trasformazioni lineari non omogenee sono C^∞, ϕ è di classe C^∞. Dato che le matrici ortogonali sono invertibili, l'inversa di (A.41) si scrive subito come:

$$x_1^j = \sum_{i=1}^n (R^{-1})^j{}_i x_2^i + c^j\ ,\ \text{ con } c^j = \textstyle\sum_{i=1}^n (R^{-1})^j{}_i b^i.$$

Dato che la matrice inversa di una matrice ortogonale è ancora ortogonale, per l'esercizio precedente, la trasformazione inversa di ϕ è ancora una trasformazione affine che conserva le distanze, in particolare è ancora C^∞ e quindi ϕ è un diffeomorfismo. La composizione di due isometrie affini, in coordinate cartesiane ortonormali $x_2^i = \sum_{j=1}^n R^i{}_j\, x_1^j + b^i$ e $x_3^i = \sum_{j=1}^n R'^i{}_j x_2^j + b'^i$ produce banalmente la trasformazione:

$$x_3^i = \sum_{j=1}^n (R'R)^i{}_j x_1^j + \left(\sum_{j=1}^n R'^i{}_j b^j + b'^i\right)$$

Dato che le matrici ortogonali formano un gruppo $R'R$ è ortogonale e pertanto la trasformazione di sopra è ancora un'isometria per l'esercizio A.5.**3**.
La prova di (iv) è immediata tenendo conto che, dall'esercizio A.5.**3**, la rappresentazione matriciale di $d\phi$ è una matrice ortogonale quando ci si riferisce alla coppia di basi che danno luogo alle coordinate cartesiane ortonormali in cui si rappresenta ϕ.

A.5.**5**. *Soluzione.* Per l'esercizio precedente l'insieme delle isometrie affini da $\mathbb{E}^n$ a $\mathbb{E}^n$ è chiuso rispetto alla composizione ed all'inversa, inoltre la trasformazione identica è banalmente un'isometria. Di conseguenza, rispetto alla composizione di funzioni (che è associativa) l'insieme delle isometrie affini su un fissato spazio euclideo è un gruppo. Essendo le trasformazioni affini degli isomorfismi, tale gruppo è sottogruppo rispetto al gruppo degli isomorfismi di $\mathbb{E}^n$ in se stesso.

A.5.**6**. *Soluzione.* Se f è un'isometria affine allora è banalmente un'isometria. Per dimostrare che se f è isometria allora è un'isometria affine, è sufficiente dimostrare che f è una trasformazione affine. Abbiamo bisogno del seguente risultato preliminare.

Lemma. *Sia $\phi : V_1 \to V_2$, con V_1 e V_2 spazi vettoriali di dimensione finita n, dotati di prodotti scalari (reali simmetrici) $(\cdot|\cdot)_1$ e $(\cdot|\cdot)_2$ rispettivamente, che generano le norme $||\cdot||_1$ e $||\cdot||_2$ rispettivamente. ϕ è lineare e conserva il prodotto scalare se e solo se valgono le due condizioni insieme:*

(i) $\phi(0_1) = 0_2$*, essendo* 0_i *il vettore nullo di* V_i*,* $i = 1, 2$*;*
(ii) $||\phi(\mathbf{v}) - \phi(\mathbf{u})||_2 = ||\mathbf{v} - \mathbf{u}||_1$ *per ogni coppia* $\mathbf{u}, \mathbf{v} \in V_1$*.*

Dimostrazione. È chiaro che se ϕ è lineare e conserva il prodotto scalare, allora le condizioni (i) e (ii) sono verificate. Proviamo che se valgono (i) e (ii) allora ϕ conserva il prodotto scalare. Per ogni $\mathbf{w} \in V_1$

$$||\phi(\mathbf{w})||_2 = ||\phi(\mathbf{w}) - 0_2||_2 = ||\phi(\mathbf{w}) - \phi(0_1)||_2 = ||\mathbf{w} - \mathbf{0}||_1 = ||\mathbf{w}||_1 \,.$$

D'altra parte $||\phi(\mathbf{u}) - \phi(\mathbf{v})||_2^2 = ||u - v||_1^2$ si esplicita come:

$$||\phi(\mathbf{u})||_2^2 + ||\phi(v)||_2^2 + 2(\phi(\mathbf{u})|\phi(\mathbf{v}))_2 = ||\mathbf{u}||_1^2 + ||\mathbf{v}||_1^2 + 2(\mathbf{u}|\mathbf{v})_2$$

e pertanto, usando $||\phi(\mathbf{w})||_2 = ||\mathbf{w}||_1$, segue che $(\phi(\mathbf{u})|\phi(\mathbf{v}))_2 = (\mathbf{u}|\mathbf{v})_2$, che è quanto volevamo provare. Per concludere, dimostriamo che ϕ è lineare. Sia $\{\mathbf{e}_i\}_{i=1,\ldots,n}$ base ortonormale in V_1. Per la conservazione del prodotto scalare $\{\phi(\,\mathbf{e})_i\}_{i=1,\ldots,n}$ è una base ortonormale in V_2. Quindi, per ogni $\mathbf{v} \in V_1$, decomponendo $\phi(\mathbf{v})$ sulla base $\{\phi(\,\mathbf{e})_i\}_{i=1,\ldots,n}$ ed infine usando nuovamente la conservazione del prodotto scalare:

$$\phi(\mathbf{v}) = \sum_{i=1}^{n} (\phi(\mathbf{e}_i)|\phi(v))\phi(\mathbf{e}_i) = \sum_{i=1}^{n} (\mathbf{e}_i|v)\phi(\mathbf{e}_i) \,.$$

La linearità in v dei prodotti scalari $(\mathbf{e}_i|v)$ implica immediatamente che ϕ è lineare. □

Torniamo alla dimostrazione che ogni isometria è una trasformazione affine. Fissiamo $O \in \mathbb{E}_1^n$. Ogni vettore $v \in V_1$ si potrà scrivere, biunivocamente, come $\mathbf{v} = P_\mathbf{v} - O$ per $P_\mathbf{v} \in \mathbb{E}_1^n$. In particolare $\mathbf{0}_1 = O - O$. Definiamo $\phi : V_1 \to V_2$ tale che $\phi(\mathbf{v}) := f(P_\mathbf{v}) - f(O)$. In questo modo $\phi(\mathbf{0}_1) = f(O) - f(O) = \mathbf{0}_2$, inoltre

$$||\phi(\mathbf{u}) - \phi(\mathbf{v})||_2^2 = ||(f(P_\mathbf{u}) - f(O)) - (f(P_\mathbf{v}) - f(O))||_2^2 = ||f(P_\mathbf{u}) - f(P_\mathbf{v})||_2^2 \,,$$

ma l'ultimo membro uguaglia:

$$d_2(f(P_\mathbf{u}), f(P_\mathbf{v}))^2 = d_1(P_\mathbf{u} - P_\mathbf{v})^2 = ||\mathbf{u} - \mathbf{v}||_1^2 \,.$$

Sono dunque verificate (i) e (ii) del lemma. Applicando il lemma, concludiamo che ϕ è lineare (e conserva i prodotti scalari). Infine per la definizione di ϕ e la sua linearità :

$$f(P_\mathbf{v}) - f(P_\mathbf{u}) = f(P_\mathbf{v}) - f(O) - (f(P_\mathbf{u}) - f(O)) = \phi(\mathbf{u}) - \phi(\mathbf{v}) = \phi(\mathbf{u} - \mathbf{v}) = \phi(P_\mathbf{u} - P_\mathbf{v}) \,.$$

Dato che $P_\mathbf{w}$ spazia in tutto $\mathbb{E}_1^n$ al variare di $\mathbf{w} \in V_1$, l'identità di sopra si riscrive:

$$f(P) - f(Q) = \phi(P - Q) \,, \text{ per ogni coppia } P, Q \in \mathbb{E}^n,$$

dove ϕ è una funzione lineare. Questa identità implica che f sia invariante per traslazioni e sia una trasformazione affine con $df = \phi$.

Bibliografia

[Arnold92] V.I. Arnold, *Metodi Matematici della Meccanica Classica* II ed., Ed. Riuniti, Roma, 1992

[doCarmo92] M.P.do Carmo, *Riemannian Geometry*, Birkhäuser, Boston 1992

[Fasano-Marmi] A. Fasano, S. Marmi *Meccanica analitica*, Bollati Boringhieri, Torino, 2002

[Goldstein50] H. Goldstein, *Classical mechanics*, Addison-Wesley Cambridge, MA, (1950)

[GiustiI] E. Giusti *Analisi Matematica 1*, Bollati Boringhieri, Torino, 2003

[GiustiII] E. Giusti *Analisi Matematica 2*, Bollati Boringhieri, Torino, 2003

[Jackson] J.D. Jackson, *Elettrodinamica Classica*, Zanichelli, Trento, 1988

[Malkin] I.G. Malkin, *Teoria della stabilità del moto*, Gostezchidat, Mosca, 1952.

[Sernesi90] E. Sernesi, *Geometria 2*, Bollati Boringhieri, Torino, 1990

[Westenholtz78] C. von Westenholtz, *Differential Forms in Mathematical Physics*; North-Holland, Amsterdam, (1978)

Valter Moretti
Dipartimento di Matematica
Facoltà di Scienze M.F.N
Università di Trento

Fisica Matematica II (Introduzione alla teoria delle equazioni alle Derivate Parziali del secondo ordine)

Corso di *Fisica Matematica II* per la Laurea Triennale in Matematica
Facoltà di Scienze MFN, Università di Trento,
anno accademico 2007-2008

Indice

1 Introduzione alle equazioni differenziali a derivate parziali del secondo ordine quasi lineari. 5

1.1 Introduzione. 5
1.2 Notazioni e convenzioni. 5
1.3 Qualche motivazione fisico matematica ed un esempio. 6
1.3.1 Equazioni di Maxwell in forma integrale. 7
1.3.2 Teoremi di Gauss, Stokes ed equazioni di Maxwell in forma differenziale locale. 8
1.4 Classificazione delle equazioni differenziali del secondo ordine quasilineari 13
1.5 Il problema di Cauchy ed il Teorema di Cauchy-Kovalevskaja. 20
1.5.1 Superfici regolari in $\mathbb{R}^n$. 20
1.5.2 Il problema di Cauchy e la "ben posizione" del problema nel senso di Hadamard. 20
1.5.3 Il Teorema di Cauchy-Kovalevskaja. 22
1.5.4 Superfici caratteristiche. 26

2 Equazioni Ellittiche e funzioni armoniche in $\mathbb{R}^n$: risultati elementari. 31

2.0.5 Il problema fisico dell'elettrostatica e le equazioni di Poisson e Laplace. . 31
2.1 Principio del massimo per funzioni armoniche e principio del massimo generalizzato. 33
2.1.1 Funzione armoniche e sub armoniche in $\mathbb{R}^n$. 33
2.1.2 Principio del massimo (in forma debole). 36
2.1.3 Principio del massimo generalizzato. 37
2.1.4 Due teoremi di unicità per il probelma di Dirichlet dal principio del massimo. 39
2.2 Le identità di Green le loro conseguenze elementari. 41
2.2.1 Identità di Green. 42
2.2.2 Conseguenze del teorema di Gauss e delle identità di Green: teorema di unicità per il problema di Neumann. 42

3 Soluzioni fondamentali per l'equazione di Poisson in $\mathbb{R}^n$ e risultati ad esse legati. 48

3.1 Soluzioni fondamentali. 48

3.1.1 Proprietà elementari delle soluzioni fondamentali. 49
3.2 Ulteriori proprietà delle funzioni armoniche in $\mathbb{R}^n$. 54
3.2.1 Non esistenza di funzioni armoniche con supporto compatto. 55
3.2.2 Analiticità delle funzioni armoniche in $\mathbb{R}^n$. 56
3.2.3 Teorema della media e principio del massimo in forma forte. 59
3.2.4 Teorema di Liouville per le funzioni armoniche in $\mathbb{R}^n$. 63

4 Funzioni di Green e costruzione di soluzioni del problema di Dirichlet. 65
4.1 Ancora sul problema di Dirichlet. 65
4.1.1 Funzioni di Green e nuclei di Poisson. 66
4.2 Funzioni di Green per domini particolari. 73
4.2.1 Il metodo delle cosiddette *cariche immagine*. 73
4.2.2 La funzione di Green nella palla in $\mathbb{R}^3$. 74
4.2.3 La funzione di Green nel cerchio in $\mathbb{R}^2$. 77
4.2.4 La funzione di Green in un semispazio di $\mathbb{R}^3$. 79
4.3 *Calcolo del nucleo di Poisson per il problema del cerchio in $\mathbb{R}^2$ tramite l'analisi di Fourier 80

5 Equazioni iperboliche: alcuni risultati generali elementari per le equazioni di D'Alembert e di Klein-Gordon in $\mathbb{R} \times \mathbb{R}^n$. 85
5.1 L'equazione di D'Alembert come equazione della corda vibrante. 86
5.2 Condizioni iniziali ed al contorno. 87
5.3 Bilancio energetico e teoremi di unicità. 88
5.3.1 Densità di energia ed equazione di continuità. 88
5.3.2 Teoremi di unicità. 90

6 Equazione di D'Alembert e di Klein-Gordon in $\mathbb{R} \times \mathbb{R}$. 95
6.1 Equazione di D'Alembert sulla retta reale senza condizioni al contorno. 95
6.1.1 Assenza di sorgenti, formula di D'Alembert, domini di dipendenza. 95
6.1.2 Equazione di D'Alembert con sorgente. 102
6.2 Dalla separazione delle variabili alla serie di Fourier. 104
6.3 Alcuni risultati elementari sulla serie di Fourier. 107
6.3.1 La serie di Fourier nello spazio di Hilbert $L^2([-L/2, L/2], dx)$. 108
6.3.2 Convergenza uniforme della serie di Fourier e derivazione sotto il simbolo di serie. 110
6.4 Il problema su $\mathbb{R} \times [-L/2, L/2]$ con condizioni al bordo periodiche. 115
6.4.1 Teorema di unicità. 115
6.4.2 Esistenza delle soluzioni per dati iniziali sufficientemente regolari. 117
6.5 Il problema su $\mathbb{R} \times [-L/2, L/2]$ con condizioni al bordo di annullamento 123
6.5.1 Teorema di unicità. 123
6.5.2 Esistenza delle soluzioni per dati iniziali sufficientemente regolari. 124

A Un accenno all'approccio moderno per il problema ellittico: soluzioni in senso debole e teoremi di regolarità ellittica. 131

B Limite e derivazione sotto il segno integrale e di serie dalla teoria della misura. 135

B.1 Teoremi della convergenza monotona e dominata. 136

B.2 Derivazione sotto il segno di integrale e di serie. 140

Capitolo 1

Introduzione alle equazioni differenziali a derivate parziali del secondo ordine quasi lineari.

1.1 Introduzione.

Queste dispense sono relative al corso di Fisica Matematica II, tenuto dall'autore per il corso di Laurea Triennale in Matematica presso Facoltà di Scienze MFN dell'Università di Trento. Lo scopo del corso è di introdurre gli studenti ai primi rudimenti della teoria delle equazioni differenziali a derivate parziali. In particolare equazioni ellittiche ed iperboliche. I prerequisiti riguardano l'analisi matematica delle funzioni di più variabili e alcune nozioni elementari di teoria delle funzioni di variabile complessa.
L'autore ringrazia tutti gli studenti e le persone che hanno segnalato errori di vario genere ed hanno contribuito a migliorare le dispense con suggerimenti vari. Vorrei ringraziare in particolare (in ordine alfabetico): F. Franceschini e G. Stecca.

1.2 Notazioni e convenzioni.

Se f è una funzione definita sullo spazio topologico X, per esempio $\mathbb{R}^n$ oppure un sottoinsieme di $\mathbb{R}^n$ dotato della topologia indotta da $\mathbb{R}^n$, ed i valori di f sono assunti in $\mathbb{R}^n$, allora il **supporto di** f è , come ben noto, l'insieme:

$$\operatorname{supp} f := \overline{\{x \in X \mid f(x) \neq 0\}}\,,$$

dove, 0 indica il vettore nullo (o semplicemente lo zero se $n = 1$) e la chiusura è riferita alla topologia di X.

Definizione 1.1. Siano $n, m = 1, 2, \ldots$ e $k = 0, 1, \ldots$ fissati e sia $\Omega \subset \mathbb{R}^n$ un insieme aperto e non vuoto.

(a) Una funzione $f : \Omega \to \mathbb{R}^m$ o $\mathbb{C}^n$ è detta essere **di classe** C^k, e si scrive in tal caso $f \in C^k(\Omega; \mathbb{R}^n)$ o $C^k(\Omega; \mathbb{C}^n)$, rispettivamente, se tutte le derivate parziali (incluse quelle miste) delle componenti di f esistono e sono continue fino all'ordine k incluso. Si pone $C^k(\Omega) := C^k(\Omega; \mathbb{R})$.
(b) $f : \Omega \to \mathbb{R}^m$ o $\mathbb{C}^m$ è detta di **classe** C^∞ se è di classe C^k per ogni $k = 0, 1, \ldots$ e si definisce:

$$C^\infty(\Omega; \mathbb{R}^n) := \bigcap_{k=0,1,\ldots} C^k(\Omega; \mathbb{R}^n)\,, \quad C^\infty(\Omega; \mathbb{C}^n) := \bigcap_{k=0,1,\ldots} C^k(\Omega; \mathbb{C}^n)\,.$$

Si pone $C^\infty(\Omega) := C^\infty(\Omega; \mathbb{R})$.
(c) In riferimento alle definizioni già date $C^k_0(\Omega; \mathbb{C}^n)$ (rispettivamente $C^k_0(\Omega)$), con $k = 0, 1, \ldots, \infty$, indica l'insieme delle funzioni in $C^k(\Omega; \mathbb{C}^n)$ (rispettivamente $C^k(\Omega)$) il cui *supporto*, riferito a Ω, è compatto. $\diamondsuit$

Diamo separatamente un'altra importante definizione che riguarda le funzioni differenziabili su un insieme ottenuto dalla chiusura di un aperto $\overline{\Omega}$ o, più in generale, su un insieme qualsiasi F.

Definizione 1.2. Siano $n, m = 1, 2, \ldots$ e $k = 0, 1, \ldots$ fissati e sia $F \subset \mathbb{R}^n$ un insieme qualsiasi (non vuoto). $f : F \to \mathbb{R}^m$ o $\mathbb{C}^m$ è detta essere **di classe** C^k, con $k = 0, 1, \ldots, \infty$ e si scrive in tal caso $f \in C^k(F; \mathbb{R}^n)$ (oppure $f \in C^k(F)$ se $m = 1$), quando f è la restrizione di una funzione $g : \Omega \to \mathbb{R}^m$ di classe C^k, dove $\Omega \supset F$ è un aperto non vuoto. $\diamondsuit$.

Nel caso in cui $F = [a, b] \subset \mathbb{R}$, si può provare che $f \in C^k(F)$ (o $f \in C^k(F; \mathbb{C})$ se la funzione è a valori complessi), se e solo se $f \in C^k((a, b)) \cap C^0([a, b])$ (risp. $f \in C^k((a, b); \mathbb{C}) \cap C^0([a, b]; \mathbb{C})$) ed esistono finiti i limiti delle derivate, fino all'ordine k, in b e in a rispettivamente, tali limiti risultano coincidere con le derivate sinistre in b e destre in a rispettivamente.

1.3 Qualche motivazione fisico matematica ed un esempio.

In fisica molto spesso le leggi che descrivono la dinamica di un certo sistema fisico, ed in particolare certe grandezze differenziabili dipendenti dal tempo e dal posto, sono date in termini di *equazioni differenziali alle derivate parziali* tra queste grandezze. Una tale equazione è una (o più) relazione tra le derivate (in generale di ordine arbitrario) nello spazio e nel tempo delle grandezze considerate. Ci si aspetta che, assegnando qualche ulteriore informazione (dati iniziali e/o al contorno), le equazioni considerate ammettano una ed una sola soluzione.
L'importanza in fisica delle equazioni differenziali alle derivate parziali è evidente studiando la storia della fisica. Tale teoria permette di dare una formalizzazione adeguata delle equazioni fondamentiali che riguardano i sistemi fisici che hanno un'estensione non puntiforme e sono dunque descritti da funzioni del posto e del tempo (densità, campi di velocità, campi di forze,...). Tali sistemi non puntifirmi sono presenti in vari rami della fisica come la meccanica dei mezzi continui, la fluidodinamica, la teoria dei campi classica, relativistica e quantistica-relativistica.
Un capitolo importante sia dal punto di vista fisico che da quello matematico è quello che riguarda la teoria classica dell'*eletromagnetismo*. Nel diciannovesimo secolo la formalizzazione teorica

completa della teoria Elettromagnetica fu data da J. C. Maxwell. Le sue quattro equazioni differenziali lineari alle derivate parziali per il campo elettrrico $\mathbf{E} = \mathbf{E}(t, \mathbf{x})$ ed il campo magnetico $\mathbf{B}(t, \mathbf{x})$:

$$\begin{cases} \nabla \cdot \mathbf{E}(t, \mathbf{x}) = 4\pi\rho(t, \mathbf{x}) \\ \nabla \wedge \mathbf{E}(t, \mathbf{x}) = -\dfrac{1}{c}\dfrac{\partial \mathbf{B}}{\partial t}(t, \mathbf{x}) \\ \nabla \cdot \mathbf{B}(t, \mathbf{x}) = 0 \\ \nabla \wedge \mathbf{B}(t, \mathbf{x}) = \dfrac{4\pi}{c}\mathbf{J}(t, \mathbf{x}) + \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t}(t, \mathbf{x}) \end{cases} \tag{1.1}$$

assieme alle legge di conservazione della carica e alla forza di Lorentz, descrivono completamente (nell'ambito della fisica classica) il sistema fisico dato dal campo elettromagnetico ed dalle sue sorgenti $\rho, \mathbf{J}$ (vedi oltre). L'esempio delle equazioni di Maxwell è interessante dal punto di vista didattico perché permette di introdurre diversi operatori differenziali ed alcuni teoremi generali che vengono adoperati in tutta la teoria classica delle equazioni differenziali alle derivate parziali. Nel seguito riassumeremo alcuni aspetti matematici delle equazioni di Maxwell.

1.3.1 Equazioni di Maxwell in forma integrale.

In presenza di campi elettromagnetici $\mathbf{E} = \mathbf{E}(t, \mathbf{x})$ e $\mathbf{B}(t, \mathbf{x})$ assegnati in un riferimento $\mathscr{I}$ con coordinate cartesiane solidali $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ e coordinata temporale $t \in \mathbb{R}$, una carica q è sottoposta ad una forza detta di **forza Lorentz** descritta da, se $\mathbf{x}(t)$ e $\mathbf{v}(t)$ sono la posizione e la velocità della carica nel riferimento $\mathscr{I}$ al tempo t:

$$\mathbf{F}(t, \mathbf{x}(t), \mathbf{v}(t)) = q\,\mathbf{E}(t, \mathbf{x}(t)) + \frac{q}{c}\mathbf{v}(t) \wedge \mathbf{B}(t, \mathbf{x}(t))\,. \tag{1.2}$$

Sopra c è la velocità della luce. In riferimento alle coordinate di $\mathscr{I}$ suddette, in ogni regione spaziotemporale aperta nella quale sono definiti i campi $\mathbf{E}$ e $\mathbf{B}$, valgono le celeberrime **equazioni di Maxwell nel vuoto**:

$$\begin{cases} \displaystyle\oint_{+\partial V} \mathbf{E} \cdot \mathbf{n}\, dS = 4\pi \int_V \rho\, d^3x\,, \\ \displaystyle\oint_C \mathbf{E} \cdot d\mathbf{x} = -\frac{1}{c}\frac{d}{dt}\int_{\Sigma_C} \mathbf{B} \cdot \mathbf{n}\, dS\,, \\ \displaystyle\oint_{+\partial V} \mathbf{B} \cdot \mathbf{n}\, dS = 0\,, \\ \displaystyle\oint_C \mathbf{B} \cdot d\mathbf{x} = \frac{4\pi}{c}\int_{\Sigma_C} \mathbf{J} \cdot \mathbf{n}\, dS + \frac{1}{c}\frac{d}{dt}\int_{\Sigma_C} \mathbf{E} \cdot \mathbf{n}\, dS\,, \end{cases} \tag{1.3}$$

dove, V è un volume in quiete in $\mathscr{I}$ il cui bordo ∂V è una superficie chiusa regolare orientabile (cioè il vettore normale è definibile con continuità su tutta la supercie e non si annulla mai) ed il versore normale è indicato con $\mathbf{n}$. L'orientazione di $\mathbf{n}$ è uscente, come indicato dal segno $+$ davanti a ∂V. C è una curva regolare chiusa, in quiete in $\mathscr{I}$, che è il bordo della superficie regolare Σ. Il versore normale a Σ è orientato con la legge della mano destra rispetto al vettore tangente a C. $\rho = \rho(t, \mathbf{x})$ e $\mathbf{J}(t, \mathbf{x})$ sono rispettivamente la densità (volumetrica) di carica

elettrica (pertanto la carica complessiva presente, all'istante t, in un fissato volume si ottiene integrando la funzione $\rho(t, \mathbf{x})$ nella variabile $\mathbf{x}$ nel volume detto) ed il vettore densità di corrente elettrica. Quest'ultimo può essere pensato come della forma $\mathbf{J}(t, \mathbf{x}) = \rho(t, \mathbf{x})\mathbf{V}(t, \mathbf{x})$, dove $\mathbf{V}$ è il campo di velocità delle cariche elettriche con densità di carica ρ. Se sono presenti più tipi di portatori di carica (per esempio elettroni e ioni), questa forma elementare della densità corrente deve essere modificata. La regolarità dei campi e delle densità è supposta tale da dare senso alle equazioni scritte. Le equazioni di maxwell scritte sopra sono in forma integrale. Mostreremo tra poco come trascriverle in forma di equazioni differenziali alle derivate parziali. Per fare ciò dobbiamo ricordare qualche teorema di analisi elementare.

1.3.2 Teoremi di Gauss, Stokes ed equazioni di Maxwell in forma differenziale locale.

Ricordiamo che una funzione misurabile $f : A \to \mathbb{C}$ (o $\mathbb{R}$), con $A \subset \mathbb{R}^n$ misurabile, è detta **integrabile secondo Lebesgue**, oppure equivalentemente **assolutamente integrabile secondo Lebesgue** oppure equivalentemente **Lebesgue-integrabile**, se

$$\int_A |f| d^n x < +\infty \,,$$

dove $d^n x$ denota la misura di Lebesgue su $\mathbb{R}^n$. L'insieme delle funzioni Lebesgue integrabili si denita con $\mathscr{L}^1(A)$. Supporremo nota la teoria elementare della misura di Lebesgue in $\mathbb{R}^n$ e le sue relazioni con la teoria dell'integrazione secondo Riemann.
Ricordiamo ora i teoremi di Gauss (della divergenza) e di Stokes.

Teorema 1.1. **(di Gauss)**. *Sia Ω un aperto non vuoto di $\mathbb{R}^n$, la cui chiusura $\overline{\Omega}$ è compatta*[1] *e tale che il suo bordo $\partial\Omega$ sia una superficie regolare orientabile con versore normale $\mathbf{n}$ orientato in maniera uscente. Se $\mathbf{F} : \overline{\Omega} \to \mathbb{R}^n$ è di classe $C^1(\Omega; \mathbb{R}^n) \cap C^0(\overline{\Omega}; \mathbb{R}^n)$, allora vale la* **formula di Gauss***:*

$$\oint_{+\partial\Omega} \mathbf{F} \cdot \mathbf{n} \, dS = \int_\Omega \nabla \cdot \mathbf{F} \, d^n x \,, \tag{1.4}$$

dove il secondo integrale è un integrale di Lebesgue se $\nabla \cdot \mathbf{F}$ è Lebesgue-integrabile *su Ω (e ciò accade in particolare se $\mathbf{F} \in C^1(\overline{\Omega}; \mathbb{R}^n)$), altrimenti è da intendersi come un integrale improprio nel senso di Riemann.* $\diamondsuit$

Teorema 1.2. **(di Stokes)**. *Sia $C \subset \mathbb{R}^3$ una curva regolare orientata, bordo della superficie regolare Σ_C con $\mathbf{n}$ vettore normale a Σ_C orientato con la legge della mano destra rispetto al senso di percorrenza di C. Sia $\Omega \supset \Sigma \cup C$ un insieme aperto e limitato. Se $\mathbf{F} : \Omega \to \mathbb{R}^3$ è un campo di classe $C^1(\Omega; \mathbb{R}^3)$, allora vale la* **formula di Stokes***:*

$$\oint_C \mathbf{F} \cdot d\mathbf{x} = \int_{\Sigma_C} \nabla \wedge \mathbf{F} \cdot \mathbf{n} \, dS \,. \tag{1.5}$$

[1] Questo equivale a richiedere che l'insieme aperto non vuoto Ω sia limitato.

$\diamondsuit$

Osservazioni 1.1.
(1) La procedura per provare il teorema di Gauss nelle ipotesi indebolite data sopra è la seguente [7]. Il teorema di Gauss viene inizialmente provato nel caso in cui entrambi gli integrali esistono e sono ben definiti, cioè quando $\mathbf{F} \in C^1(\overline{\Omega}, \mathbb{R}^n)$. L'insieme Ω è misurabile secondo Lebesgue, con misura finita essendo aperto e limitato. La chiusura è anch'essa Lebesgue-misurabile ed ha la stessa misura di Ω, in quando il bordo di Ω ha misura nulla (essendo una superficie $n-1$ dimensionale regolare). La misura di Peano-Jordan-Riemann di e $\overline{\Omega}$ coincide con la misure di Lebesgue di Ω. In questo caso l'identità (1.4) è verificata interpretando il secondo membro come integrale di Riemann su $\overline{\Omega}$ oppure, equivalentemente, come integrale di Lebesgue su Ω: essendo $\nabla \cdot \mathbf{F}$ continuo su $\overline{\Omega}$, i due tipi di integrali sono ben definiti e coincidono.
Quindi si passa ad indebolire l'ipotesi di regolarità di $\mathbf{F}$ sul bordo di Ω. In questa ipotesi più debole si considera una successione di aperti $\{\Omega_m\}_{m\in\mathbb{N}}$, a chiusura compatta e bordo regolare orientabile, che soddisfino: $\overline{\Omega_m} \subset \Omega_{m+1} \subset \Omega$ e con $\partial\Omega_m$ incluso in un intorno aperto[2] di $\partial\Omega$ di raggio ϵ_m con $\epsilon_m \to 0$ se $m \to +\infty$. Evidentemente $\cup_{m\in\mathbb{N}}\Omega_m = \Omega$.
Per ogni Ω_m, l'identità (1.4) è verificata interpretando il secondo membro come integrale di Riemann su $\overline{\Omega_m}$ oppure, equivalentemente, come integrale di Lebesgue su Ω_m dato che $\mathbf{F} \in C^1(\overline{\Omega_m}, \mathbb{R}^n)$. Quindi si considerano i limiti:

$$\lim_{m\to+\infty} \oint_{+\partial\Omega_m} \mathbf{F}\cdot\mathbf{n}\, dS = \lim_{m\to+\infty} \int_{\Omega_m} \nabla\cdot\mathbf{F}\, d^n x\,.$$

Dato che $\mathbf{F}$ è continua su $\overline{\Omega}$, si prova che il limite di sinistra esiste e coincide con

$$\oint_{+\partial\Omega} \mathbf{F}\cdot\mathbf{n}\, dS\,.$$

In questo l'interpretazione di (1.4) è :

$$\oint_{+\partial\Omega} \mathbf{F}\cdot\mathbf{n}\, dS = \lim_{m\to+\infty} \int_{\Omega_m} \nabla\cdot\mathbf{F}\, d^n x\,. \tag{1.6}$$

Notare che il limite del secondo membro non dipende dalla classe degli Ω_m scelti, purché soddisfino le ipotesi dette sopra. In questo caso il secondo membro dell'identità di Gauss è interpretato come integrale di Riemann improprio.
Nel caso in cui $\nabla\cdot\mathbf{F}$ sia anche (assolutamente) integrabile nel senso di Lebesgue su Ω, possiamo dire di più . Definiamo $\chi_{\Omega_m} : \Omega \to \mathbb{R}$ come: $\chi_{\Omega_m}(x) = 1$ per $x \in \chi_{\Omega_m}$ e $\chi_{\Omega_m}(x) = 0$ altrimenti. Dato che su Ω vale:

$$|\chi_{\Omega_m}(x)\nabla\cdot\mathbf{F}(x)| \le |\nabla\cdot\mathbf{F}(x)|\,,$$

[2]Se $A \subset \mathbb{R}^n$, un intorno aperto di raggio $\epsilon > 0$ di A è l'insieme dato dall'unione di tutte le palle aperte di raggio ϵ, $B_\epsilon(x)$ centrate in $x \in A$.

il *teorema della convergenza dominata* (vedi la sezione B.2 in appendice) permette di concludere che:

$$\lim_{m\to+\infty}\int_\Omega \chi_{\Omega_m}\nabla\cdot\mathbf{F}\,d^nx = \lim_{m\to+\infty}\int_{\Omega_m}\nabla\cdot\mathbf{F}\,d^nx = \int_\Omega \nabla\cdot\mathbf{F}\,d^nx\,.$$

In riferimento a (1.6), possiamo ora dire che vale la (1.4) dove il secondo membro dell'identità di Gauss è interpretato come integrale di Lebesgue.
(2) Il teorema di Stokes si potrebbe enunciare con i potesi molto più deboli, ma non ce ne occuparemo in questa sede.

Usando questi teoremi nelle equazioni di Maxwell in forma integrale, si arriva facilmente a provare che, assumendo i campi di classe C^1 nelle 4 variabili congiuntamente, essi soddisfano le *equazioni di Maxwell in forma differenziale locale*, per ogni punto ed istante:

$$\begin{cases} \nabla\cdot\ \mathbf{E}(t,\mathbf{x}) = 4\pi\rho(t,\mathbf{x}) \\ \nabla\wedge\ \mathbf{E}(t,\mathbf{x}) = -\dfrac{1}{c}\dfrac{\partial\,\mathbf{B}}{\partial t}(t,\mathbf{x}) \\ \nabla\cdot\ \mathbf{B}(t,\mathbf{x}) = 0 \\ \nabla\wedge\ \mathbf{B}(t,\mathbf{x}) = \dfrac{4\pi}{c}\mathbf{J}(t,\mathbf{x}) + \dfrac{1}{c}\dfrac{\partial\,\mathbf{E}}{\partial t}(t,\mathbf{x}) \end{cases} \tag{1.7}$$

A titolo di esempio, usando il teorema di Gauss, la prima equazione di Maxwell in forma integrale può essere riscritta:

$$\int_V \nabla\cdot\ \mathbf{E}\,d^3x = 4\pi\int_V \rho\,d^3x\,,$$

da cui

$$\int_{V_r}(\nabla\cdot\ \mathbf{E} - 4\pi\rho)\ d^3x = 0\,, \tag{1.8}$$

per ogni palla aperta V_r di raggio finito $r>0$ centrata in $\mathbf{x}_0$. Se valesse

$$(\nabla\cdot\ \mathbf{E}(t,\mathbf{x}_0) - 4\pi\rho(t,\mathbf{x}_0)) = I > 0\,,$$

troveremmo una contraddizione. Infatti, per continuità , scegliendo r sufficientemente piccolo, l'integrando assumerebbe in V_r valori in $(I-\epsilon, I+\epsilon)$ con $I-\epsilon>0$, e pertanto

$$\int_{V_r}(\nabla\cdot\ \mathbf{E} - 4\pi\rho)\ d^3x > (I-\epsilon)\frac{4\pi r^3}{3} > 0\,,$$

che contraddirebbe l'ipotesi (1.8). Si ottiene la stessa contraddizione assumendo $I<0$. Concludiamo che, per ogni punto $(t,\mathbf{x})\in I\times\Omega$ in cui vale la prima equazione di Maxwell in forma integrale, deve valere anche la prima equazione di Maxwell in forma differenziale. Viceversa, se vale la prima equazione in forma differenziale, integrandola su un qualunque insieme V con frontiera ∂V sufficientemente regolare contenuto nel dominio spaziale di validità delle equazioni, ed usando il teorema di Gauss, si ottiene subito la prima equazione di Maxwell in forma integrale su tale volume V. La terza equazione di Maxwell in forma differenziale si ottiene dalla

terza equazione in forma integrale con la stessa procedura. Le rimanenti due si ricavano dalle corrispondenti equazioni integrali, con una analoga procedura, ma usando il teorema di Stokes in luogo del teorema della divergenza. A titolo di esempio consideriamo la seconda equazione in forma integrale:

$$\oint_C \mathbf{E} \cdot d^3x = -\frac{1}{c}\frac{d}{dt}\int_{\Sigma_C} \mathbf{B} \cdot \mathbf{n}\, dS\,.$$

se il campo $\mathbf{B}$ è C^1 (congiuntamente in tutte le variabili), tale equazione si riscrive, passando sotto il segno di integrale la derivata nel tempo:

$$\oint_C \mathbf{E} \cdot d\mathbf{x} + \int_{\Sigma_C} \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}\, dS = 0\,.$$

Usando il teorema di Stokes, dopo aver assunto anche il campo $\mathbf{E}$ di calsse C^1, si arriva quindi all'identità:

$$\int_{\Sigma_C} \left(\nabla \wedge \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}\right) \cdot \mathbf{n}\, dS = 0\,.$$

Ammettiamo per assurdo che, al fissato tempo t_0 e nel punto $\mathbf{x}_0$, valga:

$$\nabla \wedge \mathbf{E}|_{(t_0,\mathbf{x}_0)} + \frac{1}{c}\left.\frac{\partial \mathbf{B}}{\partial t}\right|_{(t_0,\mathbf{x}_0)} = \mathbf{c} \neq 0\,.$$

Scegliamo C come una circonferenza centrata in $\mathbf{x}_0$ nel piano normale a $\mathbf{c}$, e come Σ_C il cerchio associato a tale circonferenza, in modo tale che $\mathbf{c} \cdot \mathbf{n} = I > 0$. Con queste ipotesi deve essere

$$\int_{\Sigma_C} \left(\nabla \wedge \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}\right) \cdot \mathbf{n}\, dS = 0\,, \tag{1.9}$$

malgrado, esattamente al centro del cerchio, l'integrando sia strettamente positivo per ipotesi. Lo stesso ragionamento che abbiamo usato per il teorema della divergenza, basato sulla permanenza del segno di una funzione continua, produce una contraddizione. Infatti, dato che l'integrando in (1.9) è continuo, il suo segno rimarrà costante in un intorno di $\mathbf{x}_0$, quindi scegliendo il raggio $R > 0$ del cerchio Σ_C sufficientemente piccolo, in modo che su ogni punto del cerchio:

$$\int_{\Sigma_C} \left(\nabla \wedge \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}\right) \cdot \mathbf{n}\, dS > \pi R^2 (I - \epsilon) > 0\,,$$

in contraddizione con la seconda equazione di Maxwell in forma integrale. Pertanto per ogni $(t_0, \mathbf{x}_0)$ deve valere la seconda equazione di Maxwell in forma differenziale:

$$\nabla \wedge \mathbf{E}|_{(t_0,\mathbf{x}_0)} + \frac{1}{c}\left.\frac{\partial \mathbf{B}}{\partial t}\right|_{(t_0,\mathbf{x}_0)} = 0\,.$$

Per ottenere la quarta ed ultima equazione di Maxwell in forma differenziale dalla corrispondente in forma integrale si procede nello stesso modo. È evidente che procedendo in senso inverso nei

ragionamenti, lavorando in tutto lo spazio $\mathbb{R}^3$, le equazioni in forma differenziale implicano quelle in forma integrale. Tuttavia le equazioni in forma differenziale richiedono ipotesi più forti sulla regolarità dei campi.

Osservazioni 1.2.
(1) Se $I \subset \mathbb{R}$ è un intervallo aperto e $\Omega \in \mathbb{R}^3$ un aperto, per campi $\mathbf{E}$ e $\mathbf{B}$ di classe $C^1(I \times \Omega; \mathbb{R}^3)$ e sorgenti $\rho \in C^0(I \times \Omega)$, $\mathbf{J} \in C^0(I \times \Omega; \mathbb{R}^3)$, le equazioni di Maxwell in forma differenziale sono *conseguenza* di quelle integrali assunte valide nello stesso insieme $I \times \Omega$. Nel caso $\Omega = \mathbb{R}^3$ i due set di equazioni sono *equivalenti*. Nel caso locale l'equivalenza dei due set di equazioni si ha solo se Ω ha una struttura topologica opportuna (vedi osservazione 2.1 più avanti).
(2) Le equazioni di Maxwell in forma differenziale, se si assumono i campi elettrico e magnetico di classe C^2 e le sorgenti ρ e $\mathbf{J}$ di classe C^1, implicano facilmente la validità dell'equazione che vincola le sorgenti dei campi:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \,. \tag{1.10}$$

(Per ottenere tale equazione è sufficiente calcolare la divergenza dei due membri dell'ultima equazione di Maxwell in forma differenziale, tenere conto del fatto che $\nabla \cdot \nabla \wedge = 0$ ed infine usare nel risultato ottenuto l'identità data dalla prima equazione di Maxwell differenziale.) Questa equazione è detta **equazione di continuità della carica elettrica** ed esprime matematicamente la legge di conservazione locale della carica elettrica. Essa ha un'equivalente forma integrale che si ottiene integrando i due membri dell'equazione su un volume $V \subset \mathbb{R}^3$ dato da un aperto a chiusura compatta dal bordo dato dalla superficie chiusa regolare ed orientabile ∂V, applicando il teorema della divergenza ed, infine, portando la derivata nel tempo fuori dal segno di integrazione spaziale:

$$\frac{d}{dt} \int_V \rho(t, \mathbf{x}) d^3x = \oint_{+\partial V} \mathbf{J}(t, \mathbf{x}) \cdot \mathbf{n} \, dS \,.$$

Questa equazione dice che la variazione per unità di tempo della carica totale nel volume V è pari al flusso uscente della densità di corrente attraverso la frontiera di V, all'istante considerato. Procedendo in senso contrario nelle implicazioni e tenendo conto dell'arbitarietà del volume V si verifica che l'identità di sopra equivale alla (1.10) se i campi ρ e $\mathbf{J}$ sono di classe C^1.
(3) Nelle ipotesi di campi elettrico e magnetico di classe C^2, le quattro equazioni di Maxwell in forma differenziale, nelle regioni spaziotemporali in cui le sorgenti sono nulle, implicano la validità dell'*equazione delle onde di D'Alembert* per ogni componente del campo elettromagnetico:

$$-\frac{1}{c^2}\frac{\partial^2 E^i}{\partial t^2} + \Delta E^i = 0 \,, \quad \text{per } i = 1, 2, 3$$

e

$$-\frac{1}{c^2}\frac{\partial^2 B^i}{\partial t^2} + \Delta B^i = 0 \,, \quad \text{per } i = 1, 2, 3.$$

Per ottenere tali equazioni, in assenza delle sorgenti ρ e $\mathbf{J}$, è sufficiente partire, rispettivamente dalla quarta e dalla seconda equazione di Maxwell differenziale, derivarne i due membri rispetto

a t e quindi usare nel primo membro le identità date, rispettivamente dalla seconda e dalla terza equazione. Si deve quindi usare l'identità operatoriale:

$$\nabla \wedge (\nabla \wedge \mathbf{A}) = -\Delta \mathbf{A} + \nabla(\nabla \cdot \mathbf{A}) \,,$$

usando infine il fatto, dato dalla prima e terza equazione di Maxwell, che $\nabla \cdot \mathbf{E} = 0$ e $\nabla \cdot \mathbf{B} = 0$. Queste equazioni implicano che, in assenza di sorgenti e nel vuoto, i campi descrivano propagazioni ondose che si propagano alla velocità della luce c (la luce è un'onda elettromagnetica!).
(4) Nel ventesimo secolo la teoria di Maxwell ha avuto ulteriori sorprendenti sviluppi, infatti nel 1905 l'articolo di Einstein sulla relatività ristretta ha sottolineato la covarianza relativistica delle equazioni di Maxwell, ovvero il fatto che queste sono valide in un qualsiasi sistema di riferimento inerziale pur di cambiare profondamente la struttura geometrica dello spaziotempo ed entrando nella teoria della relatività speciale. Lo sviluppo della teoria da un punto di vista quantistico inizia negli anni venti con l'equazione di Dirac e culmina qualche decennio più tardi con quella parte della teoria dei campi quantizzati nota come elettrodinamica quantistica. Non bisogna assolutamente pensare però che tutto ciò che c'era da scoprire è stato ormai compreso: l'elettodinamica quantistica è una teoria che, pur avendo ottenuto dei successi sperimentali sorprendenti, manca ancora di una formalizzazione matematica rigorosa che vada oltre la teoria perturbativa.

1.4 Classificazione delle equazioni differenziali del secondo ordine quasilineari

In questo corso ci concentreremo essenzialmente su una classe di equazioni differenziali che andiamo a descrivere. Un'equazione differenziale su $\Omega \subset \mathbb{R}^n$ aperto non vuoto, a derivate parziali (PDE) del secondo ordine ed in forma **quasi lineare**, è un'equazione della forma:

$$\sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \Phi(x, u(x), \nabla_x u) \,, \tag{1.11}$$

dove $u = u(x)$, con $x := (x^1, \ldots, x^n)$, è la funzione reale incognita da determinarsi. Si suppone $u \in C^2(\Omega)$ mentre le funzioni reali *assegnate* a^{ij} e Φ sono (almeno) di classe C^0 rispettivamente su Ω su $\Omega \times \mathbb{R} \times \mathbb{R}^n$. Infine $\nabla_x u$ denota in gradiente della funzione u.
Ovviamente è supposto che le funzioni a^{ij} non siano tutte identicamente nulle su Ω (in tal caso non avrebbe senso chiamare l'equazione di sopra "del secondo ordine"). La matrice $A(x)$ i cui coefficienti sono i numeri $a^{ij}(x)$ si dice **matrice caratteristica**, nel punto x, dell'equazione (1.11).
L'equazione (1.11) è detta **lineare** quando ha la forma specifica:

$$\sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^{n} b^k(x) \frac{\partial u}{\partial x^k} + c(x)u(x) = f(x) \,, \tag{1.12}$$

dove, a parità delle altre condizioni, le funzioni *assegnate* b^k, c e f sono (almeno) di classe C^0 su Ω. Nel caso in cui la funzione f è identicamente nulla su Ω, l'equazione linerare si dice **omogenea**.
Di particolare interesse è il caso in bidimensionale in cui $\Omega \subset \mathbb{R}^2$, useremo in questo caso le coordinate (x, y) in luogo di (x^1, x^2). In questa situazione l'equazione (1.11) si riscrive:

$$a(x,y)u_{xx} + 2b(x,y)u_{xy} + c(x,y)u_{yy} + \Phi(x,y,u(x,y),u_x,u_y) = 0\,,$$

dove con u_x e u_y sono rispettivamente $\partial u/\partial x$ e $\partial u/\partial y$ mentre con u_{xx} abbiamo indicato la derivata seconda in x di u, con u_{xy} abbiamo indicato la derivata seconda mista (in x e y) di u e con u_{yy} abbiamo indicato la derivata seconda in y di u. Le funzioni a^{ij} che determinano la matrice caratteristica dell'equazione, si possono ora scrivere come:

$$a^{xx} = a\,, \quad a^{xy} = a^{yx} = b\,, \quad a^{yy} = c\,.$$

Osservazioni 1.3. Le equazioni di Maxwell prima descritte, non sono equazioni differenziali del secondo ordine, ma possono essere riscritte in modo che lo diventino, introducendo delle grandezze ausiliarie dette *potenziali elettromagnetici*, dei quali non ci occuperemo in questa sede se non in una versione ridotta, discutendo le equazioni del secondo ordine di tipo ellittico. Tratteremo invece altre equazioni che discendono dalle equazioni di Maxewll, come quella di D'Alembert precedentemente introdotta.

Passiamo ora alla classificazione delle PDE del secondo ordine quasi lineari [7]. La classificazione è dovuta alle proprietà della forma quadratica indotta dalla matrice di coefficienti $a^{ij}(x)$ e dipende dal punto considerato. Per arrivare ad enunciare tale classificazione dobbiamo studiare come l'equazione (1.11) cambia al variare delle coordinate utilizzate. Consideriamo una trasformazione di coordinate $y = y(x)$ dove $x \in \Omega$. Assumiamo che la trasformazione sia (almeno) di classe $C^2(\Omega)$, che sia invertibile e che la sua inversa sia una funzione di classe $C^2(\Omega')$, dove $\Omega' \subset \mathbb{R}^n$ è, per ipotesi, un insieme aperto connesso su cui variano le coordinate y. Di conseguenza avremo che la matrice Jacobiana della trasformazione è non nulla in ogni punto $x \in \Omega$. Infatti, nelle ipotesi fatte possiamo scrivere:

$$y^i = y^i(x(y))\,,$$

ed, applicando la regola di derivazione di funzioni di funzioni abbiamo che

$$\delta^i_j = \frac{\partial y^i}{\partial y^j} = \sum_{k=1}^{n} \frac{\partial y^i}{\partial x^k}\frac{\partial x^k}{\partial y^j}\,.$$

Equivalentemente, se J è la matrice jacobiana della trasformazione $y = y(x)$ e J' quella della trasformazione inversa $x = x(y)$, l'identità di sopra si scrive

$$I = JJ'\,.$$

In particolare quindi: $1 = \det J \det J'$ e pertanto $\det J \neq 0$. Faremo uso tra poco di questo risultato.

Osservazioni 1.4. Prima di procedere oltre, è interessante notare come invece di richiedere dall'inizio che l'inversa di $y = y(x)$ esista e sia di classe C^2 su qualche aperto Ω', avremmo potuto chiedere, con lo stesso risultato finale, che la funzione $y = y(x)$ fosse in $C^2(\Omega)$, invertibile e con matrice jacobiana J ovunque non singolare. Infatti, sotto tali ipotesi, per il teorema del Dini, (1) la funzione $y = y(x)$ è funzione aperta e pertanto l'immagine Ω' di Ω secondo $y = y(x)$ è aperto (e connesso visto che la funzione considerata è continua e Ω è connesso); (2) la funzione inversa $x = x(y)$ è in $C^2(\Omega')$.

La funzione u potrà essere espressa in funzione delle nuove coordinate $y^1, \dots y^n$ sull'insieme Ω':

$$u'(y) := u(x(y)) \, .$$

L'equazione differenziale (1.11) può essere trascritta per la funzione u' preservando la sua forma, ma cambiando le funzioni che in essa appaiono. Vediamo come procedere. Intanto osserviamo che, nelle ipotesi fatte:

$$\frac{\partial u}{\partial x^i} = \sum_{k=1}^{n} \frac{\partial y^k}{\partial x^i} \frac{\partial u'}{\partial y^k} \, ,$$

e quindi

$$\frac{\partial^2 u}{\partial x^h \partial x^i} = \sum_{k,r=1}^{n} \frac{\partial y^r}{\partial x^h} \frac{\partial^2 u'}{\partial y^r \partial y^k} \frac{\partial y^k}{\partial x^i} + \sum_{k=1}^{n} \frac{\partial^2 y^k}{\partial x^h \partial x^i} \frac{\partial u'}{\partial y^k} \, .$$

Inserendo queste identità nella (1.11), otteniamo che la stessa equazione differenziale può essere riscritta per la funzione u', come:

$$\sum_{p,q=1}^{n} a'^{pq}(y) \frac{\partial^2 u'}{\partial y^p \partial y^q} + \Phi'(y, u'(y), \nabla_y u') = 0 \, , \tag{1.13}$$

dove, per $k, r = 1, \dots, n$:

$$a'^{kr}(y(x)) := \sum_{i,j=1}^{n} \frac{\partial y^k}{\partial x^i} a^{ij}(x) \frac{\partial y^r}{\partial x^j} \, , \tag{1.14}$$

mentre:

$$\Phi'(y, u'(y), \nabla_y u') := \Phi\left(y(x), u(x(y)), J^{-1}(y)(\nabla_x u(x))_{x=x(y)}\right) + \sum_{i,j,p=1}^{n} a^{ij}(x(y)) \left.\frac{\partial^2 y^p}{\partial x^i \partial x^j}\right|_{x=x(y)} \frac{\partial u'}{\partial y^p}.$$

Per costruzione, l'equazione (1.11) è completamente equivalente all'equazione (1.13): $u = u(x)$ soddisfa la prima su Ω se e solo se $u'(y) := u(x(y))$ soddisfa la seconda su Ω'.
Fissando un punto $x_0 \in \Omega$, consideriamo una trasformazione lineare di coordinate

$$y^i = \sum_{k=1}^{n} J^i_k x^k \, ,$$

dove, se J è la matrice di coefficienti dati dalle *costanti* J^i_k, vale $\det J \neq 0$. In tal caso J è proprio la matrice Jacobiana della trasformazione considerata e sono soddisfatte le ipotesi sopra richieste dalle trasformazioni $y = y(x)$. La (1.14) valutata nel punto x_0 si può ora trascrivere come:

$$A'(y_0) = JA(x_0)J^t \,, \tag{1.15}$$

dove $A'(y_0)$ è la matrice di coefficienti $a'^{pq}(y_0)$ mentre $A(x_0)$ è la matrice di coefficienti $a^{ij}(x_0)$ ed infine $y_0 = y(x_0)$. Per il teorema di Sylvester, possiamo sempre scegliere la matrice non singolare J in modo tale che $A'(y_0)$ abbia *forma canonica di Sylvester*, cioè sia una matrice diagonale del tipo:

$$A'(y_0) = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0) \,,$$

in cui i numeri 1 compaiono r volte, i numeri -1 compaiono s volte ed i numeri 0 compaiono t volte. È noto, dalla teoria delle forme quadratiche, che la terna (r, s, t), detta *segnatura*, è una proprietà della forma quadratica associata alla matrice $A(x_0)$, e quindi una proprietà dell'equazione differenziale in riferimento al punto x_0. In altre parole, se esiste un'altra scelta della matrice non singolare J che riduce la matrice $A(x_0)$ tramite la (1.15) a forma canonica di Sylvester, il numero di volte in cui appariranno i numeri $1, -1, 0$ sulla diagonale principale saranno sempre, rispettivamente, i numeri r, s, t trovati sopra. Ovviamente $r + s + t = n$.
In modo del tutto analogo alla classificazione delle coniche tramite lo studio della forma quadratica associata si ha la seguente classificazione.

Definizione 1.3. Se $\Omega \in \mathbb{R}^n$ è un insieme aperto non vuoto e connesso, si consideri l'equazione quasi lineare del secondo ordine nella funzione a valori reali u:

$$\sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \Phi(x, u(x), \nabla_x u) = 0 \,, \tag{1.16}$$

dove $u \in C^2(\Omega)$ e $a^{ij} \in C^0(\Omega)$ e $\Phi \in C^0(\Omega \times \mathbb{R} \times \mathbb{R}^n)$ sono funzioni a valori reali assegnate. In riferimento alla matrice caratteristica $A(x_0)$ di coefficienti $a^{ij}(x_0)$, valutata nel punto $x_0 \in \Omega$, si dice che:

(a) l'equazione differenziale è di **tipo ellittico** in x_0 se la segnatura di $A(x_0)$ è $(n, 0, 0)$ oppure $(0, n, 0)$;

(b) l'equazione differenziale è di **tipo iperbolico** in x_0 se la segnatura di $A(x_0)$ è $(r, s, 0)$ con $r \neq 0$ e $s \neq 0$, in particolare si dice che l'equazione differenziale è di **tipo iperbolico normale** in x_0 se la segnatura di $A(x_0)$ è $(1, n-1, 0)$ con $n > 1$ oppure $(n-1, 1, 0)$ con $n > 1$;

(c) l'equazione differenziale è di **tipo parabolico** in x_0 se la segnatura di $A(x_0)$ è (r, s, t) con $t \neq 0$, in particolare si dice che l'equazione differenziale è di **tipo parabolico normale** in x_0 se la segnatura di $A(x_0)$ è $(n-1, 0, 1)$ con $n > 1$ oppure $(0, n-1, 1)$ con $n > 1$. $\diamondsuit$

Osservazioni 1.5. Ai fini della classificazione di una PDE quasilineare del secondo ordine, la matrice caratteristica può essere ridefinita moltiplicandola per una costante (più in generale

una funzione del punto in cui si valuta la matrice) diversa da zero (in ogni punto). Tale trasformazione non altera la classe di appartenenza dell'equazione differenziale come segue immediatamente dalle definizioni date sopra.

Esempi 1.1.
(1) L'**equazione di Tricomi** in $\mathbb{R}^2$ si scrive:

$$yu_{xx} + u_{yy} = 0 \, .$$

La forma quadratica associata è data dalla matrice non costante:

$$A(x, y) = \begin{bmatrix} y & 0 \\ 0 & 1 \end{bmatrix} \, .$$

Notiamo subito che, sull'asse delle ascisse, cioè $y = 0$, l'equazione differenziale è di *tipo parabolico normale.*
Consideriamo ora un punto (x_0, y_0) con $y_0 > 0$. Definiamo il nuovo sistema di coordinate cartesiane (x', y') su $\mathbb{R}^2$ dove $y' := y$ mentre $x' = x/\sqrt{y_0}$. In queste coordinate, l'equazione prende forma:

$$\frac{y'}{y_0} u_{x'x'} + u_{y'y'} = 0 \, .$$

Pertanto, esattamente in $(x, y) = (x_0, y_0)$, la matrice associata all'equazione, nelle nuove coordinate diventa:

$$A'(x', y') = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \, .$$

Concludiamo che, nel semipiano $y > 0$, l'equazione di Tricomi è di *tipo ellittico.*
Consideriamo infine un punto (x_0, y_0) con $y_0 < 0$. Definiamo il nuovo sistema di coordinate cartesiane (x', y') su $\mathbb{R}^2$ dove $y' := y$ mentre $x' = x/\sqrt{-y_0}$. In queste coordinate, l'equazione prende forma:

$$-\frac{y'}{y_0} u_{x'x'} + u_{y'y'} = 0 \, .$$

Pertanto, esattamente in $(x, y) = (x_0, y_0)$, la matrice associata all'equazione, nelle nuove coordinate diventa:

$$A'(x', y') = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \, .$$

Concludiamo che, nel semipiano $y < 0$, l'equazione di Tricomi è di *tipo iperbolico normale.*
(2) L'equazione di Poisson su $\Omega \subset \mathbb{R}^n$ si scrive:

$$\Delta u = \rho$$

dove $\rho = \rho(x)$ è una funzione almeno C^0 assegnata e l'*operatore di Laplace*, detto anche **laplaciano**, Δ è definito in coordinate cartesiane ortonormali come:

$$\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x^{i2}} \, .$$

Nel caso la funzione ρ sia identicamente nulla su Ω, l'equazione suddetta si riduce all'**equazione di Laplace** su Ω:

$$\Delta u = 0 \,,$$

le cui soluzioni sono dette **funzioni armoniche** su Ω.
In entrambi i casi, la matrice $A(x_0)$, per ogni $x_0 \in \Omega$ è individuata dalla matrice identità . In base alla classificazione suddetta, le equazioni di Poisson e di Laplace sono equazioni di *tipo ellittico* in ogni punto. Dal punto di vista fisico, se $n = 3$, $-u$ può essere pensato come il *potenziale elettrostatico* e, nel caso dell'equazione di Poisson, ρ corrsisponde alla densità di carica elettrica presente nel volume Ω.
(3) Se $\Omega := \mathbb{R}^n$, l'**equazione delle onde** o **equazione di D'Alembert** si scrive:

$$-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} + \Delta u = 0 \,.$$

Sopra le coordinate su $\mathbb{R}^n$ sono state decomposte come: $x = (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^{n-1}$. La costante $c > 0$ si interpreta fisicamente come la velocità di propagazione della perturbazione ondosa descritta dalla funzione $u = u(t, \mathbf{x})$, dove t è il tempo e $\mathbf{x}$ lo spazio (almeno nel caso di $n = 3$) di un riferimento. L'equazione di D'Alembert descrive tutti i fenomeni di propagazione ondosa (armonica) conosciuti: dalla propagazione della luce a quella del suono, ma anche la propagazione di una deformazione in un mezzo continuo elastico, fino ad arrivare alla propagazione delle onde gravitazionali nella teoria della relatività generale. La matrice $A(x_0)$, per ogni $x_0 \in \Omega$, ha la forma $\mathrm{diag}(-1/c^2, 1, \ldots, 1)$. Passando a coordinate $y^1 := ct$, $\mathbf{y} := \mathbf{x} \in \mathbb{R}^{n-1}$, si ottiene che la matrice $A'(y_0)$ ha forma canonica di Sylvester in ogni punto y_0, con segnatura $(n-1, 1, 0)$. Pertanto l'equazione di D'Alembert è , ovunque, di *tipo iperbolico normale.*
Un caso particolare dell'equazione di D'alembert è dato, per $n = 4$, quando si lavora con il *campo elettrico* $\mathbf{E} = \mathbf{E}(t, \mathbf{x})$ e con il *campo magnetico* $\mathbf{B} = \mathbf{B}(t, \mathbf{x})$ che descrivono i fenomeni elettromagnetici. L'equazione di D'alembert per questi due campi (nel vuoto) descrive la propagazione delle onde elettromagnetiche (e quindi della luce in particolare)

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} - c^2 \Delta \mathbf{E} = 0 \,, \quad \frac{\partial^2 \mathbf{B}}{\partial t^2} - c^2 \Delta \mathbf{B} = 0 \,,$$

c è la velocità della luce nel vuoto. Le due equazioni di sopra, che seguono dalle più generali *equazioni di Maxwell*, si devono intepretare componente per componente e significano che *ogni componente* del campo $\mathbf{E}$ e del campo $\mathbf{B}$ soddisfa separatamente l'equazione di D'Alembert.
(4) Se $\Omega := \mathbb{R}^4$, l'**equazione di Klein-Gordon** si scrive:

$$-\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} + \Delta \phi - \frac{m^2 c^4}{\hbar^2}\phi = 0 \,.$$

Come prima, le coordinate su $\mathbb{R}^4$ sono state decomposte come: $x = (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3$. La costante $c > 0$ si interpreta fisicamente come la velocità della luce nel vuoto, $\hbar = h/2\pi$ dove h è la *costante di Planck* t è il tempo e $\mathbf{x}$ lo spazio di un sistema di riferimento (Minkowskiano della teoria della Relatività Speciale). L'equazione di Klein-Gordon descrive una campo scalare ϕ a valori in $\mathbb{R}$

associato a particelle quantistiche di massa $m > 0$ senza spin ed eletricamente neutre (in realtà ϕ descrive sia la particella che l'anti particella). La matrice $A(x_0)$, per ogni $x_0 \in \Omega$, ha la forma $\mathrm{diag}(-1/c^2, 1, 1, 1)$. Passando a coordinate $y^1 := ct$, $\mathbf{y} := \mathbf{x} \in \mathbb{R}^3$, si ottiene che la matrice $A'(y_0)$ ha forma canonica di Sylvester in ogni punto y_0, con segnatura $(3, 1, 0)$. Pertanto l'equazione di D'Alembert è , ovunque, di *tipo iperbolico normale.*
(5) Se $\Omega := \mathbb{R}^n$, l'**equazione del calore** si scrive:

$$-\frac{\partial u}{\partial t} + a^2 \Delta u = q \, .$$

dove $a > 0$ è una costante e $q = q(t, \mathbf{x})$ una funzione assegnata. Sopra le coordinate su $\mathbb{R}^n$ sono ancora state decomposte come: $x = (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^{n-1}$. Almeno per $n = 3$, u si intepreta fisicamente come la temperatura in un mezzo continuo le cui caratteristiche termodinamiche sono riassunte dai parametri a e dalla funzione q che corrisponde ad una sorgente di calore.
La matrice $A(x_0)$, per ogni $x_0 \in \Omega$, ha la forma $\mathrm{diag}(0, a^2, \ldots, a^2)$. Passando a definire coordinate $y^1 := t$, $\mathbf{y} := a\mathbf{x} \in \mathbb{R}^{n-1}$, si ottiene in tali coordinate che la matrice $A'(y_0)$ ha forma canonica di Sylvester in ogni punto y_0, con segnatura $(0, n-1, 1)$: si osservi infatti che nell'equazione del calore *non* compare la derivata seconda nella prima variabile, questo spiega l'ultimo 1 e l'$n-1$ nella segnatura. Pertanto l'equazione del calore è , ovunque, di *tipo parabolico normale.*
(6) L'**equazione di Schrödinger** per la *funzione d'onda* $\psi = \psi(t, \mathbf{x})$, con $(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3$:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi + V(\mathbf{x}) \psi \, ,$$

non ricade nella classificazione suddetta in quanto: (1) $\psi = \psi(t, \mathbf{x})$ è una funzione a valori complessi e (2) il coefficiente della derivata temporale è immaginario puro. Tuttavia tale equazione ha caratteristiche simili all'equazione del calore.
(7) L'**equazione di Dirac** per il *il campo fermionico* $\Psi = \Psi(x) \in \mathbb{C}^4$, con $x \equiv (x^0 = ct, x^1, x^2, x^3)$:

$$i\hbar \sum_{\mu=0}^{3} \gamma^\mu \frac{\partial \Psi}{\partial x^\mu} - mc\Psi = 0$$

che descrive il campo quantistico associato agli elettroni ed i positroni, non ricade nella classificazione suddetta in quanto: (1) $\Psi = \Psi(t, \mathbf{x})$ è una funzione a valori in $\mathbb{C}^4$ e (2) i coefficienti delle derivate sono matrici complesse. Tuttavia tale equazione, per taluni aspetti, ha caratteristiche simili all'equazione di Klein-Gordon. Le 4 matrici complesse γ^μ, per $\mu = 0, 1, 2, 3$ sono dette *matrici di Dirac* e soddisfano le relazioni di Dirac (o Clifford)

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = -2g^{\mu\nu} \, ,$$

dove la matrice dei coefficienti $g_{\mu\nu}$ è quella della metrica di Lorentz $diag(-1, 1, 1, 1)$.

1.5 Il problema di Cauchy ed il Teorema di Cauchy-Kovalevskaja.

1.5.1 Superfici regolari in $\mathbb{R}^n$.

Una superficie regolare Σ di dimensione $n-1$ in $\mathbb{R}^n$ (equivalentemente detta sottovarietà embedded di $\mathbb{R}^n$ di dimensione $n-1$ e classe di differenziabilità C^k con $k \geq 1$) è un sottoinsieme di $\mathbb{R}^n$ tale che nell'intorno A_p di di ogni suo punto $p \in \Sigma$ può essere espressa come luogo dei punti che annullano una funzione S di classe C^k, con $k \geq 1$, con $dS \neq 0$. In altre parole, per ogni $p \in \Sigma$, esiste un suo intorno aperto $A_p \in \Omega$ ed una funzione $S : A_p \to \mathbb{R}$, $S \in C^1(A_p)$, tale che:

$$\Sigma \cap A_p = \{x \in A_p \mid S(x) = 0\} ,$$

unitamente a:

$$dS(x) \neq 0 , \quad \text{per } S(x) = 0.$$

Con le definizioni poste, dS si identifica con ∇S (ne ha le stesse componenti lavorando in basi ortonormali) ∇S (cioè dS) è un vettore normale a Σ, nel senso che, se $\mathbf{e}_i$ è il versore i-esimo della base canonica di $\mathbb{R}^n$, il vettore mai nullo per $x \in \Sigma \cap A_p$:

$$\nabla S(x) := \sum_{i=1}^{n} \frac{\partial S}{\partial x^i}|_x \, \mathbf{e}_i$$

è normale a $\Sigma \cap A_p$ in ogni punto.
Esistono diverse definizioni equivalenti di superfici regolari $n-1$ dimensionali in $\mathbb{R}^n$. Si osservi che ogni funzione di classe C^k $(k \geq 1)$, $x^1 = f(x^2, \ldots, x^n)$ individua una superficie regolare. In tal caso la funzione S può essere scelta come:

$$S(x^1, x^2, \ldots, x^n) := x^1 - f(x^2, \ldots, x^n) .$$

1.5.2 Il problema di Cauchy e la "ben posizione" del problema nel senso di Hadamard.

Se $\Omega \subset \mathbb{R}^n$ è un aperto connesso non vuoto e $\Sigma \subset \Omega$ è una superficie regolare (con qualche ordine di differenziabilità $k \geq 1$) di dimensione $n-1$ che divide Ω in due parti connesse[3], il **problema di Cauchy** del secondo ordine riferito a Σ consiste nel sistema:

$$\begin{cases} F\left(x, u(x), \frac{\partial u}{\partial x^1}, \cdots, \frac{\partial u}{\partial x^n}, \frac{\partial^2 u}{\partial x^1 \partial x^1}, \frac{\partial^2 u}{\partial x^1 \partial x^2} \cdots, \frac{\partial^2 u}{\partial x^n \partial x^n}\right) = 0 , \\ u \restriction_\Sigma = u_0 , \\ \dfrac{\partial u}{\partial \mathbf{n}} \restriction_\Sigma = u_1 , \end{cases} \tag{1.17}$$

in cui u_0 e u_1 sono funzioni assegnate su Σ di qualche ordine di differenziabilità da definirsi, $u \in C^2(\Omega)$ è la funzione incognita da determinare, F è una funzione nota che determina l'equazione

[3] Questa richiesta è in realtà inessenziale a questo punto della teoria.

differenziale ed abbiamo usato la notazione standard:

$$\frac{\partial u}{\partial \mathbf{n}} := \sum_{i=1}^{n} n^i \frac{\partial u}{\partial x^i}$$

dove $\mathbf{n} = (n^1, \ldots, n^n)$ è il *versore* normale a Σ.

Secondo l'impostazione data da Hadamard, un problema di Cauchy (1.17) è **ben posto** se sono verificate le seguenti tre richieste.
(a) Una soluzione del problema, in una fissata classe di funzioni esiste.
(b) La soluzione in (a) è l'unica soluzione nella classe di funzioni suddetta.
(c) La soluzione dipende *con continuità* dai dati di Cauchy in qualche topologia di spazi di funzioni definita negli spazi di funzioni considerate.

La condizione (c) deriva dal fatto che, nella pratica, i dati di Cauchy sono sempre noti con una certa approssimazione e si richiede che, pertanto, le soluzioni varino di poco se le condizioni di Cauchy variano di poco.
Nel caso una delle tre condizioni di sopra sia violata, si dice che il problema di Cauchy è *mal posto nel senso di Hadamard.* Mostriamo che il problema di Cauchy ellittico può essere mal posto perché viola la condizione (c), mettendo nello spazio dei dati di Cauchy limitati la topologia naturale dell'estremo superiore indotta dalla norma $||\cdot||_\infty$.
Cosideriamo in $\mathbb{R}^2$, con coordinate (x, y), il problema di Laplace:

$$\begin{cases} \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = 0 \\ u(x,0) = 0\,, \\ \dfrac{\partial u}{\partial y}(x,0) = \dfrac{1}{k}\sin(kx)\,, \end{cases} ,$$

dove $k > 0$ è fissata e si cercano soluzioni $u \in C^2(\mathbb{R}^2)$. Si dimostra che l'unica soluzione del problema posto è :

$$u^{(k)}(x,0) := \frac{1}{k^2}\sin(kx)\ \sinh(ky)\,.$$

Si noti che, nel limite $k \to +\infty$ il dato di Cauchy:

$$u_1^{(k)}(x) = \frac{1}{k}\sin(kx)$$

soddisfa $||u_1^{(k)}||_\infty \to 0$ per $k \to +\infty$, mentre l'altro dato è nullo per ipotesi e pertanto $||u_0^{(k)}||_\infty = 0$. Tuttavia la soluzione $u^{(k)}(x, y)$ non tende, puntualmente, ad alcun limite se $x \neq 0$ e $y \neq 0$.
Si può mostrare che, le patologie del problema di Cauchy per equazioni ellittiche riguardano anche l'esistenza e l'unicità della soluzione, in particolare quando il problema di Cauchy è imposto in regioni limitate. Vedremo ciò nel prossimo capitolo. In realtà per problemi di Cauchy di carattere ellittico, si ha anche la perdita di validità della richiesta (a) almeno quando la superficie Σ

è una superficie chiusa che contorna Ω. In tal caso, lavorando per esempio con l'equazione di Laplace ed assegnando dati di Cauchy analitici su Σ, per ogni punto di Σ c'è un intorno in cui esiste una soluzione[4] del problema. La parte dell'intorno che interseca Ω produce localmente una soluzione del problema di Cauchy. Quello che però accade è che le diverse soluzioni ottenute localmente in questo modo non si "saldano" tra di loro per determinare un'unica soluzione su Ω: su un punto interno a Ω coperto da due intorni suddetti si trovano generalmente valori distinti per u a seconda dell'intorno scelto.

1.5.3 Il Teorema di Cauchy-Kovalevskaja.

Ci occuperemo ora di studiare il problema della risolubilità del problema di Cauchy sopra scritto nel caso in cui l'equazione differenziale che compare in esso sia del tipo (1.11), introducendo il *teorema di Cauchy-Kovalevskaja.* A tal fine consideriamo il caso più semplice nel quale $\Omega = \mathbb{R}^n$ e Σ è il piano $x^1 = 0$.
Un equazione differenziale alle derivate parziali del secondo ordine, nella funzione incognita $u = u(x)$, si dice che è scritta **in forma normale rispetto alla variabile** t, se $x = (t, \mathbf{x}) \in \Omega \subset \mathbb{R}^n$ (con $\mathbf{x} = (x^2, \dots, x^n)$) se è rappresentata nella forma:

$$\frac{\partial^2 u}{\partial t^2} = F\left(t, \mathbf{x}, u(t,\mathbf{x}), \frac{\partial u}{\partial t}, \frac{\partial u}{\partial x^2}, \dots, \frac{\partial u}{\partial x^n}, \frac{\partial^2 u}{\partial t \partial x^2}, \cdots, \frac{\partial^2 u}{\partial x^n \partial x^n}\right) . \tag{1.18}$$

Si noti che il secondo membro è funzione delle derivate al più del secondo ordine in cui la derivata in t appare al più al primo ordine. Si può dare un'analoga definizione di equazione in forma normale rispetto ad una coordinata, anche per equazioni di ordine superiore al secondo.
Consideriamo il **Problema di Cauchy** per la funzione $u \in C^2(\mathbb{R}^n)$ con dati di Cauchy sulla superficie Σ individuata da $t = 0$ e con equazione in forma normale nella variabile t:

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = & F\left(t, \mathbf{x}, u(t,\mathbf{x}), \dfrac{\partial u}{\partial t}, \dfrac{\partial u}{\partial x^2}, \dots, \dfrac{\partial u}{\partial x^n}, \dfrac{\partial^2 u}{\partial t \partial x^2}, \cdots, \dfrac{\partial^2 u}{\partial x^n \partial x^n}\right) , \\ u(0,\mathbf{x}) = & u_0(\mathbf{x}) , \\ \dfrac{\partial u}{\partial t}(0,\mathbf{x}) = & u_1(\mathbf{x}) , \end{cases}$$

in cui $u_0 \in C^1(\mathbb{R}^{n-1})$ e $u_1 \in C^0(\mathbb{R}^{n-1})$ sono funzioni assegnate su Σ.
Esistono vari teoremi che assicurano l'esistenza e l'unicità della soluzione. Il primo di tutti questi teoremi è quello dovuto a Cauchy e Kovalevskaja. Tale teorema non ha grande utilità nelle applicazioni moderne, in quanto richiede ipotesi estremamente forti per funzionare, ma è comunque di grande utilità teorica, in quanto serve come lemma intermedio per provare teoremi più moderni basati su ipotesi molto deboli. Enunciamo il teorema per equazioni del secondo ordine, anche se l'ordine dell'equazione non è essenziale.
Ricordiamo che una funzione a valori reali $f = f(x)$ è detta **analitica (reale) nell'insieme**

[4] Questa è l'unica soluzione in tale intorno per il teorema di Cauchy-Kovalevskaja, visto che le funzioni che soddisfano l'equazione di Laplace sono analitiche come proveremo più avanti.

aperto $\Omega \subset \mathbb{R}^n$, se per ogni $x_0 \in \Omega$, in un intorno aperto di x_0 incluso in Ω, la funzione f si può scrivere come serie di Taylor centrata in x_0:

$$f(x) = \sum_{N=0}^{+\infty} \sum_{\alpha_1+\cdots+\alpha_n=N} \frac{1}{\alpha_1! \cdots \alpha_n!} \left. \frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\partial x^{1\alpha_1} \cdots \partial x^{n\alpha_n}} \right|_{x_0} (x^1 - x_0^1)^{\alpha_1} \cdots (x^n - x_0^n)^{\alpha_n} .$$

Teorema 1.3. **(Cauchy-Kovalevskaja).** *Si consideri il problema di Cauchy del secondo ordine nella funzione incognita $u = u(t, \mathbf{x})$ con $(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^{n-1}$, con dati di Cauchy assegnati sulla superficie determinata da $t = 0$ e con equazione differenziale scritta in forma normale nella variabile t:*

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = & F\left(t, \mathbf{x}, u(t, \mathbf{x}), \dfrac{\partial u}{\partial t}, \dfrac{\partial u}{\partial x^2}, \ldots, \dfrac{\partial u}{\partial x^n}, \dfrac{\partial^2 u}{\partial t \partial x^2}, \cdots, \dfrac{\partial^2 u}{\partial x^n \partial x^n}\right) , \\ u(0, \mathbf{x}) = & u_0(\mathbf{x}) , \\ \dfrac{\partial u}{\partial t}(0, \mathbf{x}) = & u_1(\mathbf{x}) . \end{cases} \tag{1.19}$$

Se, per un punto $(0, \mathbf{x})$, le funzioni u_0 e u_1 sono analitiche nell'intorno di $\mathbf{x}_0$ e la funzione F è analitica nell'intorno del punto:

$$\left(0, \mathbf{x}_0, u_0(\mathbf{x}_0), u_1(\mathbf{x}_0), \frac{\partial u_0}{\partial x^2}(x_0), \ldots, \frac{\partial u_0}{\partial x^n}(x_0), \frac{\partial u_1}{\partial x^2}(\mathbf{x}_0), \cdots, \frac{\partial u_1}{\partial x^n}(\mathbf{x}_0), \frac{\partial^2 u_1}{\partial x^2 \partial x^2}(\mathbf{x}_0) \ldots, \frac{\partial^2 u_1}{\partial x^n \partial x^n}(\mathbf{x}_0)\right)$$

allora in un intorno di $(0, \mathbf{x}_0)$ esiste una soluzione del sistema (1.19). Tale soluzione è l'unica soluzione analitica nell'intorno considerato. $\diamondsuit$

Idea della dimostrazione. Se esiste una soluzione u deve valere:

$$u(0, \mathbf{x}_0) = u_0(\mathbf{x}_0) ,$$

inoltre, se $i = 2, \ldots, n$:

$$\frac{\partial u}{\partial x^i}(0, \mathbf{x}_0) = \frac{\partial u_0}{\partial x^i}(\mathbf{x}_0) , \quad \frac{\partial u}{\partial t}(0, \mathbf{x}_0) = u_1(\mathbf{x}_0) .$$

Consideriamo poi l'equazione differenziale con entrambi i membri valutati in $(t, \mathbf{x}_0)$. In tale situazione, il secondo membro dell'equazione differenziale è funzione dei dati di Cauchy unicamente:

$$\frac{\partial^2 u}{\partial t^2}|_{(0,\mathbf{x}_0)} =$$

$$F\left(0, \mathbf{x}_0, u_0(\mathbf{x}_0), u_1(\mathbf{x}_0), \frac{\partial u_0}{\partial x^2}(x_0), \ldots, \frac{\partial u_0}{\partial x^n}(x_0), \frac{\partial u_1}{\partial x^2}(\mathbf{x}_0), \cdots, \frac{\partial u_1}{\partial x^n}(\mathbf{x}_0), \frac{\partial^2 u_1}{\partial x^2 \partial x^2}(\mathbf{x}_0), \ldots, \frac{\partial^2 u_1}{\partial x^n \partial x^n}(\mathbf{x}_0)\right) .$$

In questo modo abbiamo ottenuto la derivata seconda nel tempo della soluzione, ammesso che esista, nel punto $(0, \mathbf{x}_0)$, in funzione dei dati di Cauchy. Le rimanenti derivate seconde della soluzione sono invece note direttamente dai dati di Cauchy $(i, k = 2, \ldots, n)$:

$$\frac{\partial^2 u}{\partial t \partial x^i}|_{(0,\mathbf{x}_0)} = \frac{\partial u_1}{\partial x^i}|_{\mathbf{x}_0} \ , \quad \frac{\partial^2 u}{\partial x^k \partial x^i}|_{(0,\mathbf{x}_0)} = \frac{\partial^2 u_0}{\partial x^k \partial x^i}|_{\mathbf{x}_0}$$

In definitiva abbiamo ottenuto tutte le derivate fino all'ordine 2 incluso, della eventuale soluzione, valutate in $(0, \mathbf{x}_0)$, in funzione dei dati di Cauchy. Nell'ipotesi di F analitica, e quindi infinitamente differenziabile, possiamo ottenre *tutte* le derivate di ogni ordine della eventuale soluzione, valutate in $(0, \mathbf{x}_0)$, in funzione dei dati di Cauchy. Per fare ciò è sufficiente derivare entrambi i membri dell'equazione differenziale e iterare la procedura seguita sopra per l'identità che si è ottenuta. Si osservi che in tutta la procedura, il fatto che l'equazione sia scritta in forma normale è di centrale importanza. In questo modo si può scrivere una serie di Taylor formale della eventuale soluzione, sviluppata nel punto $(0, \mathbf{x}_0)$. Si vede che, nelle ipotesi fatte (in particolare, tenendo conto dell'analiticità di F) tale serie converge effettivamente e, per costruzione, converge ad una soluzione del problema di Cauchy considerato. L'unicità è conseguenza di teoremi di unicità validi per funzioni analitiche. $\square$

Torniamo ora al problema di Cauchy (1.17) specializzato al caso di un'equazione del secondo ordine in forma quasi lineare (1.11) e riferito ad una superficie regolare $\Sigma \subset \Omega$ di dimensione $n-1$ che divide $\Omega \subset \mathbb{R}^n$ in due parti connesse:

$$\begin{cases} \sum_{i,j=1}^n a^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \Phi(x, u(x), \nabla_x u) = 0 \ , \\ u \restriction_\Sigma = u_0 \ , \\ \dfrac{\partial u}{\partial \mathbf{n}} \restriction_\Sigma = u_1 \ , \end{cases} \tag{1.20}$$

$\mathbf{n}$ come già detto, è il versore normale a Σ.
Vogliamo ridurci alla situazione in cui Σ appare come il piano $t = 0$, per poter cercare di applicare il teorema di Cauchy-Kovalevskaja, almeno nell'intorno di $p \in \Sigma$. Per far ciò usiamo un opportuno sistema di coordinate dette *coordinate normali Riemanniane*. Sia $\xi^2, \ldots, \xi^n$ un qualsiasi sistema di coordinate su Σ (di classe C^k come S) definito nell'intorno di $p \in \Sigma$. Per ogni punto $q \in \Sigma$ nell'intorno considerato, tracciamo la retta normale a Σ e passante per q. Sia $t \in \mathbb{R}$ la lunghezza d'arco su tale retta ponendo come origine di essa il punto q in cui la retta interseca Σ. Si dimostra che viene a definirsi in questo modo un sistema di coordinate $(t, \xi^1, \ldots, \xi^n)$, di classe C^k, in un intorno aperto $B_p \subset \Omega$ di p. Un punto $r \in B_p$ è individuato in questo modo: si considera l'unica retta γ_r perpendicolare a Σ che passa per r, la lunghezza del segmento tra r e l'intersezione $q_r = \gamma_r \cap \Sigma$ è la coordinata t_r di r, le rimanenti coordinate $\xi_r^2, \ldots, \xi_r^n$ non sono altro che le coordinate di q_r su Σ. Questo sistema di coordinate ha la particolarità che, data la natura di t, vale:

$$\Sigma \cap B_p = \{(t, \xi^2, \ldots, \xi^n) \ | \ t = 0\} \ .$$

Ulteriormente, se $(\xi^2, \dots \xi^n)$ sono fissate, la curva $t \mapsto x^i(t, \xi^2, \dots \xi^n)$ risulta essere normale a Σ per costruzione ed il suo vettore tangente

$$\mathbf{n} := \sum_{i=1}^{n} \frac{\partial x^i}{\partial t}|_{t=0} \, \mathbf{e}_i$$

non è altro che il *versore* normale a Σ nel punto di coordinate normali riemanniane $(0, \xi^2, \dots \xi^n)$. Di conseguenza

$$\frac{\partial u}{\partial \mathbf{n}} := \sum_{i=1}^{n} n^i \frac{\partial u}{\partial x^i} = \sum_{i=1}^{n} \frac{\partial x^i}{\partial t}|_{t=0} \frac{\partial u}{\partial x^i} = \frac{\partial u}{\partial t}|_{t=0} \, , \tag{1.21}$$

dove $u = u(t, \xi^2, \dots, \xi^n)$ è una funzione arbitraria definita nell'intorno di p espressa in coordinate normali riemanniane. Lavorando nell'intorno di p in coordinate normali riemanniane, il problema di Cauchy (1.20) diventa della forma:

$$\begin{cases} a'^{tt}(t, \boldsymbol{\xi}) \dfrac{\partial^2 u}{\partial t^2} = -\displaystyle\sum_{i=2}^{n} 2a'^{it}(t, \boldsymbol{\xi}) \frac{\partial^2 u}{\partial t \partial \xi^i} - \sum_{i,j=2}^{n} a'^{ij}(t, \boldsymbol{\xi}) \frac{\partial^2 u}{\partial \xi^i \partial \xi^j} \\ +\Phi' \left(t, \boldsymbol{\xi}, u(t, \boldsymbol{\xi}), \dfrac{\partial u}{\partial t}, \dfrac{\partial u}{\partial \xi^2}, \dots, \dfrac{\partial u}{\partial \xi^n}, \dfrac{\partial^2 u}{\partial t \partial \xi^2}, \dots, \dfrac{\partial^2 u}{\partial \xi^n \partial \xi^n} \right) , \\ u(0, \boldsymbol{\xi}) = u_0(\boldsymbol{\xi}) \, , \\ \dfrac{\partial u}{\partial t} \restriction_{\Sigma} = u_1(0, \boldsymbol{\xi}) \, , \end{cases} \tag{1.22}$$

Notiamo che se $a'^{tt}(t, \boldsymbol{\xi}) \neq 0$ in un intorno di p, allora possiamo dividere entrambi i membri dell'equazione differenziale per $a'^{tt}(t, \boldsymbol{\xi})$ ottenendo un'equazione differenziale in forma normale nella variabile t. Il problema di Cauchy risultante in questa situazione è un caso particolare del problema che appare nelle ipotesi del teorema di Cauchy-Kovalevskaja. Per poter applicare il teorema detto sono comunque ancora necessarie ipotesi di analiticità della quali non ci occuperemo. Tornando in coordinate cartesiane e facendo uso della (1.14), la condizione $a'^{tt}(0, \boldsymbol{\xi}) \neq 0$, tenendo conto che è :

$$a'^{tt}(t, \boldsymbol{\xi}) = \sum_{i,j=1} a^{ij}(x) \frac{\partial t}{\partial x^i} \frac{\partial t}{\partial x^j} \, ,$$

può essere riscritta, per $x \in \Sigma$:

$$\sum_{i,j=1} a^{ij}(x) \frac{\partial t}{\partial x^i} \frac{\partial t}{\partial x^j} \neq 0 \, , \tag{1.23}$$

Concludiamo che, se in un punto $x \in \Sigma$ vale la (1.23), allora possiamo applicare il teorema di Cauchy-Kovalevskaja nell'intorno di quel punto riducendosi a lavorare in coordinate normali riemanniane purché siano soddisfatte le necessariie ipotesi di analiticità .

1.5.4 Superfici caratteristiche.

Ci interessa ora approfondire la situazione in cui, in riferimento al problema di Cauchy problema di Cauchy (1.20), valga

$$\sum_{i,j=1} a^{ij}(x) \frac{\partial t}{\partial x^i} \frac{\partial t}{\partial x^j} = 0 \,. \tag{1.24}$$

ovunque sulla superficie Σ. Questo significa che, passando in coordinate normali riemanniane, $a'^{tt}(0, \boldsymbol{\xi}) = 0$. In tal caso il teorema di Cauchy-Kovalevskaja non può essere applicato e non ci sono garanzie, per tale via, sull'esistenza e l'unicità di una soluzione nell'intorno del punto singolare detto. Si osservi che la funzione $t = t(x)$ individua la superficie regolare Σ su cui diamo le condizioni di Cauchy, tramite la richiesta $t(x) = 0$, di cui $dt = \sum_i \frac{\partial t}{\partial x^i} \mathbf{e}_i$ è vettore normale mai nullo. Nel caso in cui valga (1.24), non possiamo applicare il teorema di Cauchy-Kovalevskaja per dati di Cauchy assegnati la superficie individuata da $t = 0$. Ora focalizzeremo l'attenziaone su questa classe di superfici. Tuttavia, per definirire tale classe di superfici, l'uso delle coordinate Riemanniane e della coordinata t è scomodo per vari motivi e pertanto vogliamo individuare tali superfici senza fare esplicito riferimento a tali coordinate. Diamo a tal fine la seguente definizione in cui rimpiazziamo la coordinata riemanniana t con una generica funzione S.

Definizione 1.4. Se $\Omega \in \mathbb{R}^n$ è un insieme aperto non vuoto e connesso, si consideri l'equazione quasi lineare del secondo ordine nella funzione a valori reali u:

$$\sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \Phi(x, u(x), \nabla_x u) = 0 \,, \tag{1.25}$$

dove $u \in C^2(\Omega)$ e $a^{ij} \in C^0(\Omega)$ e $\Phi \in C^0(\Omega \times \mathbb{R} \times \mathbb{R}^n)$ sono funzioni a valori reali assegnate. Una superficie regolare $\Sigma \subset \Omega$ di dimensione $n-1$ è detta **superficie caratteristica** per l'equazione differenziale (1.25), se nell'intorno di ogni $p \in \Sigma$ può essere espressa come il luogo dei punti x in cui $S(x) = 0$ dove S è una funzione almeno C^1 definita nell'intorno di p, soddisfacente $dS \neq 0$ su Σ e

$$\sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} = 0 \,, \quad \text{per } x \in \Sigma \,. \tag{1.26}$$

$\diamondsuit$

Mostriamo ora che, effettivamente, le superfici caratteristiche non permettono di applicare il teorema di Cauchy-Kovalevskaja quando i dati di Cauchy sono assegnati su di esse. Successivamente, nelle osservazioni, mostreremo che il problema di Cauchy con dati iniziali su una superficie caratteristica è , in generale, affetto da varie patologie.
Se Σ è una superficie caratteristica individuata dal luogo degli zeri della funzione S, possiamo usare su Σ le coordinate normali Riemanniane $t, \xi^1, \ldots, \xi^n$. In questo caso abbiamo che $S(0, \xi^1, \ldots, \xi^n) = 0$ per ogni scelta delle ξ^k, dato che il luogo dei punti a $S = 0$ coincide con il

luogo dei punti a $t = 0$ e coincide con Σ per costruzione. Di conseguenza

$$\left.\frac{\partial S}{\partial \xi^k}\right|_{t=0} = 0\,, \qquad \left.\frac{\partial S}{\partial t}\right|_{t=0} \neq 0\,.$$

La seconda condizione deriva dal fatto che, se non fosse vera, avremmo che tutte le derivate di S sono nulle su Σ e questo è impossibile per la richiesta $dS \neq 0$ su Σ. Di conseguenza abbiamo in particolare che, esattamente su Σ,

$$\sum_{i,j=1}^{n} a^{ij}\,\left.\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j}\right|_{\Sigma} = \sum_{i,j=1}^{n} a^{ij}\left(\frac{\partial S}{\partial t}\frac{\partial t}{\partial x^i} + \sum_{k=2}^{n}\frac{\partial S}{\partial \xi^k}\frac{\partial \xi^k}{\partial x^i}\right)\left(\frac{\partial S}{\partial t}\frac{\partial t}{\partial x^j} + \sum_{h=2}^{n}\frac{\partial S}{\partial \xi^h}\frac{\partial \xi^h}{\partial x^j}\right)$$

$$= \sum_{i,j=1}^{n} a^{ij}\left(\frac{\partial S}{\partial t}\frac{\partial t}{\partial x^i} + 0\right)\left(\frac{\partial S}{\partial t}\frac{\partial t}{\partial x^j} + 0\right) = \sum_{i,j=1}^{n} a^{ij}\,\left.\frac{\partial t}{\partial x^i}\frac{\partial t}{\partial x^j}\right|_{\Sigma}\left(\frac{\partial S}{\partial t}\right)^2 = \left(\frac{\partial S}{\partial t}\right)^2 a'^{tt}\,,$$

dove $\frac{\partial S}{\partial t} \neq 0$ nelle nostre ipotesi, e quindi:

$$a'^{tt}(0,\boldsymbol{\xi}) = \left(\frac{\partial S}{\partial t}\right)^{-2}\sum_{i,j=1}^{n} a^{ij}(x)\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j}$$

dove il coefficiente davanti alla somma a secondo membro è ben definito. Concludiamo che, come preannunciato, la richiesta che Σ sia una superficie caratteristica (ed in particolare la (1.26)) implica che, lavorando in coordinate riemanniane normali attorno a Σ, il coefficiente a'^{tt} in (1.22) si annulli su Σ (cioè per $t = 0$) e quindi non si possa applicare il teorema di Cauchy-Kovalevskaja. Si deve notare che noi abbiamo assunto che $a'^{tt}(0,\boldsymbol{\xi}) = 0$ su tutta Σ. In realtà , perché si annulli $a'^{tt}(0,\boldsymbol{\xi})$ in un punto $p = (0,\boldsymbol{\xi}) \in \Sigma$, è sufficiente che Σ che compare in (1.20) sia *tangente* in p ad una superficie caratteristica.

Osservazioni 1.6.
(1) Indipendentemente dal teorema di Cauchy-Kovalevskaja, possiamo concludere, che il problema di Cauchy (1.20) *non ha alcuna soluzione* quando i dati di Cauchy sono assegnati su una superficie caratteristica Σ (determinata localmente da $t = 0$), se essi non soddisfano una certa equazione supplettiva. Infatti, passando a coordinate Riemanniane in modo che l'equazione differenziale si possa scrivere come in (1.22), il fatto che Σ sia caratteristica e che quindi $a'^{tt}(0,\boldsymbol{\xi}) = 0$, implica che l'equazione in (1.22) si riduca, per $t = 0$, a:

$$-\sum_{i=2}^{n} 2a'^{it}(0,\boldsymbol{\xi})\frac{\partial u_1}{\partial \xi^i} - \sum_{i,j=2}^{n} a'^{ij}(0,\boldsymbol{\xi})\frac{\partial^2 u_0}{\partial \xi^i \partial \xi^j}$$

$$+\Phi'\left(0,\boldsymbol{\xi},u_0(\boldsymbol{\xi}),u_1(\boldsymbol{\xi}),\frac{\partial u_1}{\partial \xi^2},\cdots,\frac{\partial u_1}{\partial \xi^n},\cdots,\frac{\partial^2 u_0}{\partial \xi^2\partial \xi^2},\cdots\frac{\partial^2 u_0}{\partial \xi^n\partial \xi^n}\right) = 0\,.$$

Si osservi che questa condizione coinvolge unicamente quantità assegnate e non la soluzione, incognita, dell'equazione. Se questa equazione non è soddisfatta dalle condizioni di Cauchy non

può , evidentemente, esserci alcuna soluzione del problema di Cauchy.
(2) Mostriamo ora, con un esempio elementare, che viceversa vi sono casi in cui assegnare dati di Cauchy su superfici caratteristiche per un'equazione iperbolica comporta che esistano *infinite soluzioni* al problema posto. Consideriamo il problema di Cauchy per $u \in C^2(\mathbb{R}^2)$ dove $\mathbb{R}^2$ ha coordinate standard (x, y):

$$\begin{cases} u_{xy} = 0 \,, \\ u(x,0) = u_0(x) \,, \\ u_y(x,0) = u_1(x) \,, \end{cases} \tag{1.27}$$

dove $u_0 \in C^2(\mathbb{R})$ e $u_1 \in C^1(\mathbb{R})$ sono assegnate.
La matrice caratteristica A dell'equazione considerata è simmetrica, costante e vale a meno di un fattore moltiplicativo inessenziale:

$$A(x,y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \,.$$

Gli autovalori di A sono ± 1. Questo significa che esiste una matrice ortogonale R, 2×2, reale tale che:

$$RAR^t = \mathrm{diag}(1, -1) \,.$$

Di conseguenza l'equazione è ovunque di tipo iperbolico normale. Le superfici caratteristiche sono ora *curve caratteristiche*. Determiniamole. Consideriamo la solita funzione $S = S(x, y)$ i cui zeri determinano le curve caratteristiche. L'equazione (1.26) si riduce ora a:

$$\frac{\partial S}{\partial x}\frac{\partial S}{\partial y} = 0 \,.$$

La soluzione $S =$ *costante* non determina alcuna caratteristica in quanto l'insieme $S(x, y) = 0$ è vuoto se la costante è non nulla, oppure è tutto $\mathbb{R}^2$ se la costante è nulla, ma in tale caso $dS = 0$ ovunque. Pertanto deve essere $S = S_1(x)$ con $dS_1/dx \neq 0$ sull'insieme:

$$\{(x,y) \in \mathbb{R}^2 \ \ |S_1(x) = 0 \ \} \,,$$

oppure $S = S_2(y)$ con $dS_2/dy \neq 0$ sull'insieme:

$$\{(x,y) \in \mathbb{R}^2 \ \ |S_2(y) = 0 \ \} \,.$$

In particolare, per ogni coppia di costanti $c, d \in \mathbb{R}$, le funzioni $S_1(x) = x - c$ e $S_2(x) = y - d$ soddisfano le condizioni poste. In definitiva le rette $x =$ costante e $y =$ costante sono curve caratteristiche.
In particolare la retta $y = 0$ è una curva caratteristica e pertanto le condizioni di Cauchy del problema (1.27) sono assegnate su una superficie caratteristica.
La soluzione generale dell'equazione $u_{xy} = 0$ in $\mathbb{R}^2$ con $u \in C^2(\mathbb{R}^2)$ è (provarlo per esercizio):

$$u(x,y) = f(x) + g(y) \,,$$

dove $f, g \in C^2(\mathbb{R})$ sono funzioni arbitrarie. Pertanto ogni eventuale soluzione di (1.27) si deve ricondurre a questa forma. In particolare dovrà essere:

$$\begin{aligned} u(x,0) &= f(x) + g(0) = u_0(x)\,, & (1.28)\\ u_y(x,0) &= g'(0) = u_1(x)\,. & (1.29) \end{aligned}$$

Concludiamo immediatamente da (1.29) che: *se u_1 non è una funzione costante, allora il problema di Cauchy (1.27) non ha soluzioni.*
Tuttavia, nel caso in cui u_1 sia una funzione costante, il problema (1.27) ha *infinite* soluzioni. Infatti ogni funzione della forma:

$$u(x,y) = u_1 y + h(y) + u_0(x)\,,$$

per $h \in C^2(\mathbb{R})$ *arbitrariamente* scelta purché $h(0) = 0$ e $h'(0) = 0$, risolve il problema di Cauchy (1.27).

Esempi 1.2.
(1) Consideriamo l'equazione delle onde su $\mathbb{R}^n$, con velocità di propagazione delle onde data dalla costante $c > 0$:

$$-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} + \Delta u = 0\,.$$

Le coordinate su $\mathbb{R}^n$ sono state decomposte come: $x = (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^{n-1}$. L'equazione è scritta in forma normale rispetto alla variabile t nel modo seguente:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u\,.$$

Ci si aspetta pertanto, che assegnando come dati di Cauchy $u(0,\mathbf{x})$ e $\frac{\partial u}{\partial t}(0,\mathbf{x})$ esista una soluzione (ed una sola) del problema. Questo accade effettivamente, quando i dati di Cauchy sono in una determinata classe di funzioni. Passiamo a studiare le superfici caratteristiche sulle quali il problema di Cauchy è , in generale, mal posto. Per $(t_0, \mathbf{x}_0) \in \mathbb{R}^n$ fissato, si consideri la superficie conica di vertice $(t_0, \mathbf{x}_0)$:

$$\Gamma_{(t_0,\mathbf{x}_0)} := \left\{(t,\mathbf{x}) \in \mathbb{R}^n \;\mid\; c^2(t-t_0)^2 = (\mathbf{x}-\mathbf{x}_0)^2\right\}\,.$$

Posto $S(t,x) = c^2(t-t_0)^2 - (\mathbf{x}-\mathbf{x}_0)^2$, abbiamo che il luogo dei punti $(t,\mathbf{x})$ per cui $S(t,\mathbf{x}) = 0$ coincide con $\Gamma_{(t_0,\mathbf{x}_0)}$, inoltre

$$dS = 2c^2(t-t_0)dt + 2(\mathbf{x}-\mathbf{x}_0)\cdot d\mathbf{x}$$

non si annulla su $\Gamma_{(t_0,\mathbf{x}_0)} \setminus \{(t_0,\mathbf{x}_0)\}$. Infine, se a^{ij} definiscono la matrice caratteristica dell'equazione delle onde, si ha che:

$$\sum_{ij=1}^{n} a^{ij}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j} = -\frac{1}{c^2}\left(\frac{\partial S}{\partial t}\right)^2 + \sum_{i=2}^{n}\left(\frac{\partial S}{\partial x^i}\right)^2 = -2c^2(t-t_0)^2 + 2(\mathbf{x}-\mathbf{x}_0)^2\,,$$

si annulla esattamente su $\Gamma_{(t_0,\mathbf{x}_0)}$. Concludiamo che $\Gamma_{(t_0,\mathbf{x}_0)} \setminus \{(t_0,\mathbf{x}_0)\}$ è una superficie caratteristica dell'equazione delle onde. Tale superficie è detta *cono di luce* di vertice $(t_0,\mathbf{x}_0)$.
È importante precisare che i coni di luce non sono le sole superfici caratteristiche dell'equazione delle onde. Altre semplici superficie caratteristiche sono date dai piani di equazione

$$ct + \mathbf{x}\cdot\mathbf{b} = d\,,$$

dove il vettore costante $\mathbf{b} \in \mathbb{R}^{n-1}$ soddisfa $||\mathbf{b}|| = 1$ e $d \in \mathbb{R}$ è arbitrario. La funzione S in questo caso è banalmente:

$$S(t,\mathbf{x}) := ct + \mathbf{x}\cdot\mathbf{b} - d\,.$$

(2) Le equazioni ovunque di tipo ellittico in $\Omega \subset \mathbb{R}^n$, ed in particolare l'equazione di Poisson e quella di Laplace *non ammettono superfici caratteristiche.* Infatti, essendo la matrice dei coefficienti a^{ij} definita positiva, l'equazione (1.26) ammette come unica soluzione:

$$\frac{\partial S}{\partial x^i} = 0\,, \quad \text{per } i = 1,2,\ldots,n,$$

per cui, di conseguenza $dS = 0$ ovunque sul dominio di S. La non esistenza di superfici caratteristiche non significa, come già sottolineato, che i problemi di Cauchy siano sempre ben posti.
(3) Se $\Omega := \mathbb{R}^n$, consideriamo infine l'equazione del calore:

$$-\frac{\partial u}{\partial t} + a^2\Delta u = q\,.$$

dove $a > 0$ è una costante e $q = q(t,\mathbf{x})$ una funzione assegnata. Sopra le coordinate su $\mathbb{R}^n$ sono ancora state decomposte come: $x = (t,\mathbf{x}) \in \mathbb{R}\times\mathbb{R}^{n-1}$. In questo caso, la (1.26) si riduce a:

$$\sum_{i=2}^{n}\left(\frac{\partial S}{\partial x^i}\right)^2 = 0\,,$$

che ammette come soluzione ogni funzione regolare $S = S(t)$. In particolare, per ogni fissata costante $c \in \mathbb{R}$, la funzione $S(t,\mathbf{x}) = t - c$ soddisfa tutti i requisiti per definire superfici caratteristiche. Concludiamo che i piani $t =$ costante sono superficie caratteristiche dell'equazione del calore.

Capitolo 2

Equazioni Ellittiche e funzioni armoniche in $\mathbb{R}^n$: risultati elementari.

In questo capitolo affronteremo il problema della ricerca delle soluzioni di una certa classe di equazioni differenziali alle derivate parziali ellittiche e dello studio di alcune delle loro proprietà.

2.0.5 Il problema fisico dell'elettrostatica e le equazioni di Poisson e Laplace.

Noi ci occuperemo dell'analisi matematica del problema fondamentale dell'elettrostatica. Il *regime elettrostatico* si ottiene studiando il campo $\mathbf{E}$ unicamente tramite le prime due equazioni di Maxwell. In tale ambito si assume che $\mathbf{E}$ unitamente a ρ, pensata come funzione assegnata, ed al campo $\mathbf{B}$ (che in tal modo sparisce dalle equazioni suddette) siano funzioni *indipendenti dal tempo*.
Consideriamo un insieme aperto $\Omega \subset \mathbb{R}^3$ non vuoto e connesso (anche se alcuni teoremi che dimostreremo in seguito sono validi anche se Ω non è connesso). La seconda equazione di Maxwell in forma integrale, se si assume $\mathbf{E} = \mathbf{E}(\mathbf{x})$ di classe $C^0(\Omega)$, assicura che $\mathbf{E}$ sia conservativo. Pertanto, in queste ipotesi esisterà una funzione $\varphi : \Omega \to \mathbb{R}$ di classe $C^1(\Omega)$, determinata a meno di costanti additive, tale che:

$$\mathbf{E}(\mathbf{x}) = -\nabla_{\mathbf{x}}\varphi \,. \tag{2.1}$$

Osservazioni 2.1. La seconda equazione in forma differenziale porta allo stesso risultato se si assume che $\mathbf{E}$ sia $C^1(\Omega)$ e che Ω sia *semplicemente connesso* (per esempio $\Omega = \mathbb{R}$). Questo fatto mostra che le equazioni di Maxwell differenziali implicano la validità di quelle integrali solo sotto opportune ipotesi topologiche sul dominio spaziale Ω.

La funzione φ è detta **potenziale elettrostatico**. Determinare φ significa determinare $\mathbf{E}$, pertanto passiamo a studiare φ che, essendo un campo scalare, è più facile da maneggiare di un campo vettoriale quale è $\mathbf{E}$. Supponiamo di conoscere la densità di carica elettrica in Ω,

descritta da una funzione $\rho \in C^0(\Omega)$. Se assuminamo che φ sia di classe $C^2(\Omega)$ (cioè $\mathbf{E}$ di classe $C^1(\Omega)$), la prima equazione di Maxwell in forma differenziale, tenuto conto della (2.1) ed del fatto che

$$\nabla \cdot \nabla = \Delta \,,$$

implica immediatamente che φ soddisfi su Ω l'**equazione di Poisson**

$$\Delta\varphi = -4\pi\rho \quad \text{ossia} \quad \sum_{i,j=1}^{3} \delta^{ij} \frac{\partial\varphi}{\partial x^i \partial x^j} = -4\pi\rho \,. \tag{2.2}$$

Tale equazione è un tipico esempio di equazione ellittica, infatti la forma quadratica associata al termine del secondo ordine è costantemente definita positiva.
In particolare se la densità di carica è nulla in tutti i punti della regine Ω, cioè $\rho = 0$, allora l'equazione (2.2) diviene l'**equazione di Laplace**:

$$\Delta\varphi = 0 \,. \tag{2.3}$$

Come già detto, la conoscenza del potenziale φ comporta anche la conoscenza del campo elettrico e quindi la soluzione del nostro problema. Le tre domande fondamentali alle quali un fisico vorrebbe avere risposta (dai matematici) sono: data una generica densità di carica $\rho \in C^0(\Omega)$, esiste la soluzione dell'equazione (2.2)? È unica? Come possiamo calcolarla?
Per dare una risposta positiva alle prime due domande (esistenza e unicità) è necessario fare delle ipotesi sulla struttura della regione Ω e imporre delle condizioni aggiuntive che il potenziale φ dovrà soddisfare: le *condizioni al contorno*. Sappiamo che non possiamo dare condizioni di Cauchy complete perché il problema risulta, in generale essere malposto. Dobbiamo pertanto "indebolire" le condizioni assegnate nel problema di Cauchy. Ci sono diversi modi di far ciò. Vediamo i due esempi più importanti.

Problema di Dirichlet. Risolvere il problema di Dirichlet significa risolvere l'equazione di Poisson su una regione Ω, insieme aperto a chiusura compatta imponendo i valori che il potenziale φ deve assumere sul bordo $\partial\Omega$ di Ω:

$$\begin{cases} \Delta\varphi = f \quad \text{su } \Omega \,, \\ \varphi \restriction_{\partial\Omega} = \varphi_0 \ \text{(assegnato)} \,. \end{cases} \tag{2.4}$$

Osservazioni 2.2.
(1) Una tipica situazione fisica il cui è necessario risolvere un problema di Dirichlet è quella in cui $\partial\Omega$ è una superficie conduttrice, sulla quale il potenziale elettrico ha un valore costante che può essere assegnato arbitrariamente dall'esterno collegando la superficie ad una batteria e in Ω è anche presente una densità di carica $\rho = f/4\pi$ assegnata. Vedremo che in questo caso la soluzione φ del problema è unica, e quindi il campo $\mathbf{E} = -\nabla\varphi$ è unicamente determinato.
(2) C'è anche un altro caso fisicamente importante in cui rientra l'equazione di Laplace. Si tratta del caso in cui si vuole determinare la temperatura T in un mezzo omogeneo di volume Ω limitato (pensato come un aperto a chiusura compatta di $\mathbb{R}^3$), quando è tenuta fissa tramite termostati,

ed è nota, la temperatura al contorno del mezzo $\partial\Omega$ e il sistema fisico si trova in situazione di regime (non ci sono più variazioni di temperatura nel tempo). In tal caso, all'interno di Ω, la temperatura T soddisfa l'equazione di Laplace e pertanto il problema fisico si riduce ad un problema di Dirichlet:

$$\begin{cases} \Delta T = 0 \quad \text{su } \Omega\,, \\ T \restriction_{\partial\Omega} = T_0 \ \text{(funzione assegnata)}\,. \end{cases} \tag{2.5}$$

Problema di Neumann. Risolvere il problema di Neumann significa risolvere l'equazione di Poisson su una regione Ω, con Ω aperto a chiusura compatta e bordo $\partial\Omega$ regolare, imponendo i valori che la derivata normale alla superficie potenziale $\frac{\partial\varphi}{\partial n} = \nabla\varphi \cdot \mathbf{n}$ deve assumere sul bordo $\partial\Omega$ di Ω:

$$\begin{cases} \Delta\varphi = f \quad \text{su } \Omega\,, \\ \dfrac{\partial\varphi}{\partial\mathbf{n}} \restriction_{\partial\Omega} = \varphi_1 \ \text{(assegnato)}\,. \end{cases} \tag{2.6}$$

Osservazioni 2.3.
(1) Notiamo che nei problemi di elettrostatica $-\nabla\varphi \cdot \mathbf{n}$ è la componente del campo elettrico ortogonale alla superficie $\partial\Omega$.
(2) Una tipica situazione fisica il cui è necessario risolvere un problema di Neumann è quella in cui $\partial\Omega$ è una superficie conduttrice ed è anche presente in Ω una densità di carica $\rho = -f/4\pi$ assegnata. (In tal caso si dimostra che il campo elettrico è sempre ortogonale a $\partial\Omega$ ed è proporzionale alla densità di carica superficiale su $\partial\Omega$). Vedremo che in questo caso la soluzione φ del problema è unica a meno di una costante additiva, e quindi il campo $\mathbf{E} = -\nabla\varphi$ è *comunque* unicamente determinato.

2.1 Principio del massimo per funzioni armoniche e principio del massimo generalizzato.

2.1.1 Funzione armoniche e sub armoniche in $\mathbb{R}^n$.

Le funzioni reali di classe C^2 che soddisfano l'equazione di Laplace, anche in dimensione maggiore di 3, sono di grandissima rilevanza in matematica, per le loro molteplici proprietà analoghe a quelle delle funzioni analitiche (*olomorfe*) complesse di variabile complessa [6, 3]. Queste funzioni sono dette *funzioni armoniche*.

Definizione 2.1. Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto e $\varphi : \Omega \to \mathbb{R}$. φ è detta **armonica** se $\varphi \in C^2(\Omega)$, e soddisfa $\Delta\varphi = 0$ in Ω. $\diamondsuit$

Le funzioni *subarmoniche* hanno minor rilevanza ma sono un utile strumento tecnico in alcune dimostrazioni.

Definizione 2.2. Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto e $\varphi : \Omega \to \mathbb{R}$. φ è detta **subarmonica** se $\varphi \in C^2(\Omega)$ e soddisfa $\Delta\varphi \geq 0$ in Ω. $\diamondsuit$

Osservazioni 2.4. A volte si richiede che l'insieme aperto non vuoto Ω sul quale sono definite le funzioni armoniche e sub armoniche sia anche connesso. Noi non faremo questa assunzione dato che non è strettamente necessaria. Quando essa risulterà necessaria ci ridurremo a lavorare in una componente connessa di Ω.

Esempi 2.1.
1. In dimensione $n = 1$ le funzioni armoniche sono tutte e sole le funzioni che si restringono a funzioni lineari (non necessariamente omogenee) su ogni componente connessa del dominio. Infatti gli aperti sono unioni di aperti connessi e gli aperti connessi di $\mathbb{R}$ sono gli intervalli aperti (a, b), inoltre, se su (a, b) vale, per $\varphi \in C^2((a, b))$:

$$\frac{d^2}{dx^2}\varphi(x) = 0$$

allora, integrando,

$$\frac{d}{dx}\varphi(x) = m \quad \text{costante},$$

per cui, per qualche costante $q \in \mathbb{R}$:

$$\varphi(x) = mx + q\,, \quad \text{per ogni } x \in (a, b).$$

Viceversa ogni funzione lineare su (a, b) è sicuramente di classe $C^2((a, b))$ ed ha derivata seconda nulla. Su ogni componente connessa di un aperto non vuoto $\Omega \subset \mathbb{R}$, ogni funzione armonica su Ω è una funzione lineare non omogenea. Si osservi che le costanti m, q possono essere diverse a seconda della componente connessa di Ω considerata.
Per $n = 1$, le funzioni subarmoniche sono invece le funzioni di classe C^2 definite in aperti non vuoti e ivi *convesse*.
2. Sia $n = 2$. Esiste un legame interessante tra funzioni analitiche complesse e funzioni reali armoniche. Si consideri $f : \Omega \to \mathbb{C}$ con $\Omega \subset \mathbb{C}$ aperto, f è detta **analitica** oppure, indifferentemente, **olomorfa**[1] su Ω, se per ogni $z_0 \in \Omega$ la funzione f ammette sviluppo di Taylor centrato in z_0 e convergente a f in un intorno aperto di z_0. L'esistenza della serie di Taylor necessita in particolare dell'esistenza della derivata:

$$f'(z_0) := \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \in \mathbb{C}\,, \tag{2.7}$$

per ogni $z_0 \in \Omega$. Il limite è definito nella topologia di $\mathbb{R}^2$: per ogni $\epsilon > 0$ esiste $\delta > 0$ t.c. se $|z - z_0| < \delta$ allora:

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| < \epsilon\,,$$

[1] Il fatto che esistano due nomi per la stessa classe di funzioni è dovuto al particolare percorso storico che la teoria delle funzioni di variabile complessa ha seguito. Inizialmente si davano due differenti definizioni per le funzioni olomorfe e per quelle analitiche, più tardi è stato dimostrato che si tratta della stessa classe di funzioni.

e quindi è uniforme in tutte le direzioni. In particolare quindi, se $f'(z_0)$ esiste nel senso scritto sopra, può essere calcolata derivando lungo una fissata direzione. Decomponendo (2.7) in parte reale ed immaginaria, si può scrivere:

$$z = x + iy \in \Omega, \quad f(z) = f(x+iy) = u(x,y) + iv(x,y)$$

dove $u, v : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ sono la parte reale e immaginaria della funzione f. Tenendo conto dell'indipendenza direzionale del limite per calcolare $f'(z_0)$, abbiamo:

$$f'(z_0) = \lim_{h\to 0} \frac{f(x_0+iy_0+h) - f(x_0+iy_0)}{h} = \lim_{h\to 0} \frac{f(x_0+iy_0+ih) - f(x_0+iy_0)}{ih} \,,$$

Sostituendo la decomposizione di f in parte reale ed immaginaria, abbiamo dunque l'identità :

$$\frac{\partial}{\partial x}\left(u(x,y) + iv(x,y)\right) = \frac{1}{i}\frac{\partial}{\partial y}\left(u(x,y) + iv(x,y)\right) \,,$$

dove le derivate sono calcolate nel generico punto $(x_0, y_0) \in \Omega \subset \mathbb{R}^2$. Tenendo conto che u e v sono reali, raccogliendo separatamente parte reale ed immaginaria nell'identità trovata, abbiamo immediatamente che devono valere le **condizioni di Cauchy-Riemann**:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \,, \quad \text{per ogni } (x,y) \in \Omega \,. \tag{2.8}$$

Notiamo che u e v sono sicuramente $C^2(\Omega)$, in realtà sono sempre $C^\infty(\Omega)$, dato che f è analitica e quindi infinitamente differenziabile. Abbiamo infine direttamente dalle condizioni di Cauchy-Riemann:

$$\Delta u = 0 \,, \qquad \Delta v = 0.$$

Concludiamo che la parte reale ed immaginaria di ogni funzione olomorfa è una funzione armonica.

3. In dimensione $n > 1$ le funzioni armoniche sono moltissime. Alcuni polinomi di grado n sono armonici in dimensione $n > 1$: $P(x^1, \cdots, x^n) = x^1 \cdots x^n$. Vi sono anche funzioni non polinomiali, come per esempio, se $n > 1$:

$$\varphi(x^1, \ldots, x^n) = \sin\left(\sqrt{n-1}x^1\right) \sinh(x^2 + \cdots + x^n) \,.$$

Osservazioni 2.5.

(1) Si può dimostrare che condizione necessaria e sufficiente – e la sufficienza è uno dei risultati più notevoli della teoria delle funzioni analitiche complesse – affinché $f : \Omega \to \mathbb{C}$, con $\Omega \subset \mathbb{C}$ aperto non vuoto, sia olomorfa su Ω è che, per ogni $z_0 \in \Omega$, esista la derivata $f'(z_0)$, definita in (2.7).

(2) Si può provare [6], in riferimento a $f : \Omega \to \mathbb{C}$, con $\Omega \subset \mathbb{C}$ aperto non vuoto, che f è olomorfa in Ω se e solo se la sua parte reale ed immaginaria u e v soddisfano: (i) $u, v \in C^1(\Omega)$ (pensando Ω come sottoinsieme di $\mathbb{R}^2$) unitamente a (ii) le condizioni di Cauchy-Riemann (2.8) su Ω.

(3) Le funzioni olomorfe sono molto comuni: tutti i polinomi della variabile $z \in \mathbb{C}$ sono funzioni

olomorfe su $\mathbb{C}$, i rapporti tra polinomi sono funzioni olomorfe su tutto $\mathbb{C}$ escludendo gli zeri del polinomio a denominatore. Vi sono poi funzioni olomorfe definite sommando serie che estendono nel piano complesso funzioni reali di variabile reale. Per esempio:

$$e^z := \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

estende a valori complessi la funzione esponenziale $\mathbb{R} \ni \mapsto e^x$ definendo una funzione olomorfa su tutto il piano complesso. Nello stesso modo, cioè usando la stessa serie di Taylor che le definisce nei reali, ma valutandola per valori complessi della variabile, si defniscono le funzioni olomorfe *sinz* e *cosz* su tutto il piano complesso e si verifica la relazione di Eulero valida per tutti i complessi $z = x + iy \in \mathbb{C}$:

$$e^z = e^x(\cos y + i \sin y) \, .$$

Sono infine olomorfe le composizioni di funzioni olomorfe. Quindi, per esempio:

$$\mathbb{C} \setminus \{0\} \ni z \mapsto \sin\left(e^{\frac{z^2}{z-1}}\right) \, ,$$

è una funzione olomorfa sul dominio indicato.

2.1.2 Principio del massimo (in forma debole).

Nel caso $n = 1$, le funzioni armoniche sono, come visto sopra, della forma $\varphi : (a, b) \ni x \mapsto mx+q$. Notiamo che, nel caso considerato, se estendiamo φ all'intervallo chiuso $[a, b]$, supposto a, b finiti, accade che il massimo ed il minimo di φ sono assunti sul bordo di tale intervallo. Questa è una notevole proprietà delle funzioni armoniche che vale nel caso generale e cade sotto il nome di *principio del massimo* (anche se è un teorema). Si osservi che, sempre nel caso $n = 1$, ogni funzione subarmonica $\varphi : (a, b) \to \mathbb{R}$ (che è convessa come prima precisato), se è estendibile per continuità ai valori estremi dell'intervallo (a, b) supposto finito, assume valore massimo in uno dei due estremi. Anche questa è una proprietà generale delle funzioni subarmoniche che useremo per provare il principio del massimo per funzioni armoniche e che enunciamo in un unico teorema.

Teorema 2.1. **(Principio del massimo)**. *Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto con $\overline{\Omega}$ compatto. Sia $\varphi : \overline{\Omega} \to \mathbb{R}$ di classe $C^0(\overline{\Omega})$. Allora vale quanto segue.*
(a) *Se φ è subarmonica in Ω allora:*

$$\max_{\overline{\Omega}} \varphi = \max_{\partial\Omega} \varphi \, .$$

(b) *Se φ è armonica in Ω allora:*

$$\max_{\overline{\Omega}} \varphi = \max_{\partial\Omega} \varphi \, , \qquad \min_{\overline{\Omega}} \varphi = \min_{\partial\Omega} \varphi \qquad e \qquad \max_{\overline{\Omega}} |\varphi| = \max_{\partial\Omega} |\varphi| \, .$$

$\diamondsuit$

Dimostrazione. Prima di tutto notiamo che l'ipotesi di compattezza di $\overline{\Omega}$ (e quindi del sottoinsieme chiuso $\partial\Omega$) assicura l'esistenza dei massimi e minimi di cui si parla nella tesi essendo essi relativi a funzioni continue. Supponiamo inizialmente che $\Delta\varphi > 0$ su Ω. Sia $x_0 \in \bar{\Omega}$ un punto di massimo assoluto. Proviamo che $x_0 \in \partial\Omega$ e quindi la validità della tesi (a) nel caso $\Delta\varphi > 0$.
Se per assurdo fosse $x_0 \in \Omega$, allora la matrice hessiana di coefficienti $\frac{\partial^2\varphi}{\partial x^i \partial x^j}|_{x_0}$ sarebbe semidefinita negativa e quindi la sua traccia sarebbe non positiva: $\Delta\varphi|_{x_0} \leq 0$, che contraddice $\Delta\varphi|_{x_0} > 0$, quindi $x_0 \in \partial\Omega$.
Consideriamo ora il caso $\Delta\varphi \geq 0$ e definiamo la funzione $\tilde{\varphi} = \varphi + \epsilon|x|^2$, dove $\epsilon > 0$. Allora $\Delta\tilde{\varphi} = \Delta\varphi + 2n\epsilon > 0$, quindi applicando il risultato appena dimostrato, si ha per ogni fissato $x \in \overline{\Omega}$:

$$\tilde{\varphi}(x) \leq \max_{\partial\Omega} \tilde{\varphi} \leq \max_{\partial\Omega} \varphi + \epsilon R^2,$$

dove $R^2 = \max_{x\in\partial\Omega} |x|^2$. Quindi $\forall x \in \overline{\Omega}$

$$\varphi(x) \leq \varphi(x) + \epsilon|x|^2 \leq \max_{\partial\Omega} \varphi + \epsilon R^2,$$

da cui:

$$\varphi(x) \leq \max_{\partial\Omega} \varphi + \epsilon R^2 \ .$$

Dato che, per ogni fissato $x \in \overline{\Omega}$, ciò vale per ogni $\epsilon > 0$, dovrà anche essere, per quel valore di x:

$$\varphi(x) \leq \max_{\partial\Omega} \varphi$$

da cui segue la tesi in (a). Se φ è armonica allora φ e $-\varphi$ sono subarmoniche, da cui:

$$\max_{\partial\Omega} \varphi = \max_{\overline{\Omega}} \varphi,$$

$$\min_{\partial\Omega} \varphi = -\max_{\partial\Omega}(-\varphi) = -\max_{\overline{\Omega}}(-\varphi) = \min_{\overline{\Omega}} \varphi.$$

Inoltre, dato che $\max|\varphi| = \max(|\max\varphi|, |\min\varphi|)$ vale anche:

$$\max_{\partial\Omega} |\varphi| = \max_{\overline{\Omega}} |\varphi|.$$

□

2.1.3 Principio del massimo generalizzato.

Mostriamo ora che il teorema precedente si generalizza a funzioni che non sono necessariamente armoniche, ma che sono soluzioni di una particolare classe di equazioni del secondo ordine lineari ed ellittiche.

Teorema 2.2. **(Principio del massimo generalizzato)**. *Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto con $\overline{\Omega}$ compatto. Sia:*

$$L = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{k=1}^{n} b^k(x) \frac{\partial}{\partial x^k}$$

un operatore differenziale del secondo ordine tale che:

(i) la matrice caratteristica di coefficienti $a^{ij}(x)$ è ovunque definita positiva su Ω;

(ii) $\sum_{k=1}^{n} x^k b^k(x) \geq 0$ per ogni $x \in \Omega$ (in particolare può essere $b^k(x) = 0$ per $k = 1, 2, \ldots, n$ ed ogni $x \in \Omega$).

Si consideri una funzione $\varphi : \overline{\Omega} \to \mathbb{R}$ di classe $C^0(\overline{\Omega}) \cap C^2(\Omega)$. Allora valgono i fatti seguenti.

(a) *Se $L\varphi \geq 0$ su Ω allora:*

$$\max_{\overline{\Omega}} \varphi = \max_{\partial\Omega} \varphi$$

(b) *Se $L\varphi = 0$ su Ω allora vale, in aggiunta alla precedente, anche*:

$$\min_{\overline{\Omega}} \varphi = \min_{\partial\Omega} \varphi \qquad e \qquad \max_{\overline{\Omega}} |\varphi| = \max_{\partial\Omega} |\varphi| \ .$$

$\diamondsuit$

Dimostrazione. Si procede come nella dimostrazione del teorema precedente. Se vale $L\varphi > 0$ su Ω e $x_0 \in \Omega$ è un punto di massimo assoluto della funzione φ su $\overline{\Omega}$, allora $\nabla\varphi|_{x_0} = 0$ e la matrice di coefficienti $\frac{\partial^2\varphi}{\partial x^i \partial x^j}|_{x_0}$ è semidefinita negativa. Dato che la matrice $A(x_0)$ di coefficienti $a^{ij}(x_0)$ è definita positiva, per il teorema di Sylvester, si potrà scrivere come $A(x_0) = DID^t = DD^t$ dove D è una matrice quadrata non singolare. Inoltre, se $H(x_0)$ è la matrice quadrata simmetrica di coefficienti $\frac{\partial^2\varphi}{\partial x^i \partial x^j}|_{x_0}$, si ha (tenendo conto che la parte del prim'ordine di L non fornisce contributo in quanto tutte le derivate prime di φ si annullano in x_0):

$$L\varphi|_{x_0} = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2\varphi}{\partial x^i \partial x^j}|_{x_0} = \mathrm{tr}\left(A(x_0)H(x_0)\right) \ ,$$

per cui:

$$L\varphi|_{x_0} = \mathrm{tr}\left(DD^t H(x_0)\right) = \mathrm{tr}\left(D^t H(x_0) D\right) \ .$$

Dato che D è non singolare e $H(x_0)$ è semidefinita negativa, $D^t H(x_0) D$ sarà ancora semidefinita negativa[2] e quindi avrà autovalori non positivi. La traccia di tale matrice sarà dunque non positiva. In definitiva:

$$L\varphi|_{x_0} = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2\varphi}{\partial x^i \partial x^j}|_{x_0} \leq 0,$$

[2]Per ipotesi $u^t H(x_0) u \leq 0$ per ogni $u \in \mathbb{R}^n$, ma dato che $D : \mathbb{R}^n \to \mathbb{R}^n$ è biettiva, dovrà anche valere: $v^t D^t H(x_0) D v = (Dv)^t H(x_0) D v \leq 0$ per ogni $v \in \mathbb{R}^n$.

che è assurdo perché per ipotesi $L\varphi|_{x_0} > 0$.
Supponiamo ora che $L\varphi \geq 0$ su Ω e definiamo $\tilde{\varphi}(x) = \varphi(x) + \epsilon|x|^2$. Per ogni $\epsilon > 0$ vale:

$$L\tilde{\varphi}(x) = L\varphi(x) + 2\epsilon \, \mathrm{tr}A(x) + 2\epsilon \sum_{k=1}^{n} x^k b^k(x) > 0,$$

dove abbiamo usato il fatto che la traccia della matrice $A(x)$ di coefficienti $a^{ij}(x)$ è strettamente positiva in quanto, per ipotesi, tale matrice è definita positiva ed inoltre $\sum_{k=1}^{n} x^k b^k(x) \geq 0$ sempre per ipotesi. Per la prima parte della dimostrazione, abbiamo quindi che:

$$\max_{\overline{\Omega}} \tilde{\varphi} = \max_{\partial\Omega} \tilde{\varphi} \, ,$$

e allora, per ogni fissato $x \in \overline{\Omega}$:

$$\varphi(x) \leq \tilde{\varphi}(x) \leq \max_{\overline{\Omega}} \tilde{\varphi} \leq \max_{\partial\Omega} \varphi + \epsilon R^2,$$

con $R^2 = \max_{\partial\Omega} |x|^2$. La dimostrazione si conclude come quella del teorema precedente. □

Osservazioni 2.6. Si noti che non è stata fatta alcuna ipotesi di regolarità sulle funzioni $\Omega \ni x \mapsto a^{ij}(x)$ e $\Omega \ni x \mapsto b^k(x)$ usate nella definizione di L.

2.1.4 Due teoremi di unicità per il probelma di Dirichlet dal principio del massimo.

In questa sezione applichiamo il principio del massimo per dimostrare l'unicità delle soluzioni dell'equazione di Poisson (2.2) nel caso del problema di Dirichlet.

Teorema 2.3. **(Unicità per il problema di Dirichlet 1)**. *Si consideri il seguente problema di Dirichlet per la funzione $\varphi : \overline{\Omega} \to \mathbb{R}$, riferita all'aperto $\Omega \subset \mathbb{R}^n$ non vuoto a chiusura compatta:*

$$\begin{cases} \Delta\varphi = f & su \ \Omega \, , \\ \varphi \restriction_{\partial\Omega} = \psi \, , \end{cases} \qquad \varphi \in C^0(\overline{\Omega}) \cap C^2(\Omega) \, , \tag{2.9}$$

con $f \in C^0(\Omega)$, $\psi \in C^0(\partial\Omega)$ assegnate. Se esiste una soluzione φ al problema posto, questa è unica. ♢

Dimostrazione. Siano ϕ_1 e ϕ_2 due soluzioni del problema, dimostriamo che la funzione $\phi_1 - \phi_2$ è identicamente nulla: $\phi_1 - \phi_2 = 0$.
$\phi_1 - \phi_2$ è armonica su Ω, infatti $\phi_1 - \phi_2 \in C^2(\Omega)$ e $\Delta(\phi_1 - \phi_2) = 0$ su Ω; inoltre $(\phi_1 - \phi_2)|_{\partial\Omega} = 0$.
Per il principio del massimo

$$\max_{\overline{\Omega}} |\phi_1 - \phi_2| = \max_{\partial\Omega} |\phi_1 - \phi_2| = 0,$$

da cui $\phi_1 = \phi_2$ su $\overline{\Omega}$. $\diamondsuit$

Osservazioni 2.7.
(1) Notare che nel teorema precedente non abbiamo fatto alcuna ipotesi sulla regolarità di $\partial\Omega$, in particolare non è necessario che $\partial\Omega$ sia una superficie regolare. Non è necessario inoltre supporre che Ω sia connesso.
(2) Si supponga di essere riusciti a provare, e questo è possibile sotto opportune ipotesi di regolarità di $\partial\Omega$, che Ω ed f del teorema precedente sono tali che, per ogni $\psi \in C^0(\partial\Omega)$ esista una soluzione (e dunque una *sola* soluzione) φ del problema di Dirichlet considerato. In questo caso, la dimostrazione data del teorema dimostra anche che il problema di Dirichlet è ben posto nel senso di Hadamard (considerando ovviamente la sola condizione al bordo di Dirichlet), se si dota lo spazio delle condizioni iniziali e lo spazio delle soluzioni della topologia metrica indotta dalla norma $||\cdot||_\infty$. In realtà questa topologia non è appropriata nello spazio delle soluzioni dato che non considera le derivate delle funzioni φ (in linea di principio, si potrebbe avere una successione di dati al bordo che tende a zero uniformemente, mentre le derivate delle soluzioni associate non tendono ad alcun limite). Pertanto il problema della dipendenza continua dai dati iniziali non può essere affrontato con il solo principio del massimo.
(3) Con la stessa dimostrazione, ma usando il principio del massimo generalizzato si dimostra il seguente teorema più generale:

Teorema 2.4. *Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto con $\overline{\Omega}$ compatto. Sia:*

$$L = \sum_{i,j=1}^{n} a^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j} + \sum_{k=1}^{n} b^k(x)\frac{\partial}{\partial x^k}$$

un operatore differenziale del secondo ordine su Ω tale che:
(i) la matrice caratteristica di coefficienti $a^{ij}(x)$ è ovunque definita positiva su Ω;
(ii) $\sum_{k=1}^{n} x^k b^k(x) \geq 0$ per ogni $x \in \Omega$ (in particolare $b^k(x) = 0$ per $k = 1, 2, \ldots, n$ ed ogni $x \in \Omega$).
Si consideri il problema di Dirichlet per $\varphi : \overline{\Omega} \to \mathbb{R}$:

$$\begin{cases} L\varphi = f \quad su\ \Omega\,, \\ \varphi \restriction_{\partial\Omega} = \psi\,, \end{cases} \qquad \varphi \in C^0(\overline{\Omega}) \cap C^2(\Omega)\,, \tag{2.10}$$

con $\psi \in C^0(\partial\Omega)$ assegnate. Se esiste una soluzione al problema posto, questa è unica. $\diamondsuit$

Passiamo a provare un teorema di unicità della soluzione del problema di Dirichlet su una regione non limitata (problema di Dirichlet *esterno*) nell'ipotesi che la soluzione φ tenda a 0 quando $x \to \infty$ *uniformemente* (nelle possibili direzioni), in altre parole, per ogni $\epsilon > 0$ esiste $R_\epsilon > 0$ tale che $|\varphi(x)| < \epsilon$ se $||x|| > R_\epsilon$.

Teorema 2.5. **(Unicità per il problema di Dirichlet 2)**. *Se $\Omega \neq \emptyset$ è un aperto di $\mathbb{R}^n$ a chiusura compatta, si consideri il problema di Dirichlet per $\varphi : \overline{\mathbb{R}^n \setminus \overline{\Omega}} \to \mathbb{R}$, dove vale*

$\varphi \in C^2(\mathbb{R}^n \setminus \overline{\Omega}) \cap C^0\left(\overline{\mathbb{R}^n \setminus \overline{\Omega}}\right)$:

$$\begin{cases} \Delta\varphi = f \quad su\ \mathbb{R}^n \setminus \overline{\Omega} \\ \varphi \restriction_{\partial\Omega} = \psi\,, \quad \varphi \to 0 \ uniformemente\ quando\ x \to \infty. \end{cases} \tag{2.11}$$

dove $f \in C^0(\mathbb{R}^n \setminus \Omega)$ e $\psi \in C^0(\partial\Omega)$ sono funzioni assegnate.
Se esiste una soluzione essa è unica. $\diamondsuit$

Dimostrazione. Sia B_R una palla di raggio $R > 0$ centrata nell'origine di $\mathbb{R}^n$ e contenente $\overline{\Omega}$. Siano ϕ_1 e ϕ_2 due soluzioni del problema 2.11 tendenti uniformemente a 0 quando $x \to \infty$. Allora $\phi_1 - \phi_2 = 0$ su $\partial\Omega$ e $|\phi_1 - \phi_2|_{|\partial B_R} \to 0$ quando $R \to \infty$ uniformemente. $\phi_1 - \phi_2$ inoltre è armonica in $B_R \setminus \bar{\Omega}$, per cui, fissato $x \in \mathbb{R}^n \setminus \Omega$,

$$|\phi_1(x) - \phi_2(x)| \leq \max_{\partial\Omega \cup \partial B_R} |\phi_1 - \phi_2| = \max_{\partial B_R} |\phi_1 - \phi_2| \to 0 \quad \text{se} \quad R \to \infty\,.$$

Ciò prova che $\phi_1(x) = \phi_2(x)$ per ogni $x \in \mathbb{R}^n \setminus \Omega$.$\square$

Osservazioni 2.8. Anche in questo caso, nello stesso modo, ma usando il principio del massimo generalizzato possiamo provare il seguente teorema più generale.

Teorema 2.6. *Se $\Omega \neq \emptyset$ è un aperto di $\mathbb{R}^n$ a chiusura compatta e si consideri il problema di Dirichlet per per $\varphi : \overline{\mathbb{R}^n \setminus \overline{\Omega}} \to \mathbb{R}$, dove vale $\varphi \in C^2(\mathbb{R}^n \setminus \overline{\Omega}) \cap C^0\left(\overline{\mathbb{R}^n \setminus \overline{\Omega}}\right)$:*

$$\begin{cases} L\varphi = f \quad su\ \mathbb{R}^n \setminus \overline{\Omega}\,, \\ \varphi \restriction_{\partial\Omega} = \psi\,, \quad \varphi \to 0 \ uniformemente\ quando\ x \to \infty. \end{cases} \tag{2.12}$$

dove $\psi \in C^0(\partial\Omega)$ è assegnata e

$$L = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{k=1}^{n} b^k(x) \frac{\partial}{\partial x^k}$$

è un operatore differenziale del secondo ordine tale che:
(i) la matrice caratteristica di coefficienti $a^{ij}(x)$ è ovunque definita positiva su $\mathbb{R}^n \setminus \overline{\Omega}$;
(ii) $\sum_{k=1}^{n} x^k b^k(x) \geq 0$ per ogni $x \in \Omega$ (in particolare $b^k(x) = 0$ per $k = 1, 2, \ldots, n$ ed ogni $x \in \mathbb{R}^n \setminus \overline{\Omega}$).
Se esiste una soluzione essa è unica. $\diamondsuit$

2.2 Le identità di Green le loro conseguenze elementari.

In questa sezione introduciamo delle formule utili dette di *identità di Green*, che hanno alcune importanti conseguenze sulle funzioni armoniche e sulle soluzioni dell'equazione di Poisson.

2.2.1 Identità di Green.

Diamo la forma delle identità di Green. Le ipotesi non sono il caso più generale possibile.

Teorema 2.7. *Sia $V \subset \mathbb{R}^n$ aperto, la cui chiusura $\overline{V}$ sia compatta e tale che il suo bordo ∂V sia una superficie regolare (di dimensione $n-1$) orientabile. Siano $\phi, \psi : \overline{V} \to \mathbb{R}$ due funzioni di classe $C^1(\overline{V})$. Allora valgono le due* **identità di Green**, *dove* $\mathbf{n}$ *è il versore normale a ∂V orientato in modo uscente:*

$$\int_V \phi \Delta \psi d^n x + \int_V \nabla \phi \cdot \nabla \psi d^n x = \oint_{\partial V} \phi \nabla \psi \cdot \mathbf{n} \, dS \tag{2.13}$$

se $\psi \in C^2(V)$,

$$\int_V (\phi \Delta \psi - \psi \Delta \phi) \, d^n x = \oint_{\partial V} (\phi \nabla \psi - \psi \nabla \phi) \cdot \mathbf{n} \, dS \tag{2.14}$$

se $\psi, \phi \in C^2(V)$.
Gli integrali $\int_V \phi \Delta \psi d^n x$ e $\int_V (\phi \Delta \psi - \psi \Delta \phi) \, d^n x$ sono da intendersi nel caso generale come integrali impropri di Riemann, oppure come integrali nel senso di Lebesgue, qualora le funzioni integrande siano (assolutamente) integrabili nel senso di Lebesgue. $\diamondsuit$

Dimostrazione. La prima identità di Green (2.13) si ottiene applicando la formula di Gauss alla funzione $\mathbf{E} : \overline{V} \to \mathbb{R}^n$ (che soddisfa le ipotesi per applicare il teorma di Gauss su Ω). La tesi segue subito dall'identità:

$$\nabla \cdot (\phi \nabla \psi) = \nabla \phi \cdot \nabla \psi + \psi \Delta \psi \, .$$

La seconda identità di Green (2.14) si ottiene dalla (2.13), scritta due volte invertendo, la seconda volta, il ruolo di ψ e ϕ e sottraendo membro a membro i risultati. $\square$

2.2.2 Conseguenze del teorema di Gauss e delle identità di Green: teorema di unicità per il problema di Neumann.

Vediamo ora altre proprietà delle funzioni armoniche e delle soluzioni dell'equazione di Poisson che derivano dal teorema di Gauss e dalle identità di Green.

Teorema 2.8. *Sia φ armonica in $\Omega \subset \mathbb{R}^n$ non vuoto aperto.*
Allora, per ogni superficie regolare $S \subset \Omega$ che sia il bordo orientabile di un aperto $V \subset \Omega$ con $\overline{V} \subset \Omega$ compatto, vale:

$$\oint_S \nabla \varphi \cdot \mathbf{n} \, dS = 0$$

dove $\mathbf{n}$ *è il versore normale uscente alla superficie S. Il risultato vale considerando anche il caso limite $S = \partial \Omega$ se $\overline{\Omega}$ è compatto, S regolare e $\varphi \in C^1(\overline{\Omega})$.* $\diamondsuit$

Dimostrazione. Dato che $\varphi \in C^2(\Omega)$, possiamo applicare la formula di Gauss al volume $V \subset \Omega$ con $\partial V = S \subset \Omega$:

$$\oint_S \nabla\varphi \cdot \mathbf{n}\, dS = \int_V \nabla \cdot \nabla\varphi d^n x = \int_V \Delta\varphi\, d^n x = 0\,.$$

Il caso limite segue nello stesso modo sempre dal teorema di Gauss. □

Conseguenza delle identità di Green è anche il seguente teorema di unicità della soluzione del problema di Neumann. Ricordiamo preventivamente un lemma.

Lemma 2.1. *Se $f : \Omega \to \mathbb{R}$ è differenziabile sull'aperto $\Omega \subset \mathbb{R}^n$ e per ogni punto di Ω vale $\nabla f = 0$ su Ω allora f è costante su ogni componente connessa di Ω.* ♢

Dimostrazione. Ogni componente connessa di un aperto è un insieme aperto connesso. Ogni insieme aperto connesso di $\mathbb{R}^n$ è connesso per archi differenziabili. Sia Ω_0 una componente connessa di Ω e $p, q \in \Omega_0$. Sia infine $\gamma : [a, b] \to \Omega_0$ una curva differenziabile tale che $\gamma(a) = p$ e $\gamma(b) = q$ e $\gamma([a, b]) \subset \Omega_0$. La funzione $g(t) := f(\gamma(t))$ per $t \in [a, b]$ è differenziabile e, nelle nostre ipotesi:

$$g'(t) = \dot\gamma(t) \cdot \nabla f(\gamma(t)) = 0\,, \quad \text{per ogni } t \in [a, b]\,.$$

Ne consegue (dal teorema di Lagrange) che g è costante ed in particolare $f(p) = g(\gamma(a)) = g(\gamma(b)) = f(q)$. L'arbitrarietà dei punti $p, q \in \Omega_0$ implica che f è costante su Ω_0. □

Teorema 2.9. **(Unicità per il problema di Neumann 1)**. *Si consideri il seguente problema di Neumann per la funzione $\varphi : \overline{\Omega} \to \mathbb{R}$:*

$$\begin{cases} \Delta\varphi = f \quad \text{su } \Omega\,, \\ \dfrac{\partial\varphi}{\partial\mathbf{n}} \restriction_{\partial\Omega} = \psi \end{cases} \qquad \varphi \in C^1(\overline{\Omega}) \cap C^2(\Omega) \tag{2.15}$$

con $f \in C^0(\Omega)$, $\psi \in C^0(\partial\Omega)$ assegnate, per $\emptyset \neq \Omega \subset \mathbb{R}^n$ aperto, connesso, con chiusura $\overline{\Omega}$ compatta e $\partial\Omega$ superficie $n-1$ dimensionale regolare e orientabile e $\mathbf{n}$ è il versore normale uscente. Allora, se esiste una soluzione φ al problema posto, questa è unica a meno di costanti addittive. □

Dimostrazione. Siano ϕ_1 e ϕ_2 due soluzioni del problema, dimostriamo che la funzione $u = \phi_1 - \phi_2$ è costante. Per dimostrare ciò , notiamo inizialmente che $\phi_1 - \phi_2$ è armonica su Ω, infatti $\phi_1 - \phi_2 \in C^2(\Omega)$ e $\Delta(\phi_1 - \phi_2) = 0$ su Ω. Applicando la prima identità di Green:

$$0 = \int_\Omega u\Delta u d^n x = -\int_\Omega \nabla u \cdot \nabla u\, d^n x + \oint_{\partial\Omega} u \frac{\partial u}{\partial\mathbf{n}}\, dS\,.$$

Dato che, nelle nostre ipotesi $\frac{\partial\phi_1 - \phi_2}{\partial\mathbf{n}}|_{\partial\Omega} = 0$ ossia $\frac{\partial u}{\partial\mathbf{n}}|_{\partial\Omega} = 0$, l'identità di sopra implica:

$$\int_\Omega ||\nabla u||^2 d^n x = 0\,,$$

e dunque, essendo ∇u continua, $\nabla u = 0$ su Ω. Concludiamo che, su ogni componente connessa di Ω, u deve essere costante. Questo implica immediatamente la tesi. $\square$

Osservazioni 2.9. Si consideri il problema di Neumann (2.15). Notiamo che dalla formula di Gauss applicata al campo vettoriale $\nabla\varphi$ segue

$$\int_\Omega \Delta\varphi d^n x = \int_\Omega \nabla \cdot \nabla\varphi \, d^n x = \oint_{\partial\Omega} \frac{\partial\varphi}{\partial \mathbf{n}} dS$$

dal fatto che φ deve essere soluzione del problema di Neumann abbiamo che

$$\int_\Omega \Delta\varphi d^n x = \int_\Omega f d^n x \qquad \text{e} \qquad \oint_{\partial\Omega} \frac{\partial\varphi}{\partial \mathbf{n}} dS = \oint_{\partial\Omega} \psi dS \, .$$

Combinando i due risultati otteniamo la proposizione seguente.

Proposizione 2.1. *Condizione necessaria affinchè il problema di Neumann (2.15) ammetta soluzione è che:*

$$\int_\Omega f d^n x = \oint_{\partial\Omega} \psi dS \, .$$

$\diamondsuit$

Questo esempio ci porta a concludere che non tutti i problemi contenenti l'equazione di Poisson ammettono soluzione.
Concludiamo con due teorema di unicità per il problema di Neumann *esterno*.

Teorema 2.10. **(Unicità per il problema di Neumann 2)**. *Se $\Omega \neq \emptyset$ è un aperto di $\mathbb{R}^n$ a chiusura compatta con bordo $\partial\Omega$ dato da una superficie $n-1$ dimensionale regolare ed infine $\mathbb{R}^n \setminus \overline{\Omega}$ è connesso, si consideri il problema di Neumann per $\varphi : \overline{\mathbb{R}^n \setminus \overline{\Omega}} \to \mathbb{R}$:*

$$\begin{cases} \Delta\varphi = f \quad su\ \mathbb{R}^n \setminus \overline{\Omega} \, , \\ \dfrac{\partial\varphi}{\partial \mathbf{n}} \restriction_{\partial\Omega} = \psi \\ |\varphi(x)| \leq K \, , \quad ||\nabla\varphi(x)|| \leq \dfrac{K}{(1+||x||)^{n-1+\alpha}} \quad su\ \mathbb{R}^n \setminus \overline{\Omega} \end{cases} \qquad \varphi \in C^1\left(\overline{\mathbb{R}^n \setminus \overline{\Omega}}\right) \cap C^2(\mathbb{R}^n \setminus \overline{\Omega}) \, , \tag{2.16}$$

con $f \in C^0(\mathbb{R}^n \setminus \overline{\Omega})$, $\psi \in C^0(\partial\Omega)$ assegnate e dove $\mathbf{n}$ è il versore normale uscente e $K > 0$, $\alpha > 0$ sono costanti reali dipendenti da φ in generale. Allora, se esiste una soluzione φ al problema posto, questa è unica a meno di costanti addittive. $\square$

Dimostrazione. Sia B_R una palla aperta di raggio $R > 0$ centrata nell'origine di $\mathbb{R}^n$ e contenente $\overline{\Omega}$. Siano ϕ_1 e ϕ_2 due soluzioni del problema considerato dimostriamo che la funzione $u := \phi_1 - \phi_2$ è costante. Per dimostrare ciò , notiamo inizialmente che, per costruzione u è

armonica su $\Omega_R := (\mathbb{R}^n \setminus \overline{\Omega}) \cap B_R$. Applicando la prima identità di Green:

$$0 = \int_{\Omega_R} u \Delta u d^n x = -\int_{\Omega_R} \nabla u \cdot \nabla u \, d^n x + \oint_{\partial\Omega} u \frac{\partial u}{\partial \mathbf{n}} \, dS + \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}} \, dS \, .$$

Dato che, nelle nostre ipotesi $\frac{\partial \phi_1 - \phi_2}{\partial \mathbf{n}}|_{\partial\Omega} = 0$ ossia $\frac{\partial u}{\partial \mathbf{n}}|_{\partial\Omega} = 0$, l'identità di sopra implica:

$$\int_{\Omega_R} ||\nabla u||^2 d^n x = \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}} \, dS \, .$$

D'altra parte

$$\left| \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}} \, dS \right| \leq \oint_{\partial B_R} |u \mathbf{n}_{\partial B_R} \cdot \nabla u| \, dS \leq \frac{C}{(1+R)^{n-1+\beta}} \oint_{\partial B_R} dS \, ,$$

per qualche $C, \beta > 0$. Nell'ultimo passaggio abbiamo usato:

$$|u \mathbf{n} \cdot \nabla u| = |u| \, |\mathbf{n} \cdot \nabla u| \leq |u| \, ||\nabla u|| = ||u \nabla u||$$

e quindi

$$||u \nabla u|| \leq |\phi_1| \, ||\nabla \phi_1|| + |\phi_2| \, ||\nabla \phi_2|| + |\phi_1| \, ||\nabla \phi_2|| + |\phi_2| \, ||\nabla \phi_1|| \, ,$$

e, dato che tutte le funzioni sono valutate per $||x|| = R$,

$$|\phi_1| \, ||\nabla \phi_1|| + |\phi_2| \, ||\nabla \phi_2|| + |\phi_1| \, ||\nabla \phi_2|| + |\phi_2| \, ||\nabla \phi_1|| \leq \frac{H}{(1+R)^{n-1+\beta}} \, ,$$

dove $H > 0$ e $\beta > 0$ esistono per ipotesi (per esempio β è la più piccola tra le costanti α di ϕ_1 e ϕ_2 e H è 4 volte il quadrato della più grande delle costanti K di ϕ_1 e ϕ_2). Torniamo alla disuguaglianza:

$$\left| \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}} \, dS \right| \leq \frac{H}{(1+R)^{n-1+\beta}} \oint_{\partial B_R} dS \, .$$

L'ultimo membro tende a 0 per $R \to +\infty$ in quanto l'ultimo integrale vale CR^{n-1} per qualche costante $C > 0$. Per cui:

$$\lim_{R \to +\infty} \int_{\Omega_R} ||\nabla u||^2 d^n x = 0 \, .$$

Il teorema della convergenza monotona, assicura che la funzione $x \mapsto ||\nabla u||^2$ sia integrabile su $\mathbb{R}^n \setminus \overline{\Omega} = \Omega_\infty$ con integrale nullo. Pertanto la funzione deve essere quasi ovunque nulla. Essendo $u \mapsto ||\nabla u||$ continua, deve essere $\nabla u = 0$ su $\mathbb{R}^n \setminus \overline{\Omega}$. Concludiamo che, su ogni componente connessa di Ω, u deve essere costante. Questo implica immediatamente la tesi. $\square$

Ecco il secondo teorema che fornisce l'unicità senza l'ambiguità della costante additiva arbitraria.

Teorema 2.11. **(Unicità per il problema di Neumann 3)**. *Se $\Omega \neq \emptyset$ è un aperto di $\mathbb{R}^n$ a chiusura compatta con bordo $\partial\Omega$ dato da una superficie $n-1$ dimensionale regolare ed infine $\mathbb{R}^n \setminus \overline{\Omega}$ è connesso, si consideri il problema di Neumann per $\varphi : \overline{\mathbb{R}^n \setminus \overline{\Omega}} \to \mathbb{R}$:*

$$\begin{cases} \Delta\varphi = f \quad su\ \mathbb{R}^n \setminus \overline{\Omega}\,, \\ \dfrac{\partial\varphi}{\partial\mathbf{n}} \restriction_{\partial\Omega} = \psi \\ |\varphi(x)| \leq \dfrac{K}{(1+||x||)^\alpha}\,, \quad ||\nabla\varphi(x)|| \leq \dfrac{K}{(1+||x||)^{n-1}} \quad su\ \mathbb{R}^n \setminus \overline{\Omega} \end{cases} \qquad \varphi \in C^1\left(\overline{\mathbb{R}^n \setminus \overline{\Omega}}\right) \cap C^2(\mathbb{R}^n \setminus \overline{\Omega}), \tag{2.17}$$

con $f \in C^0(\mathbb{R}^n \setminus \overline{\Omega})$, $\psi \in C^0(\partial\Omega)$ assegnate e dove $\mathbf{n}$ è il versore normale uscente e $K > 0$, $\alpha > 0$ sono costanti reali dipendenti da φ in generale. Allora, se esiste una soluzione φ al problema posto, questa è unica. $\square$

Dimostrazione. La dimostrazione procede analogamente a prima. Sia B_R una palla aperta di raggio $R > 0$ centrata nell'origine di $\mathbb{R}^n$ e contenente $\overline{\Omega}$. Siano ϕ_1 e ϕ_2 due soluzioni del problema considerato dimostriamo che la funzione $u := \phi_1 - \phi_2$ è costante. Per dimostrare ciò , notiamo inizialmente che, per costruzione u è armonica su $\Omega_R := (\mathbb{R}^n \setminus \overline{\Omega}) \cap B_R$. Applicando la prima identità di Green:

$$0 = \int_{\Omega_R} u\Delta u d^n x = -\int_{\Omega_R} \nabla u \cdot \nabla u\, d^n x + \oint_{\partial\Omega} u \frac{\partial u}{\partial \mathbf{n}}\, dS + \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}}\, dS\,.$$

Dato che, nelle nostre ipotesi $\frac{\partial \phi_1 - \phi_2}{\partial \mathbf{n}}|_{\partial\Omega} = 0$ ossia $\frac{\partial u}{\partial \mathbf{n}}|_{\partial\Omega} = 0$, l'identità di sopra implica:

$$\int_{\Omega_R} ||\nabla u||^2 d^n x = \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}}\, dS\,.$$

D'altra parte

$$\left| \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}}\, dS \right| \leq \oint_{\partial B_R} |u \mathbf{n}_{\partial B_R} \cdot \nabla u|\ dS \leq \frac{H}{(1+R)^{n-1+\beta}} \oint_{\partial B_R} dS\,,$$

per costanti $H, \beta > 0$. Nell'ultimo passaggio abbiamo usato:

$$|u\mathbf{n} \cdot \nabla u| = |u|\, |\mathbf{n} \cdot \nabla u| \leq |u|\, ||\nabla u|| = ||u \nabla u||$$

e quindi

$$||u\nabla u|| \leq |\phi_1|\, ||\nabla\phi_1|| + |\phi_2|\, ||\nabla\phi_2|| + |\phi_1|\, ||\nabla\phi_2|| + |\phi_2|\, ||\nabla\phi_1||\,,$$

e, dato che tutte le funzioni sono valutate per $||x|| = R$,

$$|\phi_1|\, ||\nabla\phi_1|| + |\phi_2|\, ||\nabla\phi_2|| + |\phi_1|\, ||\nabla\phi_2|| + |\phi_2|\, ||\nabla\phi_1|| \leq \frac{H}{(1+R)^\beta} \frac{1}{(1+R)^{n-1}}\,,$$

dove $H > 0$ e $\beta > 0$ esistono per ipotesi (β è la più piccola tra le costanti α di ϕ_1 e ϕ_2 e H è 4 volte il quadrato della più grande delle costanti K di ϕ_1 e ϕ_2). Torniamo alla disuguaglianza:

$$\left| \oint_{\partial B_R} u \frac{\partial u}{\partial \mathbf{n}} \, dS \right| \leq \frac{H}{(1+R)^{n-1+\beta}} \oint_{\partial B_R} dS \, .$$

L'ultimo membro tende a 0 per $R \to +\infty$ in quanto l'ultimo integrale vale CR^{n-1} per qualche costante $C > 0$. Per cui:

$$\lim_{R \to +\infty} \int_{\Omega_R} ||\nabla u||^2 d^n x = 0 \, .$$

Il teorema della convergenza monotona, assicura che la funzione $x \mapsto ||\nabla u||^2$ sia integrabile su $\mathbb{R}^n \setminus \overline{\Omega} = \Omega_\infty$ con integrale nullo. Pertanto la funzione deve essere quasi ovunque nulla. Essendo $u \mapsto ||\nabla u||$ continua, deve essere $\nabla u = 0$ su $\mathbb{R}^n \setminus \overline{\Omega}$. Concludiamo che, su ogni componente connessa di Ω, u deve essere costante. Questo implica immediatamente la tesi dato che $u(x) \to 0$ per $||x|| \to +\infty$ nelle nostre ipotesi, e quindi l'eventuale costante che differenzia ϕ_1 da ϕ_2 è nulla. $\square$

Capitolo 3

Soluzioni fondamentali per l'equazione di Poisson in $\mathbb{R}^n$ e risultati ad esse legati.

Ci occuperemo ora di definire, e studiarne le proprietà , delle cosiddette *soluzioni fondamentali*, dell'equazione di Poisson. Tali strumenti matematici sono utili per vari motivi come vedremo. Esse permettono di ottenere altre importanti proprietà delle funzioni armoniche, ma fondamentalmente sono usate in una delle procedure classiche per determinare le soluzioni del problema di Dirichlet e Neumann.

3.1 Soluzioni fondamentali.

In questa sezione, una volta definite le soluzioni fondamentali, mostreremo come da esse si ricavino nuovi risultati sulle funzioni armoniche: i teoremi della media, un rafforzamento dell'enunciato del principio del massimo, la prova del fatto che le funzioni armoniche sono C^∞ ed addirittura analitiche, ed il risultato che stabilisce che se due funzioni armoniche coincidono su un aperto non vuoto A allora coincidono in ogni aperto connesso che contiene A sul quale sono entrambe definite (in particolare ogni funzione armonica nulla su un aperto non vuoto è nulla).

Definizione 3.1. Per $n = 2, 3, \ldots$ fissato, si definiscono **soluzioni fondamentali** su $\mathbb{R}^n$ dell'equazione di Poisson rispetto al punto $x_0 \in \mathbb{R}^n$ le funzioni:

$$\mathbb{R}^n \ni x \mapsto G_n(x, x_0) := \mathscr{G}_n(||x - x_0||) \,,$$

dove, per $r > 0$,

$$\mathscr{G}_n(r) := \begin{cases} \dfrac{1}{(2-n)\omega_n} r^{2-n} & \text{se } n > 2 \,, \\ \dfrac{1}{(2\pi)} \log r & \text{se } n = 2 \,, \end{cases} \tag{3.1}$$

in cui ω_n è la misura della superficie della sfera di raggio unitario in $\mathbb{R}^n$ pari a:

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)} \, .$$

$\diamondsuit$

Osservazioni 3.1.
(1) Γ è la nota *funzione gamma* di Eulero:

$$\Gamma(z) := \int_0^\infty e^{-t} t^{z-1} dt \, , \quad \text{per } z > 0,$$

che si prolunga analiticamente univocamente in una funzione olomorfa su tutto $\mathbb{C}$ con l'esclusione di singolarità (poli) nei punti $z = 0, -1, -2, \ldots$. Γ soddisfa, in particolare:

$$\Gamma(n) = (n-1)! \quad \text{e} \quad \Gamma(n+1/2) = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n} \sqrt{\pi} \quad \text{per } n = 1, 2, \ldots.$$

(2) Se $B_R \subset \mathbb{R}^n$ è una palla aperta di raggio R, $Vol(\partial B_R)$ e $Vol(B_R)$ denotano rispettivamente la misura della superficie di ∂B_R e del volume di B_R. Tali valori sono rispettivamente:

$$Vol(\partial B_R) = \omega_n R^{n-1} \quad \text{e} \quad Vol(B_R) = \frac{\omega_n R^n}{n} \, .$$

3.1.1 Proprietà elementari delle soluzioni fondamentali.

Vediamo ora le proprietà principali delle soluzioni fondamentali appena definite. Ricordiamo che una funzione $f : \mathbb{R}^n \to \mathbb{C}$ è detta **di classe** $\mathscr{L}^1_{loc}(\mathbb{R}^n)$ ovvero, equivalentemente, **localmente integrabile**, se è misurabile e soddisfa:

$$\int_B |\tilde{f}\,| d^n x < +\infty \, , \quad \text{per ogni aperto limitato } B \text{ di } \mathbb{R}^n,$$

dove l'integrale è quello di Lebesgue.
Nel seguito, riferendoci ad una funzione $f : A \to \mathbb{C}$ (o $\mathbb{R}$) con $A \subset \mathbb{R}^n$, diremo che essa è **di classe** $\mathscr{L}^1_{loc}(\mathbb{R}^n)$, ovvero equivalentemente **localmente integrabile**, se, estendendo tale funzione alla funzione nulla fuori da A, la funzione $\tilde{f}$ ottenuta in tal modo è localmente integrabile.

Teorema 3.1. *Per $n = 2, 3, \ldots$ e $x_0 \in \mathbb{R}^n$ fissati, le soluzioni fondamentali $G_n(x, x_0)$ soddisfano le seguenti proprietà .*

(a) *Le funzioni $\mathbb{R}^n \setminus \{x_0\} \ni x \mapsto G_n(x, x_0)$ e $\mathbb{R}^n \setminus \{x_0\} \ni x \mapsto G_n(x_0, x)$ sono di classe $\mathscr{L}^1_{loc}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n \setminus \{x_0\})$, anche se G_0 diverge per $x \to x_0$. Se $f \in C^0_0(\mathbb{R}^n)$ è pertanto ben definito:*

$$\int_{\mathbb{R}^n} G_n(x, y) f(y) d^n y \, .$$

(b) *Se* $x \neq x_0$ *allora:*

$$\Delta_x G_n(x, x_0) = 0 \,.$$

(c) *Se* $\rho \in C_0^2(\mathbb{R}^n)$ *allora:*

$$\int_{\mathbb{R}^n} G_n(x, y) \Delta_y \rho(y) d^n y = \rho(x) \,.$$

(d) *Se* $\rho \in C_0^2(\mathbb{R}^n)$ *allora:*

$$\Delta_x \int_{\mathbb{R}^n} G_n(x, y) \rho(y) d^n y = \rho(x) \,.$$

$\diamondsuit$

Dimostrazione. Cominciamo con il dimostrare (b). Dato che $G_n(x, x_0) := \mathscr{G}_n(||x - x_0||)$, è conveniente traslare l'origine delle cooordinate in x_0, che è tenuto fisso, introdurre un sistema di coordinate polari sferiche centrato in $x_0 = O$ e sfruttare il fatto che $\mathscr{G}_n$ dipende esplicitamente solo dalla variabile:

$$r := \sqrt{\sum_{i=1}^n (x^i)^2} = ||x - x_0|| \,. \tag{3.2}$$

Per computo diretto, calcolando il laplaciano in coordinate cartesiane ortonormali, si verifica subito che se $f = f(r)$ allora:

$$\Delta f(r) = \sum_{i,j=1}^n \delta^{ij} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} f(r) = \sum_{i,j=1}^n \delta^{ij} \frac{\partial}{\partial x^i} \left(\frac{\partial r}{\partial x^j} \frac{\partial}{\partial r} f(r) \right) = \sum_{i,j=1}^n \delta^{ij} \frac{\partial}{\partial x^i} \left(\frac{x^j}{r} \frac{\partial}{\partial r} f(r) \right)$$

$$= \sum_{i,j=1}^n \delta^{ij} \frac{x^j}{r} \frac{\partial}{\partial x^i} \frac{\partial f(r)}{\partial r} + \sum_{i,j=1}^n \delta^{ij} \frac{\partial}{\partial x^i} \left(\frac{x^j}{r} \right) \frac{\partial f(r)}{\partial r}$$

$$= \sum_{i,j=1}^n \delta^{ij} \frac{x^j}{r} \frac{\partial r}{\partial x^i} \frac{\partial}{\partial r} \frac{\partial f(r)}{\partial r} + \sum_{i,j=1}^n \delta^{ij} \frac{\partial}{\partial x^i} \left(\frac{x^j}{r} \right) \frac{\partial f(r)}{\partial r}$$

$$= \sum_{i,j=1}^n \delta^{ij} \frac{x^j x^i}{r^2} \frac{\partial}{\partial r} \frac{\partial f(r)}{\partial r} + \sum_{i,j=1}^n \delta^{ij} \frac{\delta^{ij} r - \frac{x^i x^j}{r}}{r^2} \frac{\partial f(r)}{\partial r} \,,$$

da cui, notando che

$$\sum_{i,j=1}^n \delta^{ij} \delta^{ij} = \sum_{i=1}^n \delta^{ii} = \sum_{i=1}^n 1 = n \,,$$

si trova alla fine a:

$$\Delta f(r) = \frac{\partial^2 f(r)}{\partial r^2} + \frac{n-1}{r} \frac{\partial f(r)}{\partial r}.$$

Ora, tramite il conto esplicito usando (3.1), si verifica la proprietà (b):

$$\Delta \mathscr{G}_n(r) = \frac{\partial^2 \mathscr{G}_n(r)}{\partial r^2} + \frac{n-1}{r}\frac{\partial \mathscr{G}_n(r)}{\partial r} = 0.$$

Da (3.1) risulta ovvio che $\mathscr{G}_n \in C^\infty(\mathbb{R}^n \setminus \{O\})$, cioè, rispristinando le coordinate cartesiane iniziali $\mathscr{G}_n \in C^\infty(\mathbb{R}^n \setminus \{x_0\})$ In particolare $\mathscr{G}_n$ è dunque misurabile. Osserviamo inoltre che, essendo l'elemento di volume in coordinare sferiche

$$d^n x = \sqrt{r^{2n-2}} dr d\Omega_n = r^{n-1} dr d\Omega_n \ ,$$

dove $d\Omega_n$ è l'elemento di volume sulla sfera unitaria $n-1$ dimensionale, si ha che la funzione $\mathscr{G}_n$ è assolutamente integrabile in ogni aperto limitato B la cui chiusura include l'origine, dato che la divergenza per $r \to 0$ di $\mathscr{G}_n$ è controbilanciata da un fattore infinitesimo per $r \to 0^+$ dovuto alla misura usata: si ha in totale un termine $r^{2-n} r^{n-1}$, ovvero $r \ln r$ se $n = 2$, da integrare rispetto alla misura dr. L'integrazione di $d\Omega_n$ produce invece ω_n finito. Su aperti limitati la cui chiusura non include l'origine, essendo $\mathscr{G}_n$ continua, la sua integrabilità è ovvia. In definitiva $\mathscr{G}_n \in \mathscr{L}^1_{loc}(\mathbb{R}^n)$. L'invarianza per traslazioni della misura di Lebesgue assicura infine l'indipendenza da punto x_0 e che quindi $G_n(\cdot, x_0) \in \mathscr{L}^1_{loc}(\mathbb{R}^n)$. Per simmetria nello scambio di x_0 e x in $G_n(x, x_0) = \mathscr{G}_n(||x - x_0||)$ si ha anche che $G_n(x_0, \cdot) \in \mathscr{L}^1_{loc}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n \setminus \{x_0\})$. In particolare sono ben definiti intgrali del tipo:

$$\int_{\mathbb{R}^n} G_n(x, y) f(y) d^n y \ ,$$

per $f \in C^0_0(\mathbb{R}^n)$ valendo:

$$\int_{\mathbb{R}^n} |G_n(x, y) f(y)| d^n y \leq \max_{\mathbb{R}^n} |f| \int_{\text{supp} f} |G_n(x, y)| d^n y < +\infty \ .$$

Abbiamo quindi provato anche la proprietà (a).
Dimostriamo ora che $(c) \Rightarrow (d)$.
Sia $\rho \in C^2_0(\mathbb{R}^n)$, se vale la (c), sfruttando l'invarianza per traslazioni della misura di Lebesgue:

$$\rho(x) = \int_{\mathbb{R}^n} \mathscr{G}_n(||x - y||) \Delta_y \rho(y) d^n y = \int_{\mathbb{R}^n} \mathscr{G}_n(||u||) \Delta_u \rho(x - u) d^n u$$

$$= \int_{\mathbb{R}^n} \mathscr{G}_n(||u||) \Delta_x \rho(x - u) d^n u = \Delta_x \int_{\mathbb{R}^n} \mathscr{G}_n(||u||) \rho(x - u) d^n u$$

$$= \Delta_x \int_{\mathbb{R}^n} G_n(x, y) \rho(y) d^n y \ .$$

La penultima uguaglianza segue dalla formula di derivazione sotto il segno di integrale nella variabile x basata sul teorema della convergenza dominata di Lebesgue (vedi la sezione B.2 in appendice). Ipostesi sufficienti per applicarla per le derivate nel punto x_0 sono che, per una palla

aperta $B_\epsilon(x_0)$ centrata nel punto x_0, la funzione $f(x,u) = \mathscr{G}_n(||u||)\rho(x-u)$ sia una funzione Lebesgue integrabile nella variabile $u \in \mathbb{R}^n$ per ogni $x \in B_\epsilon(x_0)$ e le derivate di tale funzione nella variabile x, fino all'ordine voluto (quello dell'operatore differenziale che si vuole scambiare con l'integrale), siano ciascuna rispettivamente maggiorata in valore assoluto, uniformemente in $x \in B_\epsilon(x_0)$, da una corrispondente funzione assolutamente integrabile dipendente dalla sola variabile d'integrazione $u \in \mathbb{R}^n$. Queste condizioni sono effettivamente verificate. Infatti al variare di $x \in B_\epsilon(x_0)$, i supporti delle funzioni $u \mapsto \rho(x-u)$ sono tutti contenuti in un compatto comune[1] K. Di conseguenza lo stesso accade, su K, per le derivate in x delle funzioni $u \mapsto \rho(x-u)$ (fino al secondo ordine). Definiamo una funzione C^∞ a supporto compatto $\mathbb{R}^n \ni u \mapsto g(u)$, che valga:

$$\max\left\{ \left|\frac{\partial \rho}{\partial x^k}\right| , \left|\frac{\partial^2 \rho}{\partial x^i \partial x^j}\right| \quad \middle| x \in \mathbb{R}^n \ , \ i,j,k = 0,1,2 \right\}$$

su K e si annulli rapidamente fuori da tale insieme. Per costruzione la funzione

$$\mathbb{R}^n \ni u \mapsto g(u)\mathscr{G}_n(||u||)$$

è Lebesgue integrabile e soddisfa, per $i,j = 0,1,2$:

$$\left|\frac{\partial f(u,x)}{\partial x^k}\right| , \left|\frac{\partial^2 f(u,x)}{\partial x^i \partial x^j}\right| \leq |g(u)\mathscr{G}_n(||u||)| \ , \quad (x,u) \in B_\epsilon(x_0) \times \mathbb{R}^n \ .$$

Questo giustifica lo scambio del simbolo di integrale con il laplaciano Δ_x (e con le derivate di ordine 1) eseguito sopra.
Dimostriamo infine la proprietà (c).
Ricordiamo che per ipotesi ρ ha supporto compatto, quindi ha supporto chiuso e limitato. Fissato $x \in \mathbb{R}^n$ consideriamo dunque una palla aperta $B_R(x)$, di raggio finito R e centrata in x, che include il supporto di ρ:

$$\int_{\mathbb{R}^n} G_n(x,y)\Delta_y \rho(y) d^n y = \int_{B_R(x)} G_n(x,y)\Delta_y \rho(y) d^n(y).$$

Gli integrali sono ben definiti visto che $\rho \in C^2(\overline{B_R(x)})$ è quindi limitata e pertanto, $y \mapsto G_n(x,y)\Delta_y\rho(y)$ è Lebesgue integrabile essendo G_n localmente integrabile. Siano $B_\epsilon(x)$ palle aperte di centro x con raggio $\epsilon > 0$ e $\epsilon < R$. Consideriamo la classe di funzioni, parametrizzate in $\epsilon > 0$, ottenute restringendo $y \mapsto G_n(x,y)\Delta_y\rho(y)$ agli insiemi $B_R(x)\backslash B_\epsilon(x)$ e definendole nulle fuori da tali insiemi. Per $\epsilon \to 0^+$, tali funzioni tendono puntualmente a $y \mapsto G_n(x,y)\Delta_y\rho(y)$ definita su tutta $B_R(x)$, inoltre sono maggiorate in valore assoluto dal valore assoluto di tale funzione che è integrabile per ipotesi. Applicando il teorema della convergenza dominata abbiamo

[1] Sia $B_R(0)$ una palla di raggio $R > 0$ centrata nell'origine e sufficientemente grande da includere il supporto, compatto per ipotesi, di $u \mapsto \rho(x_0 - u)$. Sia $B_S(0)$ una seconda palla, centrata nell'origine, di raggio $S > 0$ che includa $B_\epsilon(x_0)$. I supporti delle funzioni $u \mapsto \rho(x-u)$, per $x \in B_\epsilon(x_0) \subset B_S(0)$, sono sicuramente inclusi in $\cup_{u \in B_R(0)} B_S(u)$. Tale insieme è sicuramente contenuto nella palla compatta $K := \overline{B_{R+S}(0)}$ che, a maggior ragione, contiene tutti i supporti delle funzioni dette.

allora che:

$$\int_{B_R(x)} G_n(x,y)\Delta_y\rho(y)d^n(y) = \lim_{\epsilon\to 0^+}\int_{B_R(x)\setminus B_\epsilon(x)} G_n(x,y)\Delta_y\rho(y)d^n(y)\,.$$

Quindi, applicando la seconda identità di Green al risultato, troviamo:

$$\int_{B_R(x)} G_n(x,y)\Delta_y\rho(y)d^n y = \lim_{\epsilon\to 0^+}\Big[\int_{B_R(x)\setminus B_\epsilon(x)} \Delta_y G_n(x,y)\rho(y)d^n(y)+$$

$$+\oint_{+\partial(B_R(x)\setminus B_\epsilon(x))} G_n(x,y)\nabla_y\rho(y)\cdot\mathbf{n}dS - \oint_{+\partial(B_R(x)\setminus B_\epsilon(x))} \nabla_y G_n(x,y)\rho(y)\cdot\mathbf{n}dS\Big]$$

Ora $\Delta_y G_n(x,y) = \Delta_y\mathscr{G}_n(||x-y|) = \Delta_x G_n(x,y)$ e, dato che il dominio di integrazione in y è esterno a $B_\epsilon(x)$, varrà :

$$\Delta_x G_n(x,y) = 0 \quad \forall y\in B_R(x)\setminus B_\epsilon(x)\,.$$

Inoltre:

$$\rho\restriction_{\partial B_R(x)} = 0\,,\quad \nabla_y\rho\restriction_{\partial B_R(x)} = 0$$

in quanto $(\text{supp}\rho)\cap\partial B_R(x) = \emptyset$.
Dunque:

$$\int_{\mathbb{R}^n} G_n(x,y)\Delta_y\rho(y)d^n y =$$

$$= \lim_{\epsilon\to 0^+}\oint_{+\partial B_\epsilon(x)} \nabla_y G_n(x,y)\rho(y)\cdot\mathbf{n}'dS - \lim_{\epsilon\to 0^+}\oint_{+\partial B_\epsilon(x)} G_n(x,y)\nabla_y\rho(y)\cdot\mathbf{n}'dS$$

Il cambiamento di segno rispetto all'integrale precedente è dovuto al fatto che abbiamo cambiato il verso del versore normale a $\partial B_\epsilon(x)$: $\mathbf{n}$ indica il versore entrante, nell'ultimo integrale invece $\mathbf{n}' = -\mathbf{n}$ indica il versore uscente.
L'ultimo integrale soddisfa la seguente diseguaglianza:

$$\left|\oint_{+\partial B_\epsilon(x)} G_n(x,y)\nabla_y\rho(y)\cdot\mathbf{n}'dS\right| \le \sup_{\partial B_\epsilon(x)}||\nabla_y\rho||\oint_{+\partial B_\epsilon(x)}|G_n(x,y)|\,dS\,,$$

e quindi, tenendo conto che $G(x,y)$ è costante in y su $\partial B_\epsilon(x)$, mentre $\oint_{+\partial B_\epsilon(x)} dS = cost.\,\epsilon^{n-1}$,

$$\left|\oint_{+\partial B_\epsilon(x)} G_n(x,y)\nabla_y\rho(y)\cdot\mathbf{n}'dS\right| \le \sup_{\partial B_\epsilon(x)}||\nabla_y\rho||const.\epsilon^{n-1}\begin{cases}\frac{1}{\epsilon^{n-2}} & n>2\\ |\ln\epsilon| & n=2\end{cases}$$

che tende a 0 per $\epsilon\to 0^+$. Rimane quindi:

$$\int_{\mathbb{R}^n} G_n(x,y)\Delta_y\rho(y)d^n y = \lim_{\epsilon\to 0^+}\oint_{+\partial B_\epsilon(x)} \nabla_y G_n(x,y)\rho(y)\cdot\mathbf{n}'dS\,.$$

Con il solito sistema di coordinate polari sferiche centrato in x, si ha:

$$\mathbf{n}' \cdot \nabla_y G_n(x,y)|_{\partial B_\epsilon(x)} = \frac{\partial}{\partial r}\mathscr{G}_n(r)|_{\partial B_\epsilon(x)} = \begin{cases} \frac{1}{2\pi}\frac{1}{\epsilon} & n=2 \\ \frac{1}{\omega_n \epsilon^{n-1}} & n>2 \end{cases} = \frac{1}{Vol(\partial B_\epsilon(x))}\,, \tag{3.3}$$

dove abbiamo usato il fatto che, riferendosi alle coordinate polari centrate in x, vale $\mathbf{n}' = \mathbf{e}_r$ e, per funzioni della sola coordinata radiale, $\mathbf{e}_r \cdot \nabla f(r) = \frac{\partial}{\partial r} f(r)$. Quindi:

$$\int_{\mathbb{R}^n} G_n(x,y)\Delta_y \rho(y) d^n y = \lim_{\epsilon\to 0^+} \frac{1}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} \rho(y) dS$$

L'ultimo integrale può essere facilmente calcolato:

$$\lim_{\epsilon\to 0^+} \frac{1}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} \rho(y) dS(y)$$

$$= \lim_{\epsilon\to 0^+} \frac{1}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} (\rho(y)-\rho(x)) dS(y) \;+\; \rho(x) \lim_{\epsilon\to 0^+} \frac{1}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} dS(y)$$

$$= 0 + \rho(x)\,.$$

dove abbiamo usato il fatto che:

$$\left| \frac{1}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} (\rho(y)-\rho(x)) dS(y) \right| \leq \frac{\max_{\partial B_\epsilon(x)} |\rho(y)-\rho(x)|}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} dS(y) \leq \max_{\partial B_\epsilon(x)} |\rho(y)-\rho(x)|$$

In definitiva

$$\left| \frac{1}{Vol(\partial B_\epsilon(x))} \oint_{\partial B_\epsilon(x)} (\rho(y)-\rho(x)) dS(y) \right| \leq \max_{\partial B_\epsilon(x)} |\rho(y)-\rho(x)| \leq \max_{\overline{B_\epsilon(x)}} |\rho(y)-\rho(x)|$$

e l'ultimo termine tende a zero per $\epsilon \to 0^+$, dato che ρ è continua. Pertanto

$$\int_{\mathbb{R}^n} G_n(x,y)\Delta_y \rho(y) d^n y = \rho(x)\,.$$

□

3.2 Ulteriori proprietà delle funzioni armoniche in $\mathbb{R}^n$.

Nella dimostrazione del teorema 3.1, nell'espressione:

$$\int_{B_R(x)} G_n(x,y)\Delta_y \rho(y) d^n(y) = \lim_{\epsilon\to 0^+} \left[\int_{B_R(x)\setminus B_\epsilon(x)} \Delta_y G_n(x,y)\rho(y) d^n(y) + \right.$$

$$+ \oint_{+\partial(B_R(x)\setminus B_\epsilon(x))} G_n(x,y)\nabla_y\rho(y)\cdot\mathbf{n}dS - \oint_{+\partial(B_R(x)\setminus B_\epsilon(x))} \nabla_y G_n(x,y)\rho(y)\cdot\mathbf{n}dS\Bigg]$$

abbiamo trascurato gli integrali di superficie relativi a $\partial B_R(x)$, dato che la funzione ρ si annulla prima di arrivare a $\partial B_R(x)$. Tuttavia avremmo potuto considerare una palla $B_R(x)$ e più in generale un dominio Ω a chiusura compatta e con bordo regolare, sul quale ρ e le sue derivate non si annullano. Usando *essenzialmente la stessa dimostrazione* con Ω al posto di $B_R(x)$, ma senza trascurare gli integali di bordo su $\partial\Omega$, si arriva al seguente importante teorema:

Teorema 3.2. *Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto, $\overline{\Omega}$ compatto e $\partial\Omega$ una superficie regolare ed orientabile. Sia $\rho : \overline{\Omega} \mapsto \mathbb{R}$ una funzione di classe $C^2(\overline{\Omega})$. Per ogni $x \in \Omega$, vale l'identità :*

$$\rho(x) = \oint_{+\partial\Omega} (\nabla_y G_n(x,y))\rho(y)\cdot\mathbf{n}dS(y) - \oint_{+\partial\Omega} G_n(x,y)\nabla_y\rho(y)\cdot\mathbf{n}dS(y) + \int_\Omega G_n(x,y)\Delta_y\rho(y)d^n y. \quad (3.4)$$

A parità di ipotesi su Ω la stessa formula vale se, più debolmente, $\rho \in C^2(\Omega)\cap C^1(\overline{\Omega})$ con $\Delta\rho$ limitato su Ω. $\diamondsuit$

Dimostrazione. La prima affermazione si prova come per (c) del teorema 3.1 semplicemente rimpiazzando $B_R(x)$ con Ω e tenendo conto che il supporto di ρ ora può intersecare $\partial\Omega$ per cui rimangono dei contributi dovuti agli integrali di superficie trascurati nella dimostrazione di (c) del teorema 3.1. Se la funzione continua $\Delta\rho$ è limitata su Ω allora $\Omega \ni y \mapsto G_n(x,y)\Delta_y\rho(y)$ è comunque integrabile su Ω essendo G_n localmente integrabile e pertanto la dimostrazione si può ripetere similmente alla precedente dato che siamo nelle ipotesi di validità delle identità di Green. $\square$

Studiamo ora le conseguenze di questo risultato fondamentale.

3.2.1 Non esistenza di funzioni armoniche con supporto compatto.

La prima conseguenza del teorema 3.2, è la seguente proposizione che stabilisce che non esistono funzioni armoniche a supporto compatto.

Proposizione 3.1. *Se $A \subset \mathbb{R}^n$ è un aperto non vuoto non esistono funzioni armoniche a supporto compatto in A.* $\diamondsuit$

Dimostrazione. Supponiamo che $g : A \to \mathbb{R}$ sia armonica a supporto (rispetto a Ω) compatto K. Dato che la proprietà di compattezza non dipende dalla topologia indotta, K è compatto anche rispetto alla topologia di $\mathbb{R}^n$ ed è pertanto chiuso in $\mathbb{R}^n$ (e limitato). Di conseguenza K è strettamente incluso in A. Possiamo allora prolungare g su tutto $\mathbb{R}^n$ definendola come la funzione nulla fuori da A ed ottenendo una funzione $C^2(\mathbb{R}^n)$ ed armonica su tutto $\mathbb{R}^n$:

$$\Delta g(y) = 0 \quad \text{per ogni } y \in \mathbb{R}^n.$$

Applicando l'identità stabilita nel teorema 3.2 per Ω dato da una palla di raggio sufficientemente grande da includere strettamente K e x, in modo tale che g e ∇g si annullino su $\partial\Omega$, si ha:

$$g(x) = -\oint_{+\partial\Omega} G_n(x,y)\nabla_y g(y)\cdot \mathbf{n} dS(y) + \oint_{+\partial\Omega} (\nabla_y G_n(x,y))g(y)\cdot \mathbf{n} dS(y)$$

$$+\int_\Omega G_n(x,y)\Delta_y g(y) d^n y = 0\,,$$

per ogni $x \in \Omega$ (fuori da Ω la funzione è nulla per costruzione) ed in particolare $x \in A$. $\square$

3.2.2 Analiticità delle funzioni armoniche in $\mathbb{R}^n$.

Per eneunciare e dimostrare (parzialmente) il prossimo teorema, ricordiamo che una funzione di più variabili complesse $f : \Omega_{\mathbb{C}} \ni z \mapsto \mathbb{C}$, con $\Omega_{\mathbb{C}} \subset \mathbb{C}^n$ aperto non vuoto, è detta **funzione analitica di più variabili complesse** (equivalentemente **funzione olomorfa di più variabili complesse**) [6] se f è analitica (ossia olomorfa) in ciscuna variabile z^k di $z = (z^1, \ldots, z^n)$ separatamente, quando le altre sono fissate arbitrariamente. Risulta che se $f : \Omega_{\mathbb{C}} \to \mathbb{C}$ è analitica, allora, per ogni $z_0 \in \Omega_{\mathbb{C}}$, essa è sviluppabile in serie di Taylor centrata in z_0 in un intorno aperto di z_0 incluso in $\Omega_{\mathbb{C}}$. Ovviamente ci stiamo riferendo allo sviluppo di Taylor in più variabili:

$$f(z) = \sum_{N=0}^{+\infty} \sum_{\alpha_1+\cdots+\alpha_n=N} \frac{1}{\alpha_1!\cdots\alpha_n!} \left.\frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\partial z^{1\alpha_1}\cdots\partial z^{n\alpha_n}}\right|_{z_0} (z^1 - z_0^1)^{\alpha_1}\cdots(z^n - z_0^n)^{\alpha_n}\,. \tag{3.5}$$

Si osservi che, posto $z = x + iy$, *se $f(x)$ assume valori reali per ogni $x \in \mathbb{R}^n \cap \Omega_{\mathbb{C}}$*, lo sviluppo sopra scritto sviluppato attorno a $z_0 = x_0$ con $x_0 \in \mathbb{R}^n$ e calcolato per $x \in \mathbb{R}$ si riduce al solito sviluppo di Taylor reale:

$$f(x) = \sum_{N=0}^{+\infty} \sum_{\alpha_1+\cdots+\alpha_n=N} \frac{1}{\alpha_1!\cdots\alpha_n!} \left.\frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\partial x^{1\alpha_1}\cdots\partial x^{n\alpha_n}}\right|_{z_0} (x^1 - x_0^1)^{\alpha_1}\cdots(x^n - x_0^n)^{\alpha_n}\,, \tag{3.6}$$

dove abbiamo calcolato tutte le derivate parziali eseguendo i limiti sull'asse reale. In questo caso, dato che $\Omega := \mathbb{R}^n \cap \Omega_{\mathbb{C}}$ è aperto, la restrizione di f a tale dominio definisce una funzione analitica reale.

Teorema 3.3. **(Analiticità delle funzioni armoniche).** *Se $\varphi : \Omega \to \mathbb{R}$, con $\Omega \subset \mathbb{R}^n$ aperto non vuoto, è armonica su Ω, allora $\varphi \in C^\infty(\Omega)$ e più fortemente φ è analitica reale su Ω.*$\diamondsuit$

Traccia della dimostrazione. Sia $x_0 \in \Omega$, consideriamo una palla B aperta di raggio finito e con $\overline{B} \subset \Omega$ centrata in x_0 e applichiamo la formula (3.4) su B, tenendo conto che $\Delta\varphi = 0$:

$$\varphi(x) = -\oint_{+\partial B} G_n(x,s)\nabla_s\varphi(s)\cdot \mathbf{n} dS(s) + \oint_{+\partial B} (\nabla_s G_n(x,s))\varphi(s)\cdot \mathbf{n} dS(s)\,,$$

dove, in particolare, $x \in B'$, con B' palla aperta centrata in x_0 di raggio strettamente inferiore a quello di B. I due integrandi sono funzioni continue nelle variabili (x, s) e quindi limitate su $(x, s) \in \overline{B'} \times \partial B$ (che è un insieme compatto in $\mathbb{R}^n \times \mathbb{R}^n$). La stessa cosa accade alle derivate in x, di ogni ordine, degli integrandi. Notare le divergenze di $G_n(x, s)$ (e delle sue derivate) appaiono solamente quando $x = s$, cosa impossibile se $(x, s) \in \overline{B'} \times \partial B$. Per ogni derivata di ogni fissato ordine α (incluse derivate miste) nelle componenti di x, $D_x^{(\alpha)}$ esiste una costante M_α per cui $|D_x^{(\alpha)} G_n(x, s)\mathbf{n} \cdot \nabla_s \varphi(s)| \leq M_\alpha$ per ogni $s \in \partial B$ ed uniformemente in $x \in B'$. Dato che, per ogni α, ogni funzione costante $\partial B \ni s \mapsto M_\alpha \geq 0$ è sicuramente assolutamente integrabile su ∂B (che ha misura finita!), concludiamo (vedi la sezione B.2 in appendice) che possiamo passare la derivata $D_x^{(\alpha)}$ fuori dal segno di integrale in

$$-\oint_{+\partial B} D_x^{(\alpha)} G_n(x, s) \nabla_s \varphi(s) \cdot \mathbf{n} dS(s)$$

derivando per $x = x_0$. Lo stesso ragionamento si può fare per il secondo integrale nella decomposizione integrale di φ:

$$\varphi(x) = -\oint_{+\partial B} G_n(x, s) \nabla_s \varphi(s) \cdot \mathbf{n} dS(s) + \oint_{+\partial B} (\nabla_s G_n(x, s)) \varphi(s) \cdot \mathbf{n} dS(s) \,.$$

In altre parole, possiamo dunque derivare φ ad ogni ordine x, scaricando le derivate sulle funzioni $G_n(x, s)$ e $\mathbf{n} \cdot \nabla_s G_n(x, s)$, sotto il segno di integrale. In tal modo abbiamo verificato che φ è infinitamente differenziabile in x_0 e quindi, dato che ciò vale per ogni punto $x_0 \in \Omega$, abbiamo provato che $\varphi \in C^\infty(\Omega)$. Diamo ora una dimostrazione del fatto che φ può essere estesa ad una funzione analitica di più variabile complesse $z \mapsto \varphi'(z)$ con $z \in \Omega_{\mathbb{C}} \subset \mathbb{C}^n$ un aperto che include $\Omega \subset \mathbb{R}^n$. Per prima cosa notiamo che, dalla loro definizione, le funzioni $B' \ni x \mapsto G_n(x, s) = \mathscr{G}_n(||x - s||)$, con $s \in \partial B$ fissato, si estendono a funzioni analitiche complesse di più variabili (non daremo una dimostrazione rigorosa di tale fatto): $z \mapsto G_n'(z, s)$ con $z = x + iy$, dove $y \in B'$ e $x \in B'$ e possiamo anche prendere $(x, y) \in \overline{B'} \times \overline{B'}$, restringendo il raggio originale di B'. Definiamo pertanto per $(x, y) \in \overline{B'} \times \overline{B'}$:

$$\varphi'(z) := -\oint_{+\partial B} G_n(z, s) \nabla_s \varphi(s) \cdot \mathbf{n} dS(s) + \oint_{+\partial B} (\nabla_s G_n(z, s)) \varphi(s) \cdot \mathbf{n} dS(s) \,.$$

Il secondo membro è ben definito e può essere derivato in x e y passando le derivate sotto il segno di integrale, dato che la funzione G_n è infinitamente differenziabile sul compatto $\overline{B'} \times \overline{B'} \times \partial B$. Dato che per ogni fissato s, la funzione $z \mapsto G_n'(z, s)$ soddisfa in ogni variabile z^k le condizioni di Cauchy-Riemann, soddisferà le stesse condizioni la funzione φ': è sufficiente passare le derivate sotto il segno di integrale. In definitva, la funzione di variabile complessa $B' + iB' \ni z \mapsto \varphi'(z)$ è definita su un aperto, ammette derivate continue (essendo di classe C^∞) nelle variabilei x^k e y^k (dove $z = (z^1, \cdots, z^n) = (x^1 + iy^1, \cdots, x^n + iy^n)$) e soddisfa le condizioni di Cauchy-Riemann in ogni variabile z^k. Tenuto conto di quanto detto in (2) in osservazioni 2.5, φ' è una funzione olomorfa in più variabili complesse e nell'intorno di ogni punto nel suo dominio varrà lo sviluppo (3.5) con φ' al posto di f. Dato che per valori reali $z_0 = x_0 \in \mathbb{R}^n$, φ' si riduce alla funzione a

valori reali φ, concludiamo che nell'intorno di ogni $x_0 \in \Omega$ vale lo sviluppo (3.6) con φ al posto di f. In altre parole φ è una funzione analitica reale sul dominio aperto Ω. $\square$

Osservazioni 3.2.
(1) Il teorema appena dimostrato ci dice quindi che una funzione armonica può essere estesa ad una funzione analitica complessa su un opportuno dominio in $\mathbb{C}^n$.
(2) Le funzioni analitiche reale godono della proprietà dell'unicità della continuazione analitica:
Proposizione 3.2. *Se due funzioni analitiche reali φ e ψ sono entranbe definite sull'aperto non vuoto e connesso $\Omega \subset \mathbb{R}^n$ e coincidino sull'aperto non vuoto $A \subset \Omega$, allora coincidono su tutto Ω.* $\diamondsuit$
Dimostrazione. Sia $U \subset \Omega$ l'insieme dato dall'unione di tutti gli aperti inclusi in Ω su cui $\varphi \equiv \psi$. Ovviamente U è non vuoto (dato che $A \subset \Omega$ è aperto e su di esso le funzioni coincidono), aperto e $U \subset \Omega$. Supponiamo per assurdo che $U \neq \Omega$. Sia allora $q \in \Omega \setminus U$ e $p \in U$. Ci sarà un cammino continuo $\gamma : [0,1] \to \Omega$ con $\gamma(0) = p$ e $\gamma(1) = q$. Se $s = \sup\{t \in [0,1] \mid \gamma(t) \in U\}$, sarà $p' := \gamma(s) \in \partial U \cap \Omega$ per costruzione. Dimostrimolo. Se $p' = \gamma(s) \in Int(U)$, c'è una palla aperta B centrata in $\gamma(s)$ tutta contenuta in U. La contro immagine di B secondo la funzione continua γ individua un aperto $I \ni s$ con la proprietà che $\gamma(I) \subset U$, ma allora ci sarebbe un intorno destro di s la cui immagine secondo γ è inclusa in U e questo è impossibile per definizione di s. Si ottiene un analogo assurdo assumendo che $p' = \gamma(s) \in Ext(U)$: si trova un intorno sinistro di s la cui immagine secondo γ è esterna a U e questo è impossibile per definizione di s. L'unica possibilità è quindi $p' \in \partial U$, ma anche $p' = \gamma(s) \in \Omega$ per definizione di γ. In particolare ha dunque senso valutare φ e ψ in p' ed in un intorno di tale punto. Dato che φ e ψ sono continue con tutte le loro derivate di ogni ordine e che, essendo $p' \in \partial U$, esiste una successione di punti $\{p_n\}_{n\in\mathbb{N}} \subset U$ che converge a p', tutte le derivate di φ e ψ in p' possono essere calcolate prendendo i limiti di tali derivate verso $p' \in \Omega$, ma calcolandole sulla successione $\{p_n\}_{n\in\mathbb{N}} \subset U$, prima di fare i limiti. Dato che $\varphi \restriction_U \equiv \psi \restriction_U$, le derivate di φ e ψ in U coincidono ad ogni ordine (si osservi che se A, e quindi U, non fosse aperto, $\varphi \restriction_U \equiv \psi \restriction_U$ non implicherebbe necessariamente che le derivate delle due funzioni coincidono su tale insieme). In particolare tutte le derivate ad ogni ordine di φ e ψ coincideranno quando valutate in p' con la procedura di limite indicata sopra. Dato che in p' le funzioni sono analitiche, gli sviluppi di Taylor centrati in p' delle due funzioni coicidono ed abbiamo pertanto che $\varphi \equiv \psi$ in un intorno aperto $J_{p'} \subset \Omega$ di p'. Per costruzione l'insieme aperto $U \cup J_{p'} \subset \Omega$ include A, su $U \cup J_{p'}$ le due funzioni coincidono e $U \cup J_{p'}$ è più grande di U contenendo p' che non appartiene a U, dato che U è aperto e $p' \in \partial U$. Questo è impossibile per definizione di U ed abbiamo in questo modo raggiunto un assurdo. Concludiamo che deve essere $U = \Omega$. $\square$

L'osservazione (2) ha la seguente implicazione immediata in virtù del teorema 3.3.

Proposizione 3.3. *Se due funzioni armoniche reali φ e ψ sono entrambe definite sull'aperto non vuoto e connesso $\Omega \subset \mathbb{R}^n$ e coincidono su un aperto non vuoto $A \subset \Omega$, allora esse coincidono su tutto Ω.* $\diamondsuit$

3.2.3 Teorema della media e principio del massimo in forma forte.

Un'altra importante conseguenza del teorema 3.2 è il cosiddetto *teorema della media.*

Teorema 3.4. **(Teorema della media)**. *Sia $\varphi : \Omega \mapsto \mathbb{R}$, con $\Omega \subset \mathbb{R}^n$ aperto non vuoto, una funzione armonica. Allora, per ogni $x \in \Omega$ vale l'identità , detta* **formula della media superficiale***:*

$$\varphi(x) = \frac{1}{Vol(\partial B_R(x))} \oint_{\partial B_R(x)} \varphi(y) dS(y), \tag{3.7}$$

dove $B_R(x)$ è una palla aperta centrata in x di raggio finito $R > 0$ con $\overline{B_R(x)} \subset \Omega$, arbitrariamente scelta. Similmente vale anche la **formula della media volumetrica***:*

$$\varphi(x) = \frac{1}{Vol(B_R(x))} \int_{B_R(x)} \varphi(y) d^n x(y) \,. \tag{3.8}$$

$\diamondsuit$

Dimostrazione. Sia R il raggio della palla $B_R(x)$. Utilizziamo un sistema di coordinate polari sferiche centrate in x. Dalla (3.4) e tenendo conto del fatto, già notato, che $\nabla_y G_n(x,y) = \frac{\partial \mathscr{G}_n(r)}{\partial r}$, abbiamo:

$$\begin{aligned}\varphi(x) &= \oint_{\partial B_R(x)} \varphi(y) \frac{\partial \mathscr{G}_n(r)}{\partial r} dS(y) - \oint_{+\partial B_R(x)} \mathscr{G}_n(r) \nabla\varphi(y) \cdot \mathbf{n}\, dS(y) \\ &= \oint_{\partial B_R(x)} \varphi(y) \frac{\partial \mathscr{G}_n(r)}{\partial r} dS(y) - \mathscr{G}_n(R) \oint_{+\partial B_R(x)} \nabla\varphi(y) \cdot \mathbf{n}\, dS(y) \,.\end{aligned}$$

L'ultimo integrale è nullo perchè φ è armonica (teorema 2.8), mentre il primo, usando (3.3) si può scrivere:

$$\oint_{\partial B_R(x)} \varphi(y) \frac{1}{Vol(\partial B_R(x))} dS(y) = \frac{1}{Vol(\partial B_R(x))} \oint_{\partial B_R(x)} \varphi(y) dS(y) \,.$$

Passiamo alla seconda formula della media. Applichiamo la prima formula della media alla classe di palle $B_r(x)$ di raggio r, con $0 < r \leq R$, ed usiamo un sistema di coordinate polari sferiche di centro x, coordinata radiale r e coordinate angolari ω. Avremo allora che, da (3.7) vale:

$$Vol(\partial B_r(x))\varphi(x) = \oint_{\partial B_r(x)} \varphi(r,\omega) dS(r,\omega)$$

e quindi, integrando in dr da $r = 0$ a $r = R$:

$$\left(\int_0^R Vol(\partial B_r(x)) dr \right) \varphi(x) = \int_0^R \left(\oint_{\partial B_r(x)} \varphi(r,\omega) dS(r,\omega) \right) dr \,.$$

Il primo integrale produce proprio il volume della palla $B_R(x)$ moltiplicato per la costante $\varphi(x)$, mentre il secondo produce l'integrale di volume su tale palla della funzione φ, decomposto in due integrazioni in coordinate polari. In definitiva otteniamo la (3.8):

$$Vol(B_R(x))\varphi(x) = \int_{B_R(x)} \varphi(y) d^n x(y) \, .$$

□

Una conseguenza diretta del teorema della media è un rafforzamento del principio del massimo che dimostriamo in due parti.

Lemma 3.1. **(Principio del massimo forte su una palla.)** *Sia $B_R(x_0)$ una palla aperta in $\mathbb{R}^n$, di raggio $R > 0$ finito centrata in x_0 e $\varphi : \overline{B_R(x_0)} \to \Omega$ una funzione armonica in $B_R(x_0)$ e continua in $\overline{B_R(x_0)}$. Se vale uno dei seguenti fatti:*

$$\varphi(x_0) = \max_{\overline{B_R(x_0)}} \varphi \, ,$$

oppure

$$\varphi(x_0) = \min_{\overline{B_R(x_0)}} \varphi \, ,$$

oppure

$$|\varphi(x_0)| = \max_{\overline{B_R(x_0)}} |\varphi| \, ,$$

allora la funzione φ è costante su $\overline{B_R(x_0)}$. $\diamondsuit$

Dimostrazione: È sufficiente dimostrare la tesi per il caso $\varphi(x_0) = \max_{x \in B_R(x_0)} \varphi$, in quanto se vale la seconda ipotesi, cambiando segno alla funzione φ, si ricade nella prima situazione. Se vale la terza ipotesi allora deve valer la prima oppure la seconda (Dato che non è del tutto evidente dimostriamo quest'ultimo fatto. Ci sono tre casi da considerare. (i) $\varphi \geq 0$ in $\overline{B_R(x_0)}$; in questo caso $|\varphi(x)| = \varphi(x)$ e dunque, $|\varphi(x_0)| = \max_{\overline{B_R(x_0)}} |\varphi|$ equivale a dire $\varphi(x_0) = \max_{\overline{B_R(x_0)}} \varphi$. (ii) $\varphi \leq 0$ in $\overline{B_R(x_0)}$; in questo caso $|\varphi(x)| = -\varphi(x)$ e dunque, $|\varphi(x_0)| = \max_{\overline{B_R(x_0)}} |\varphi|$ equivale a dire $\varphi(x_0) = \min_{\overline{B_R(x_0)}} \varphi$. (iii) φ assume sia valori positivi che valori negativi in $\overline{B_R(x_0)}$, in questo caso il valore massimo raggiunto da φ è positivo e quello minimo è negativo. Nella situazione considerata, il valore massimo raggiunto da $|\varphi|$ deve necessariamente corrispondere al massimo valore di φ oppure al minimo valore di φ cambiato di segno, se non corrispondesse a nessuno dei due non potrebbe essere il massimo per $|\varphi|$. Allora abbiamo due sottocasi. (a) Il valore massimo che la funzione $|\varphi|$ assume è il massimo di φ; in questo caso, dato che tale valore di φ è positivo, la condizione $|\varphi(x_0)| = \max_{\overline{B_R(x_0)}} |\varphi|$ equivale a dire $\varphi(x_0) = \pm \max_{\overline{B_R(x_0)}} \varphi$. Se risultasse $\varphi(x_0) = -\max_{\overline{B_R(x_0)}} \varphi$, significherebbe $\varphi(x_0) = \min_{\overline{B_R(x_0)}} \varphi$ altrimenti ci sarebbero valori più piccoli di $\varphi(x_0)$ raggiunti da φ e quindi ci sarebbero valori più grandi di $|\varphi(x_0)|$ raggiunti da $|\varphi|$, cosa impossibile per ipotesi. (b)

Il valore massimo che la funzione $|\varphi|$ assume è, cambiato di segno, il minimo di φ; in questo caso la condizione $|\varphi(x_0)| = \max_{\overline{B_R(x_0)}} |\varphi|$ equivale a dire $\varphi(x_0) = \pm \min_{\overline{B_R(x_0)}} \varphi$. Se risultasse $\varphi(x_0) = -\min_{\overline{B_R(x_0)}} \varphi$, significherebbe $\varphi(x_0) = \max_{\overline{B_R(x_0)}} \varphi$ altrimenti ci sarebbero valori più grandi di $\varphi(x_0)$ raggiunti da φ e quindi ci sarebbero valori più grandi di $|\varphi(x_0)|$ raggiunti da $|\varphi|$, cosa impossibile per ipotesi.)
Sia dunque $\varphi(x_0) = \max_{\overline{B_R(x_0)}} \varphi$, dimostriamo che $\varphi(x) = \varphi(x_0)$ per ogni $x \in B_R(x_0)$, per continuità ciò varrà anche per $x \in \partial B_R(x_0)$.
Supponiamo *per assurdo* che esista $x_1 \in B_R(x_0)$ con $\varphi(x_1) \neq \varphi(x_0)$, allora, per le ipotesi fatte, deve essere $\varphi(x_1) < \varphi(x_0)$. Per la continuità di φ, scegliendo $0 < \epsilon < |\varphi(x_0) - \varphi(x_1)|$, esisterà una palla aperta $B_\delta(x_1) \subset B_R(x_0)$ centrata in x_1 e di raggio $\delta > 0$ tale che $|\varphi(x) - \varphi(x_1)| < \epsilon$ se $x \in B_\delta(x_1)$. Di conseguenza, se $x \in B_\delta(x_1)$, vale anche: $\varphi(x) < \varphi(x_0)$. In particolare varrà $\varphi(x) < \varphi(x_0)$, se $x \in K := \overline{B_{\delta/2}(x_1)}$ dato che $\overline{B_{\delta/2}(x_1)} \subset B_\delta(x_1)$. Applichiamo il teorema della media volumetrica:

$$Vol(B_R(x_0))\, \varphi(x_0) = \int_{B_R(x_0)} \varphi d^n x = \int_{B_R(x_0)\setminus K} \varphi d^n x + \int_K \varphi d^n x \,. \tag{3.9}$$

K è compatto per costruzione e quindi esiste $\max_K \varphi$, con $\max_K \varphi < \varphi(x_0)$ per costruzione di K. Quindi

$$\int_K \varphi d^n x \leq \left(\max_K \varphi\right) \int_K d^n x < \varphi(x_0) \int_K d^n x \,.$$

Dato che vale anche, essendo $\varphi(x_0)$ il valore massimo di φ,

$$\int_{B_R(x_0)\setminus K} \varphi d^n x \leq \varphi(x_0) \int_{B_R(x_0)\setminus K} d^n x \,,$$

da (3.9) segue subito che:

$$Vol(B_R(x_0))\, \varphi(x_0) < \varphi(x_0) \int_{B_R(x_0)\setminus K} d^n x + \varphi(x_0) \int_K d^n x = \varphi(x_0) \left(\int_{B_R(x_0)\setminus K} d^n x + \int_K d^n x \right) \,,$$

ossia

$$Vol(B_R(x_0))\, \varphi(x_0) < Vol(B_R(x_0))\, \varphi(x_0) \,,$$

che è assurdo e, pertanto, il punto x_1 con $\varphi(x_1) < \varphi(x_0)$ non può esistere in $B_R(x_0)$. $\square$

Il risultato appena dimostrato ci consente di estendere il principio del massimo, nel senso forte appena visto, a funzioni armoniche su regioni Ω diverse da un palla.

Teorema 3.5. **(Principio del massimo forte.)** *Sia Ω aperto, connesso a chiusura compatta in $\mathbb{R}^n$ e sia $\varphi \in C^2(\Omega) \cap C^0(\overline{\Omega})$ armonica su Ω. Se vale una delle seguenti condizioni per qualche $x_0 \in \Omega$:*

(i) $\varphi(x_0) = \max_{\overline{\Omega}} \varphi$, oppure

(ii) $\varphi(x_0) = \min_{\overline{\Omega}} \varphi$, *oppure*

(iii) $|\varphi(x_0)| = \max_{\overline{\Omega}} |\varphi|$,

allora φ *è costante e vale ovunque* $\varphi(x_0)$ *su* $\overline{\Omega}$. $\diamondsuit$

Dimostrazione. Dimostriamo la tesi nel caso in cui sia verificata la prima ipotesi. Se vale l'ipotesi (ii), allora possiamo ricadere in (i) cambiando segno a φ, mentre se vale (ii), allora si ricade in (i) o in (ii) con lo stesso ragionamento del lemma precedente. Notiamo infine che è sufficiente mostrare la validità della tesi in Ω, perché da questa segue, per la continuità di φ, la tesi in $\overline{\Omega}$.
Dato che $\Omega \subset \mathbb{R}^n$ aperto e connesso, allora è connesso per archi continui. Sia dunque $x_1 \in \Omega$ e $\gamma : [a, b] \to \Omega$ continua con $\gamma(a) = x_0$, $\gamma(b) = x_1$. Mostriamo che $\varphi(x_1) = \varphi(x_0)$. Ciò prova la tesi per l'arbitrarietà di $x_1 \in \Omega$. Assumendo la validità di (i), per ogni palla di raggio finito $B_R(x_0)$ centrata in x_0 e con $\overline{B_R(x_0)} \subset \Omega$ deve anche evidentemente essere:

$$\varphi(x_0) = \max_{\overline{\Omega}} \varphi = \max_{\overline{B_R(x_0)}} \varphi \, .$$

Applicando il teorema precedente concludiamo che $\varphi(x) = \varphi(x_0)$ per ogni $x \in B_R(x_0)$. La controimmagine dell'aperto $B_R(x_0)$ secondo la funzione continua γ deve essere un aperto (relativiamente alla topologia di $[a, b]$ indotta da $\mathbb{R}$) che include il punto $\gamma(a) = x_0$. Di conseguenza, ci sarà un intervallo $[a, \epsilon)$, con $a < \epsilon \leq b$ e con $\gamma(t) \in B_R(x_0)$ se $t \in [a, \epsilon)$, per cui $\varphi(\gamma(t)) = \varphi(x_0)$ se $t \in [a, \epsilon)$. L'insieme

$$S = \{s \in (a, b] \mid \varphi(\gamma(u)) = \varphi(x_0) \text{ per } u \in [a, s)\}$$

è non vuoto (per quanto appena dimostrato $\epsilon \in S$) ed è limitato superiormente da $b < \infty$, quindi esiste $L = \sup S \leq b$.
Supponiamo per assurdo che $L < b$, in tal caso $\varphi(\gamma(t)) = \varphi(x_0)$ per $t \in [a, L)$ e per continiutà $\varphi(\gamma(L)) = \varphi(x_0)$. Esisterà dunque una palla centrata in $\gamma(L)$ e di raggio $\rho > 0$, che indichiamo con $B_\rho(\gamma(L)) \subset \Omega$, tale che $\overline{B_\rho(\gamma(L))} \subset \Omega$. Come prima:

$$\varphi(\gamma(L)) = \max_{\overline{\Omega}} \varphi = \max_{\overline{B_\rho(\gamma(L))}} \varphi \, .$$

e quindi $\varphi(x) = \varphi(x_0)$ costantemente su $B_\rho(\gamma(L))$. Come prima, la controimmagine della palla aperta $B_\rho(\gamma(L))$ secondo la funzione continua γ è un aperto che continene L per costruzione. Su tale aperto $\varphi(\gamma(t)) = \varphi(x_0)$. In particolare dovrà dunque valere $\varphi(\gamma(t)) = \varphi(x_0)$ in un intorno *destro* di L, per cui L non può essere il sup di S e siamo quindi giunti ad un assurdo. Dovrà dunque essere $L = b$ e pertanto $\varphi(\gamma(t)) = \varphi(x_0)$ se $t \in [a, b)$. Per continuità: $\varphi(x_1) = \varphi(\gamma(b)) = \varphi(x_0)$. $\square$

Osservazioni 3.3. Abbiamo dato una dimostrazione del principio del massimo forte senza usare il fatto che le funzioni armoniche sono analitiche reali (risultato che non abbiamo dimostrato completamente) e quindi soddisfano la proposizione 3.3. La dimostrazione del pricipio del

massimo forte segue infatti facilmente dalla proposizione 3.3 osservando che, nelle ipotesi del teorema 3.5, la funzione φ è sicuramente costante in una palla aperta $B_R(x_0) \subset \Omega$ come provato nella parte iniziale della dimostrazione data sopra. In virtù del fatto che Ω è aperto e connesso con $B_R(x_0) \subset \Omega$, dalla proposizione 3.3 segue allora che, su tutto Ω, φ deve coincidere con la funzione ψ che vale costantemente $\varphi(x_0)$ (ed è quindi armonica) dato che $\varphi \restriction_{B_R(x_0)} \equiv \psi \restriction_{B_R(x_0)}$.

3.2.4 Teorema di Liouville per le funzioni armoniche in $\mathbb{R}^n$.

Come ultimo risultato, che segue dal teorema della media e dal fatto che le funzioni armoniche sono di classe C^∞ (per la dimostrazione è sufficiente C^3), si ha il seguente teorema.

Teorema 3.6. **(Teorema di Liouville per funzioni armoniche.)** *Ogni funzione armonica su tutto $\mathbb{R}^n$ limitata superiormente oppure inferiormente su $\mathbb{R}^n$ è costante.* $\diamondsuit$

Dimostrazione. Se $\varphi : \mathbb{R}^n \to \mathbb{R}$ è armonica e limitata inferiormente, sia $\phi(x) := \varphi(x) - \inf_{\mathbb{R}^n} \varphi$ per $x \in \mathbb{R}^n$. ϕ è armonica e non negativa. Se deriviamo ϕ rispetto a x^k avremo ancora una funzione armonica per costruzione (si tenga conto del fatto che $\phi \in C^\infty$ essendo essa armonica). Possiamo usare la formula della media volumetrica su una palla $B_r(x_0)$ centrata in x_0 di raggio finito $r > 0$ arbitrario:

$$\frac{\partial \phi}{\partial x^k}(x_0) = \frac{1}{VolB_r(x_0)} \int_{B_r(x_0)} \frac{\partial \phi}{\partial x^k} d^n x = \frac{1}{VolB_r(x_0)} \int_{B_r(x_0)} \nabla \cdot (\phi \mathbf{e}_k) d^n x = \frac{1}{VolB_r(x_0)} \oint_{+\partial B_r(x_0)} \phi n^k dS,$$

dove abbiamo usato il teorema di Gauss, $\mathbf{e}_i$ è l'i-esimo versore della base canonica di $\mathbb{R}^n$ e n^k è la k-esima componente di $\mathbf{n}$ uscente da ∂B. Quindi:

$$\left| \frac{\partial \phi}{\partial x^k}(x_0) \right| \leq \left| \frac{1}{VolB_r(x_0)} \oint_{+\partial B_r(x_0)} \phi n^k dS \right| \leq \frac{1}{VolB_r(x_0)} \oint_{+\partial B_r(x_0)} \left| \phi n^k \right| dS \, .$$

Dato che $\phi \geq 0$ e $|n^k| \leq 1$, abbiamo infine la stima:

$$\left| \frac{\partial \phi}{\partial x^k}(x_0) \right| \leq \frac{1}{VolB_r(x_0)} \oint_{+\partial B_r(x_0)} \phi dS \, .$$

Applicando al secondo membro il teorema della media superficiale abbiamo anche che:

$$\left| \frac{\partial \phi}{\partial x^k}(x_0) \right| \leq \frac{Vol\partial B_r(x_0)}{VolB_r(x_0)} \phi(x_0) \, .$$

Ossia, dato che il rapporto a secondo membro vale n/r:

$$\left| \frac{\partial \phi}{\partial x^k}(x_0) \right| \leq \frac{n}{r} \phi(x_0) \, .$$

Dato che $r > 0$ può essere scelto arbitrariamente grande (il dominio di φ è tutto $\mathbb{R}^n$) otteniamo che deve necessariamente essere:

$$\frac{\partial \phi}{\partial x^k}(x_0) = 0 \quad \text{per ogni } x_0 \in \mathbb{R}^n \ .$$

Dato che $\mathbb{R}^n$ è connesso e che quanto scritto sopra vale per ogni derivata, concludiamo che ϕ, e dunque φ, deve essere costante su $\mathbb{R}^n$. Se φ è limitata superiormente, si può ripetere la stessa dimostrazione usando $-\varphi$. $\square$

La dimostrazione contiene un risultato che è utile menzionare separatamente in un lemma.

Lemma 3.2. *Sia $\Omega \subset \mathbb{R}^n$ aperto non vuoto e $\varphi : \Omega \to \mathbb{R}$ una funzione armonica. Se $x \in \Omega$ e $B_r(x) \subset \Omega$ è una palla aperta centrata in x di raggio $r > 0$ tale che $\overline{B_r(x)} \subset \Omega$, allora vale per la derivata k-esima, $k = 1, 2, \ldots, n$:*

$$\left| \frac{\partial \varphi}{\partial x^k}(x) \right| \leq \frac{n}{r}(\varphi(x) - \min_{B_r(x)} \varphi) \ ,$$

e

$$\left| \frac{\partial \varphi}{\partial x^k}(x) \right| \leq \frac{n}{r}(\max_{B_r(x)} \varphi - \varphi(x)) \ .$$

$\diamondsuit$

Capitolo 4

Funzioni di Green e costruzione di soluzioni del problema di Dirichlet.

In questo capitolo mostreremo come costruire soluzioni del problema di Dirichlet per domini Ω specifici, usando soluzioni opportune, simili alle soluzioni fondamentali (ma che tengono conto del dominio Ω) dette *funzioni di Green* e *nuclei di Poisson*. Questo approccio è sicuramente interessante, in particolare per il significato fisico (carica immagine) e per gli sviluppi che ha avuto nella fisica matematica in riferimento a problemi di natura completamente diversa. Tuttavia, da un punto di vista puramente matematico, si tratta di un metodo che non si riesce a generalizzare nel caso di domini abbastanza arbitrari e equazioni differenziali di tipo ellittico, ma non a coefficienti costanti. Le tecniche moderne di costruzione della soluzione di problemi con dati al contorno per equazioni ellittiche, sono basate su altri approcci in cui la soluzione viene cercata e costruita in spazi funzionali *deboli* (soluzioni nel senso delle *distribuzioni* in *spazi di Sobolev*) e poi viene provata la regolarità di tali soluzioni (sfruttando proprietà di regolarità specifiche degli operatori ellittici [2] in Appendice A abbiamo dato qualche ulteriore dettaglio.

4.1 Ancora sul problema di Dirichlet.

Consideriamo il problema di Dirichlet in una regione $\Omega \subset \mathbb{R}^n$ aperto a chiusura $\overline{\Omega}$ compatta per $\varphi \in C^2(\Omega) \cap C^0(\bar{\Omega})$:

$$\begin{cases} \Delta\varphi = f, & f \in C^0(\Omega) \text{ funzione assegnata,} \\ \varphi|_{\partial\Omega} = \psi & \psi \in C^0(\partial\Omega) \text{ funzione assegnata.} \end{cases} \tag{4.1}$$

Se $\partial\Omega$ è una superficie chiusa regolare orientabile allora abbiamo a disposizione un'identità , la (3.4), che ci permette di esprimere la soluzione φ, *se esiste*, in funzione del valori che φ e il suo gradiente $\nabla\varphi$ assumono su $\partial\Omega$. Infatti vale:

$$\varphi(x) = \int_\Omega G_n(x,y)\Delta_y\varphi d^n y - \oint_{+\partial\Omega} G_n(x,y)\nabla_y\varphi(y)\cdot \mathbf{n} dS(y) +$$

$$+\oint_{+\partial\Omega} \nabla_y G_n(x,y)\cdot \mathbf{n}\varphi(y)dS(y), \tag{4.2}$$

da cui:

$$\varphi(x) = \int_\Omega G_n(x,y)f(y)d^n y - \oint_{+\partial\Omega} G_n(x,y)\nabla_y\varphi(y)\cdot \mathbf{n}dS(y)+$$
$$+\oint_{+\partial\Omega} \nabla_y G_n(x,y)\cdot \mathbf{n}\psi(y)dS(y).$$

Questa formula non può tuttavia essere utilizzata per *determinare* la soluzione al problema assegnato perchè per utilizzarla per conoscere φ in ogni punto di Ω è necessario conoscere *anche* $\nabla\varphi$ su $\partial\Omega$, che non è noto dalle condizioni al contorno. Per un problema con condizioni al contorno di Neumann si avrebbe lo stesso problema in quanto non sarebbero noti i valori che φ assume su $\partial\Omega$.

Osservazioni 4.1. Se, oltre ai valori di φ su $\partial\Omega$, fossero assegnati anche valori di $\nabla\varphi$ su $\partial\Omega$ (ad esempio $\nabla\varphi\cdot\mathbf{n}|_{\partial\Omega} = \psi_1$), e tentassimo di usare l'espressione (4.2) per scrivere una possibile soluzione,

$$\varphi(x) = \int_\Omega G_n(x,y)f(y)d^n y - \oint_{+\partial\Omega} G_n(x,y)\psi_1(y)dS(y)$$
$$+\oint_{+\partial\Omega} \nabla_y G_n(x,y)\cdot \mathbf{n}\psi(y)dS(y) \tag{4.3}$$

in generale avremmo che la funzione φ così calcolata *non* risolverebbe il problema:

$$\begin{cases} \Delta\varphi = f \text{ funzione assegnata,} \\ \varphi|_{\partial\Omega} = \psi \text{ funzione assegnata,} \\ \nabla\varphi\cdot\mathbf{n}|_{\partial\Omega} = \psi_1 \text{ funzione assegnata.} \end{cases} \tag{4.4}$$

Infatti, dai teoremi di unicità per il problema di Dirichlet e Neumann, sappiamo che in generale questo problema *non può ammettere soluzione* assegnando $\varphi|_{\partial\Omega}$ e $\nabla\varphi|_{\partial\Omega}\cdot\mathbf{n}$ contemporaneamente e arbitrariamente.

4.1.1 Funzioni di Green e nuclei di Poisson.

Per usare (4.2) al fine di scrivere la soluzione del problema di Dirichlet in funzione dei dati al bordo, possiamo cercare di modificare G_n in modo da far sparire in (4.2) il termine contenente il gradiente di φ e cercare di usare i soli dati di Dirichlet. In questo modo potremmo riuscire a produrre un candidato della soluzione del problema. Con un problema di Neumann si può procedere similmente. Abbiamo la seguente proposizione che ci porta verso la direzione voluta.

Proposizione 4.1. *Sia $\Omega\subset\mathbb{R}^n$ insieme aperto non vuoto a chiusura compatta e con bordo dato da una superficie regolare orientabile. Sia $\{v_\Omega(x,\cdot)\}_{x\in\Omega}\subset C^2(\overline{\Omega})$ una classe di soluzioni dell'equazione:*

$$\Delta_y v_\Omega(x,y) = 0\,, \quad (x,y)\in\Omega\times\Omega$$

per le quali valga anche:

$$v_\Omega(x,y) + G_n(x,y) = 0, \quad \text{se } x \in \Omega \text{ e } y \in \partial\Omega\,.$$

Allora valgono i fatti seguenti.
(a) *Per ogni funzione $\varphi \in C^0(\overline{\Omega}) \cap C^2(\Omega)$ con $\nabla\varphi$ e $\Delta\varphi$ limitate su Ω, vale:*

$$\varphi(x) = \int_\Omega G_\Omega(x,y)\Delta\varphi(y)d^n y + \oint_{\partial\Omega} N_\Omega(x,y)\varphi(y)dS(y)\,, \quad \text{per ogni } x \in \Omega, \tag{4.5}$$

dove $G_\Omega(x,y) := G_n(x,y) + v_\Omega(x,y)$ e $N_\Omega(x,y) := \nabla_y G_\Omega(x,y)\cdot\mathbf{n}|_{y\in\partial\Omega}$ (con $\mathbf{n}$ versore uscente da $\partial\Omega$) sono detti rispettivamente la **funzione di Green** *ed il* **nucleo di Poisson** *per il problema di Dirichlet del laplaciano su Ω.*
(b) *Per ogni fissato $x \in \Omega$, $\Omega \ni y \mapsto G_\Omega(x,y)$ si estende ad una funzione $C^2(\overline{\Omega}\setminus\{x\})$, armonica su $\overline{\Omega}\setminus\{x\}$ e nulla su $\partial\Omega$.*
(c) *Se $(x,y) \in \Omega\times\Omega$, allora*

$$v_\Omega(x,y) = v_\Omega(y,x) \quad \text{e, se } x \neq y, \quad G_\Omega(y,x) = G_\Omega(x,y)\,.$$

In particolare, quindi per ogni fissato $y \in \Omega$, $\Omega \ni x \mapsto G_\Omega(x,y)$ si estende univocamente per continuità ad una funzione $C^2(\overline{\Omega}\setminus\{y\})$, armonica su $\overline{\Omega}\setminus\{y\}$ e nulla su $\partial\Omega$. $\diamondsuit$

Dimostrazione. (a) Fissiamo $x \in \Omega$ e quindi scgliamo un nuovo dominio Ω_ϵ, con le stesse caratteristiche di Ω, ma tale che $\overline{\Omega_\epsilon}$ è strettamente contenuto in Ω. Assumiamo che, per $\epsilon \to 0$, Ω_ϵ invada Ω. Dato che, se $x \in \Omega$ è fissato, vale $\Delta_y v_\Omega(x,y) = 0$ quando $y \in \overline{\Omega}$, pertanto:

$$\int_{\Omega_\epsilon} \varphi(y)\Delta_y v_\Omega(x,y)d^n y = 0$$

per ogni funzione $\varphi : \Omega \to \mathbb{R}$. Se ammettiamo che $\varphi \in C^2(\Omega)$, dalla seconda identità di Green abbiamo come conseguenza:

$$0 = \int_{\Omega_\epsilon} v_\Omega(x,y)\Delta_y\varphi(y)d^n y + \oint_{+\partial\Omega_\epsilon} \varphi(y)\nabla_y v_\Omega(x,y)\cdot\mathbf{n}dS(y) - \oint_{+\partial\Omega_\epsilon} v_\Omega(x,y)\nabla_y\varphi(y)\cdot\mathbf{n}dS(y)\,.$$

Sommando membro a membro con (4.2) calcolata su Ω' e definendo $G_\Omega(x,y) = G_n(x,y) + v_\Omega(x,y)$, si ha:

$$\varphi(x) = \int_{\Omega_\epsilon} G_\Omega(x,y)\Delta_y\varphi(y)d^n y + \oint_{+\partial\Omega_\epsilon} N_\Omega(x,y)\varphi(y)dS(y) - \oint_{+\partial\Omega_\epsilon} G_\Omega(x,y)\nabla_y\varphi(y)\cdot\mathbf{n}dS(y)$$

dove $N_\Omega(x,y) := \nabla_y G_\Omega(x,y)\cdot\mathbf{n}|_{y\in\partial\Omega}$. Se ora assumiamo che $\varphi \in C^0(\overline{\Omega})\cap C^2(\Omega)$ e che $\nabla\varphi$ e $\Delta\varphi$ siano limitati su Ω, possiamo calcolare il limite per $\epsilon \to 0$ usando il teorema della convergenza dominata, ottenendo

$$\varphi(x) = \int_\Omega G_\Omega(x,y)\Delta_y\varphi(y)d^n y + \oint_{+\partial\Omega} N_\Omega(x,y)\varphi(y)dS(y)\,,$$

dove abbiamo trascurato il cotributo dovuto all'ultimo integrale di bordo che vale zero nelle nostre ipotesi in quanto $G_\Omega(x,y)=0$ se $y\in\partial\Omega$.
(b) La dimostrazione è ovvia per definizione di G_Ω, dalle proprietà note di v_Ω e G_n.
(c) Fissiamo $x_1,x_2\in\Omega$ con $x_1\neq x_2$ e consideriamo il volume, con ovvie notazioni, $\Omega_\epsilon=\Omega\setminus(B_\epsilon(x_1)\cup B_\epsilon(x_2))$. Su tale volume le funzioni: $y\mapsto G_\Omega(x_1,y)$ e $y\mapsto G_\Omega(x_2,y)$ sono regolari, armoniche e si annullano sul bordo esterno di Ω_ϵ, dato da $\partial\Omega$. In conseguenza dell'armonicità :

$$0=\int_{\Omega_\epsilon}(G_\Omega(x_1,y)\Delta_y G_\Omega(x_2,y)-G_\Omega(x_2,y)\Delta_y G_\Omega(x_1,y))\,d^n y\,.$$

Applicando la seconda identità di Green al secondo membro scritto sopra e tenendo conto che il contributo dovuto all'integrale di superficie su $\partial\Omega$ si annulla dato che $G_\Omega(x_1,y)=G_\Omega(x_1,y)=0$ se $y\in\partial\Omega$, otteniamo che:

$$\oint_{+\partial B_\epsilon(x_1)}G_\Omega(x_1,y)\mathbf{n}\cdot\nabla_y G_\Omega(x_2,y)\,dS(y)-\oint_{+\partial B_\epsilon(x_1)}G_\Omega(x_2,y)\mathbf{n}\cdot\nabla_y G_\Omega(x_1,y)\,dS(y)$$

$$+\oint_{+\partial B_\epsilon(x_2)}G_\Omega(x_1,y)\mathbf{n}\cdot\nabla_y G_\Omega(x_2,y)\,dS(y)-\oint_{+\partial B_\epsilon(x_2)}G_\Omega(x_2,y)\mathbf{n}\cdot\nabla_y G_\Omega(x_1,y))\,dS(y)=0\,.$$

Si osservi che la divergenza che si ha nel primo integrale per $\epsilon\to 0^+$ è unicamente dovuta alla divergenza di $G_n(x_1,y)$ per $y\to x_1$. Fissando coordinate polati centrate in x_1, tale divergenza è di ordine ϵ^{2-n} se $n>2$ oppure $\ln\epsilon$ se $n=2$. Tuttavia l'area della superficie $\partial B_\epsilon(x_1)$ tende a zero con rapidità ϵ^{n-1} oppure ϵ rispettivamente. Da ciò si conclude che il primo integrale tende a 0 per $\epsilon\to 0^+$. Lo stesso discorso vale per il quarto integrale. Nel secondo integrale, la divergenza è invece dovuta a $\mathbf{n}\cdot\nabla_y G_n(x_1,y)$, $y\to x_1$. Usando un sistema di coordinate polari centrate in x_1 si vede che $\mathbf{n}\cdot\nabla_y G_n(x_1,y)=1/vol(\partial B_\epsilon(x_1))$ come già osservato nella dimostrazione di (c) del teorema 3.1. Tale divergenza si compensa esattamente con il limite a 0 dell'area della superficie $\partial_\epsilon(x_1)$ quando $\epsilon\to 0^+$, lo stesso discorso vale per il secondo integrale. (Si osservi che il contributo dovuto ai termini $v_\Omega(x,y)$ è sempre nullo nel limite per $\epsilon\to 0^+$ dato che tali funzioni sono regolari nel dominio considerato e vengono integrate su domini di misura che tende a zero.) Ragionando nella dimostrazione di (c) del teorema 3.1 si verifica facilmente che, prendendo il limite per $\epsilon\to 0^+$ si ottiene un valore finale dato dai seguenti contributi :

$$0-G_\Omega(x_2,x_1)+G_\Omega(x_1,x_2)+0=0\,.$$

Da cui: $G_\Omega(x_2,x_1)=G_\Omega(x_1,x_2)$. Dato che vale anche $G_n(x_2,x_1)=G_n(x_1,x_2)$ per definizione, concludiamo che $v_\Omega(x_2,x_1)=v_\Omega(x_1,x_2)$ quando $x_1\neq x_2$. Questa identità per v_Ω vale banalmente anche nel caso $x_1=x_2$, dato che $v_\Omega(x,x)$ è definita per ogni $x\in\Omega$.
Si osservi infine che, per definizione, per ogni $x\in\Omega$, $y\mapsto G_\Omega(x,y)$ è una funzione $C^2(\overline{\Omega}\setminus\{x\})$, armonica e nulla su $\partial\Omega$. Pertanto per ogni $y\in\Omega$, $\Omega\ni x\mapsto G_\Omega(x,y)=G_\Omega(y,x)$ si estende univocamente per continuità ad una funzione $C^2(\overline{\Omega}\setminus\{y\})$, armonica e nulla su $\partial\Omega$. $\square$

L'espressione (4.5) fornisce φ in termini delle sole quantità assegnate in un problema di Dirichlet su Ω. Tale formula può essere usata per determinare la soluzione del problema di Dirichlet.

Sussiste infatti il seguente teorema.

Teorema 4.1. *Sia $\Omega \subset \mathbb{R}^n$ insieme aperto, non vuoto, a chiusura compatta e con bordo dato da una superficie regolare orientabile. Sia G_Ω una funzione di Green per il laplaciano su Ω con nucleo di Poisson N_Ω. Se valgono i seguenti fatti:*

(i) la funzione $v_\Omega := G_\Omega - G_n$ è di classe $C^3(\overline{\Omega} \times \overline{\Omega} \setminus \Delta)$, dove abbiamo definito $\Delta := \{(x,x) \mid x \in \partial\Omega\}$, e

(ii) vale l'identità :

$$\lim_{x \to x_0} \oint_{\partial\Omega} N_\Omega(x,y)\psi(y)dS(y) = \psi(x_0)\,, \quad \textit{per ogni } \psi \in C^0(\partial\Omega) \textit{ e ogni } x_0 \in \partial\Omega, \tag{4.6}$$

allora valgono i due seguenti fatti.

(a) *la funzione:*

$$\varphi(x) := \int_\Omega G_\Omega(x,y)f(y)d^n y + \oint_{\partial\Omega} N_\Omega(x,y)\psi(y)dS(y)\,, \quad \textit{per } x \in \Omega \tag{4.7}$$

estesa per continuità su $\overline{\Omega}$ è soluzione del problema di Dirichlet

$$\begin{cases} \Delta\varphi = f\,, \\ \varphi|_{\partial\Omega} = \psi\,, \end{cases} \tag{4.8}$$

per ogni scelta delle funzioni assegnate $f \in C^2_0(\Omega)$ e $\psi \in C^0(\partial\Omega)$.

(b) *La funzione di Green G_Ω per il laplaciano su Ω è l'unica soddisfacente (i) e (ii), ed è quindi unico il nucleo di Poisson, a meno di ridefinizione nei punti dell'insieme Δ.* $\diamondsuit$

Dimostrazione. (a) Fissato un qualsiasi $x' \in \Omega$ si consideri una palla aperta $B(x') \subset \Omega$ con $\overline{B(x')} \subset \Omega$. La funzione $\overline{B(x')} \times \mathrm{supp} f \ni (x,y) \mapsto v(x,y)f(y)$ è limitata e lo sono tutte le sue derivate in x di ogni ordine (essendo tale funzione con le sue derivate continue su un compatto), dunque esisterà una funzione costante $g \geq$ definita sull'insieme, di misura finita per ipotesi, Ω, e tale funzione maggiora i valori assoluti delle derivate in x fino al secondo ordine di $u(x,y)f(y)$ per tutti i valori di $y \in \Omega$ uniformemente in $x \in B(x')$. Possiamo allora derivare sotto il segno di integrale in $x \in B(x')$ due volte ottenendo, da (b) del lemma 4.1

$$\Delta_x \int_\Omega v(x,y)f(y)d^n y = \int_\Omega \Delta_x v(x,y)f(y)d^n y = 0\,.$$

Similmente, dato che f ha supporto compatto e pertanto

$$\int_\Omega G_n(x,y)f(y)d^n y = \int_{\mathbb{R}^n} G_n(x,y)f(y)d^n y$$

Per (d) del teorema 3.1 abbiamo che:

$$\Delta_x \int_\Omega G_n(x,y)f(y)d^n y = f(x)\,.$$

Mettendo tutto insieme abbiamo ottenuto che:

$$\Delta_x \int_\Omega G_\Omega(x,y) f(y) d^n y = f(x) \, .$$

La funzione $\overline{B(x')} \times \partial\Omega \ni (x,y) \mapsto N_\Omega(x,y) f(y)$ è derivabile in x e le sue derivate sono limitate, essendo funzioni continue su un compatto: per costruzione $x \neq y$ lavorando nell'insieme detto. Come prima possiamo derivare due volte sotto il segno di integrale ottenendo

$$\Delta_x \oint_{\partial\Omega} N_\Omega(x,y)\psi(y) dS(y) = \oint_{\partial\Omega} \Delta_x N_\Omega(x,y)\psi(y) dS(y) = 0 \, ,$$

dato che possiamo applicare il teorema di Schwartz, essendo v di classe C^3 e G_n di classe C^∞ (per argomenti non coincidenti):

$$\Delta_x \mathbf{n} \cdot \nabla G_\Omega(x,y) = \mathbf{n} \cdot \nabla \Delta_x G_\Omega(x,y) = 0$$

in quanto $G_\Omega(x,y)$ è , per costruzione, armonica in x se $x \neq y$. In definitiva

$$\Delta_x \left(\int_\Omega G_\Omega(x,y) f(y) d^n y + \oint_{\partial\Omega} N_\Omega(x,y)\psi(y) dS(y) \right) = f(x) \, .$$

La funzione φ è quindi in $C^2(\Omega)$ per costruzione (anzi è in $C^4(\Omega)$).
Se $B(x_0)$ è una palla aperta centrata in $x_0 \in \partial\Omega$ tale che $\partial B(x_0) \cap \operatorname{supp} f = \emptyset$, la funzione $(x,y) \mapsto G_\Omega(x,y) f(y)$ per $(x,y) \in (\overline{B(x_0)} \cap \overline{\Omega}) \times \Omega$ è continua con supporto compatto incluso in $(\overline{B(x_0)} \cap \overline{\Omega}) \times \operatorname{supp} f$ (le singolarità di G_Ω per $x = y$ non hanno effetto visto che tali punti sono fuori dal supporto). Sia $K \geq 0$ una costante che maggiora $(x,y) \mapsto G_\Omega(x,y) f(y)$ sul dominio detto. La funzione $\Omega \ni y \mapsto K$ è integrabile dato che Ω ha chiusura compatta. Possiamo allora applicare il teorema della convergenza dominata e concludere che

$$\lim_{x \mapsto x_0} \int_\Omega G_\Omega(x,y) f(y) d^n y = \int_\Omega \lim_{x \mapsto x_0} G_\Omega(x,y) f(y) d^n y = 0 \, ,$$

dove abbiamo usato il fatto che $\lim_{x \mapsto x_0} G_\Omega(x,y) = G_\Omega(x_0,y) = G_\Omega(y,x_0) = 0$ quando $x_0 \in \partial\Omega$ e $y \in \Omega$. Se infine vale anche la condizione 4.6, allora φ definita in (4.7) ed estesa su $\partial\Omega$ come $\varphi(x) := \psi(x)$ è continua su $\overline{\Omega}$ e, banalmente, soddisfa le condizioni al bordo. Mostriamone la continuità in ogni punto $x_0 \in \overline{\Omega} = \Omega \cup \partial\Omega$. Se $x_0 \in \Omega$ la funzione φ è continua in x_0 per costruzione, essendo di classe C^2, e non c'è nulla da provare. Se invece $x_0 \in \partial\Omega$ si procede come segue. Per ogni $\epsilon > 0$ possiamo trovare un intorno aperto A_ϵ di x_0 tale che, se $x \in A_\epsilon \cap \Omega$ allora $|\varphi(x) - \varphi(x_0)| < \epsilon$ come conseguenza di (4.6). Dato che ψ è a sua volta continua in x_0, per ogni $\epsilon > 0$ possiamo trovare un intorno aperto A'_ϵ di x_0 tale che, se $x \in A'_\epsilon \cap \partial\Omega$ allora $|\psi(x) - \psi(x_0)| < \epsilon$. In definitiva, tenendo conto del fatto che $\varphi = \psi$ su $\partial\Omega$, se $\epsilon > 0$, esiste un intorno di $x_0 \in \partial\Omega$, $B_\epsilon := A_\epsilon \cap A'_\epsilon$, tale che se $x \in B_\epsilon \cap \overline{\Omega} =\in B_\epsilon \cap (\Omega \cup \partial\Omega)$, allora $|\varphi(x) - \varphi(x_0)| < \epsilon$.
(b) Ovviamente è sufficiente dimostrare l'unicità della funzione di Green quando è valutata per

$x, y \in \Omega$ (con $x \neq y$), i rimanenti punti del bordo vengono inclusi nella dimostrazione per continuità . Sia G'_Ω un'altra funzione di Green soddisfacente (4.6). Per ogni problema di Dirichlet (4.8), l'unica (per il teorema 2.3) soluzione φ si potrà anche scrivere come:

$$\varphi(x) := \int_\Omega G'_\Omega(x,y) f(y) d^n y + \oint_{\partial\Omega} N'_\Omega(x,y)\psi(y) dS(y) \, .$$

Se scegliamo ψ identicamente nulla su $\partial\Omega$, per differenza con (4.7) avremmo che, per ogni $x \in \Omega$ e per ogni $f \in C_0^2(\Omega)$:

$$0 = \int_\Omega [G'_\Omega(x,y) - G_\Omega(x,y)] f(y) d^n y \, .$$

Supponiamo per assurdo che per fissati $x, y \in \Omega$ con $x \neq y$, $G'_\Omega(x,y) - G_\Omega(x,y) = l \neq 0$. Assumiamo senza perdere generalità $l > 0$. In una palla aperta di raggio finito B centrata in y e con x fissato, $G'_\Omega(x,y) - G_\Omega(x,y)$ si manterrebbe in $[l-\epsilon, l+\epsilon]$ con $l - \epsilon > 0$. Potremmo allora trovare una funzione $f \in C_0^2(\Omega)$ il cui supporto è contenuto in B che abbia integrale strettamente positivo pari a k. In modo tale che $\int_\Omega [G'_\Omega(x,y) - G_\Omega(x,y)] f(y) d^n y \geq k(l-\epsilon) > 0$ che è impossibile. Quindi $G'_\Omega(x,y) = G_\Omega(x,y)$ su $\Omega \times \Omega$ per $x \neq y$. $\square$

Osservazioni 4.2.
(1) Nelle ipotesi di validità del teorema precedente sappiamo rispovere su Ω dato *ogni* problema di Dirichlet con dati $\psi \in C^0(\partial\Omega)$ e $f \in C_0^2(\Omega)$. La conoscenza della funzione di Green non permette di risolvere un problema di Dirichlet, ma *ogni* problema di Dirichlet (se valgono tutte le ipotesi richieste).
(2) La validità della condizione (4.6) è abbastanza generale. Daremo una condizione sufficiente affincé essa valga nella prossima proposizione.
(3) Cambiamenti della definizione della funzione G_Ω sull'insieme Δ non alterano, evidentemente, i risultati in (a). Si osservi che anche la relazione (4.6) è indipendendte dal valore assunto da G_Ω su Δ.

Proposizione 4.2. *Nelle ipotesi del teorema 4.1, la condizione (4.6) è verificata se, per ogni $x_0 \in \partial\Omega$, esiste una successione di palle aperte $\{B_{r_n}(x_0)\}_{n\in\mathbb{N}}$, centrate in x_0 e di raggio $r_n > 0$ e con $r_n \to 0$ per $n \to +\infty$, tali che valga:*

$$\sup_{x \in B_{r_n}(x_0)\cap\Omega} \oint_{\partial\Omega \setminus B_{r_n}(x_0)} N_\Omega(x,y) dS(y) \to 0 \, , \quad \textit{per } n \to +\infty.$$

$\diamondsuit$

Dimostrazione. La dimostrazione usa il seguente lemma:
Lemma 4.1. *Il nucleo di Poisson $N_\Omega(x,y)$ soddisfa $N_\Omega(x,y) \geq 0$.* $\diamondsuit$

Dimostrazione del lemma. Infatti, se $x \in \Omega$ è fissato e $B_r(x)$ è una palla aperta di raggio r centrata in x con $\overline{B_r(x)} \subset \Omega$, la funzione $\Omega \setminus \overline{B_r(x)} \ni y \mapsto G_\Omega(x,y) = G_n(x,y) + v(x,y)$

è sicuramente negativa su $\partial B_r(x)$ scegliendo r sufficientemente piccolo, visto che $G_n(x,y)$ diverge a $-\infty$ quando $y \to x$, mentre $v(x,y)$ rimane limitata nell'intorno di x. Per costruzione $G_\Omega(x,y) = 0$ se $y \in \partial\Omega$. Dato che il massimo di $\overline{\Omega \setminus \overline{B_r(x)}} \ni y \mapsto G_\Omega(x,y)$ è assunto sul bordo del dominio (essendo la funzione armonica nell'interno del dominio e continua sulla chiusura), concludiamo che tale massimo è sicuramente 0 e che $G_n(x,y) < 0$ nell'interno del dominio per il principio del massimo forte. Segue facilmente che $-\mathbf{n} \cdot \nabla_y G_n(x,y) \leq 0$ quando $y \in \partial\Omega$, dove $-\mathbf{n}$ punta verso l'interno: se ciò non fosse, dato che $G_n(x,y) = 0$ su $y \in \partial\Omega$, troveremmo una curva C^1, $y = y(t)$, che entra in Ω partendo da $y(0) \in \partial\Omega$, lungo la quale $t \mapsto G_n(x, y(t))$ cresce raggiungendo valori positivi. $\square$

Proseguiamo la dimostrazione della proposizione. Per prima cosa notiamo che, se $x \in \Omega$:

$$\oint_{\partial\Omega} N_\Omega(x,y)\, dS(y) = \oint_{\partial\Omega} \mathbf{n} \cdot \nabla_y G_n(x,y)\, dS(y) + \oint_{\partial\Omega} \mathbf{n} \cdot \nabla_y v(x,y)\, dS(y)\,.$$

Consideriamo i due integrali a secondo membro. Il secondo integrale vale 0, dato che $y \mapsto v(x,y)$ è armonica su $\overline{\Omega}$ e vale il teorema 2.8. Il primo integrale vale invece 1, applicando la formula (3.4) alla funzione g che vale costantemente 1 su $\mathbb{R}^n$. In definitiva:

$$\oint_{\partial\Omega} N_\Omega(x,y)\, dS(y) = 1\,, \quad \text{per ogni } x \in \Omega.$$

Di conseguenza:

$$\oint_{\partial\Omega} N_\Omega(x,y)\psi(y) dS(y) = \psi(x_0) \oint_{\partial\Omega} N_\Omega(x,y) dS(y) + \oint_{\partial\Omega} N_\Omega(x,y)(\psi(y) - \psi(0)) dS(y)\,.$$

In altre parole:

$$\oint_{\partial\Omega} N_\Omega(x,y)\psi(y) dS(y) = \psi(x_0) + \oint_{\partial\Omega} N_\Omega(x,y)(\psi(y) - \psi(0)) dS(y)\,.$$

Per concludere dimostriamo che l'integrale a secondo membro tende a 0 se $\Omega \ni x \to x_0 \in \partial\Omega$. Infatti:

$$\oint_{\partial\Omega} N_\Omega(x,y)(\psi(y) - \psi(0)) dS(y) = \oint_{\partial\Omega \setminus B_{r_n}(x_0)} N_\Omega(x,y)(\psi(y) - \psi(0)) dS(y)$$

$$+ \oint_{\partial\Omega \cap B_{r_n}(x_0)} N_\Omega(x,y)(\psi(y) - \psi(0)) dS(y)\,.$$

Il valore assoluto del primo integrale a secondo membro, scegliendo $x \in B_{r_n}$, è sicuramente maggiorato da:

$$\max_{y \in \partial\Omega} |\psi(y) - \psi(0)| \sup_{x \in B_{r_n}(x_0) \cap \Omega} \oint_{\partial\Omega \setminus B_{r_n}(x_0)} |N_\Omega(x,y)| dS(y)\,.$$

Pertanto, nelle nostre ipotesi,

$$\oint_{\partial\Omega\setminus B_{r_n}(x_0)} N_\Omega(x,y)(\psi(y)-\psi(0))dS(y) \to 0\,, \quad \text{se } x \to x_0.$$

Il valore assoluto del secondo integrale, scegliendo $x \in B_{r_n}$, è invece maggiorato da:

$$\oint_{\partial\Omega\cap B_{r_n}(x_0)} |N_\Omega(x,y)|dS(y) \max_{y\in B_{rn}(x_0)} |\psi(y)-\psi(0)| \le \oint_{\partial\Omega} |N_\Omega(x,y)|dS(y) \max_{y\in B_{rn}(x_0)} |\psi(y)-\psi(0)|$$

$$= \oint_{\partial\Omega} N_\Omega(x,y)dS(y) \max_{y\in B_{rn}(x_0)} |\psi(y)-\psi(0)| = 1 \max_{y\in B_{rn}(x_0)} |\psi(y)-\psi(0)| \to 0\,, \quad \text{se } x \to x_0,$$

dove abbiamo usato il fatto che $N_\Omega(x,y) \ge 0$ per il lemma 4.1 e la continuità di ψ in x_0. $\square$

4.2 Funzioni di Green per domini particolari.

Non ci occuperemo della questione generale, mentre ci occuperemo solamente di determinare le funzioni di Green ed il nucleo di Poisson per domini Ω particolari.

4.2.1 Il metodo delle cosiddette *cariche immagine.*

Al fine di ottenere una funzione di Green per un dominio $\Omega \subset \mathbb{R}^3$, si deve trovare una classe di funzioni $\{v_\Omega(x,\cdot)\}_{x\in\Omega} \subset C^2(\overline{\Omega})$ (come visto nel teorema 4.1, si rinforza la richiesta di regolarità richiedendo $v_\Omega(\cdot,\cdot) \in C^3(\overline{\Omega}\times\overline{\Omega}\setminus\Delta)$) tale che siano soddisfatte le condizioni:

$$\Delta_y v_\Omega(x,y) = 0 \qquad x \in \Omega, y \in \Omega$$

e

$$G_3(x,y) + v_\Omega(x,y) = 0 \qquad x \in \Omega, y \in \partial\Omega\,.$$

Quest'ultima condizione si scrive esplicitamente:

$$-\frac{1}{4\pi||x-y||} + v_\Omega(x,y) = 0 \qquad x \in \Omega, y \in \partial\Omega\,.$$

L'intuizione fisica ci può aiutare nella ricerca della funzione $v_\Omega(x,y)$. Partiamo dal fatto che, come ben noto dai corsi di elettromagnetismo elementare, la funzione

$$\mathbb{R}^3 \ni y \mapsto \varphi(y) := \frac{e}{4\pi||x-y||}$$

è il *potenziale elettrostatico generato nel punto y da una carica elettrica e, puntiforme, posta nel punto x.* Il gradiente di tale campo, cambiato di segno,

$$\mathbf{E}(y) = \frac{e(y-x)}{4\pi||x-y||^3}\,,$$

coincide con il campo elettrostatico generato in y dalla carica detta. Infine, seguendo la formula della forza di Lorentz (1.2) il prodotto di tale campo per la carica e' di prova posta nel punto y riproduce la **legge di Coulomb**:

$$\mathbf{F}(y) = \frac{ee'(y-x)}{4\pi||x-y||^3} \,,$$

che esprime la forza elettrostatiche che la carica e in x esercita sulla carica e' in y.
La condizione:

$$G_3(x,y) + v_\Omega(x,y) = 0 \qquad x \in \Omega, y \in \partial\Omega \,,$$

dice che il potenziale elettrostatico in y totale dovuto sia ad una carica puntiforme negativa unitaria situata in x (che genera il potenziale $G_3(x,y)$ in y) unitamente ad un ulteriore potenziale incognito $v_\Omega(x,y)$, è sempre nullo sulla superficie $\partial\Omega$. Il fatto che $y \mapsto v_\Omega(x,y)$ si possa sempre pensare come un potenziale elettrostatico è dovuto alla richiesta $\Delta_y v_\Omega(x,y) = 0$ che è soddisfatta dai potenziali elettrostatici, come spiegato nella sezione 2.0.5, quando le sorgenti del campo non cadono nel punto y. *Quindi la determinazione della funzione $v_\Omega(x,\cdot)$ è legata alla determinazione di una configurazione di cariche, che non cada su $\partial\Omega$, in aggiunta a quella posta in x che annulli su $\partial\Omega$ il potenziale prodotto dalla carica unitaria posta in x.* Queste cariche, da aggiungersi a quella unitaria già presente in x, vengono chiamate *cariche immagine.* La dipendenza parametrica di v_Ω da x è dovuta al fatto che possiamo muovere a piacimento la carica in x e ci aspettiamo che ciò cambi la distribuzione di cariche immagine.

4.2.2 La funzione di Green nella palla in $\mathbb{R}^3$.

Sia Ω la palla aperta di raggio R in $\mathbb{R}^3$ centrata nell'origine. Vogliamo trovare la funzione di Green per il problema interno alla sfera. Nel caso della palla, il problema di determinare le cariche immagine è molto semplice: si verifica che è sufficiente una sola altra carica da aggiungersi a quella in x al fine di annullare il potenziale in $\partial\Omega$. Se $x'(x)$ è la posizione della carica immagine di valore q_x (può dipendere da x), deve essere:

$$-\frac{1}{4\pi||x-y||} + \frac{q_x}{4\pi||x'(x)-y||} = 0 \qquad y \in \partial\Omega.$$

Per la simmetria del problema ci aspettiamo che

$$x' = \lambda(x)x \,, \quad \text{con } |\lambda(x)|\,||x|| \geq R \,,$$

la seconda condizione è dovuta al fatto che x' deve essere fuori da $\overline{\Omega}$ altrimenti avremmo una singolarità in più per G_Ω, mentre sappiamo che essa è singolare solo per $y = x$). In questo modo, se $x \in \Omega$ allora $x'(x) \notin \overline{\Omega}$. Si osservi ancora che, per ogni $x \in \Omega$ fissato, cioè x' fissato, la funzione

$$v_\Omega(x,y) := \frac{q_x}{4\pi||x'(x)-y||}$$

è C^∞ ed armonica nella variabile $y \in \Omega$ proprio come richiesto, dato che non è altro che la funzione $y \mapsto q_x G_3(x'(x), y)$. Questa funzione, al variare di $y \in \mathbb{R}^3$ è di classe C^∞ se $y \neq x'$ ed è anche armonica in tale insieme.
Se il metodo funziona si deve avere l'annullamento del potenziale totale su $\partial\Omega$, in particolare nei due punti y_1 e y_2 intersezione della retta congiungente x e x' con $\partial\Omega$:

$$\begin{cases} -\dfrac{1}{4\pi(R - ||x||)} + \dfrac{q_x}{4\pi(||x'(x)|| - R)} = 0 \\ -\dfrac{1}{4\pi(||x|| + R)} + \dfrac{q_x}{4\pi(||x'(x)|| + R)} = 0 \end{cases} \tag{4.9}$$

In questo modo si ottiene, se $||x|| < R$:

$$\begin{cases} ||x'(x)||\,||x|| = R^2\,, \\ q_x = \dfrac{R}{||x||}\,. \end{cases} \tag{4.10}$$

Dalla prima equazione ricaviamo che

$$|\lambda(x)| = \frac{R^2}{||x||^2}\,,$$

in modo tale che, come richiesto, $||x||||\lambda| = R^2/||x|| \geq R$ se $||x|| \leq R$. Assumendo $\lambda > 0$, ci aspettiamo che la funzione di Green cercata sia:

$$G_\Omega(x, y) = -\frac{1}{4\pi||x - y||} + \frac{R||x||}{4\pi||R^2 x - ||x||^2 y||} \tag{4.11}$$

La funzione $v_\Omega(x, y) = G_\Omega(x, y) - G_n(x, y)$ soddisfa $\Delta_y v_\Omega(x, y) = 0$ dove non è singolare, essendo per costruzione, una soluzione fondamentale nelle variabili x' e y:

$$v_\Omega(x, y) := \frac{q_x}{4\pi||x'(x) - y||}$$

Ciò accade in particolare per $x \in \Omega$ e $y \in \overline{\Omega}$ come richiesto nella definizione di funzione di Green. Infatti, più fortemente si può provare che *v_Ω è di classe C^∞ su $\overline{\Omega} \times \overline{\Omega}$ se si escludono i punti (x, y) con $x = y \in \partial\Omega$*. Questo risultato si ottiene facilmente se si nota che, esplicitando i calcoli, e dove $\alpha(x, y)$ è l'angolo tra x e y:

$$v_\Omega(x, y) = \frac{R}{4\pi\sqrt{R^4 + ||x||^2||y||^2 - 2R^2||x||||y||\cos\alpha(x, y)}}\,. \tag{4.12}$$

Ma, dato che $-1 \leq \cos\alpha(x, y) \leq 1$, abbiamo anche che:

$$(R^2 - ||x||||y||)^2 = R^4 + ||x||^2||y||^2 - 2R^2||x||||y|| \leq R^4 + ||x||^2||y||^2 - 2R^2||x||||y||\cos\alpha(x, y), \tag{4.13}$$

Possiamo osservare che, perchè si annulli il denominatore nel secondo membro di (4.12) con $x, y \in \overline{\Omega}$ (e quindi $||x||, ||y|| \leq R$), deve necessariamente accadere che $(R^2 - ||x||||y||)^2 = 0$ e dunque $||x|| = ||y|| = R$, cioè $x, y \in \partial\Omega$. In questa situazione, la condizione di annullamento del denominatore del secondo membro della (4.12) fornisce:

$$R^4 + R^2R^2 - 2R^4 \cos\alpha(x, y) = 0$$

che è possibile solo se $\cos\alpha(x, y) = 1$ e quindi $x = y \in \partial\Omega$ come detto sopra. Pertanto le singolarità di $v_\Omega(x, y)$ su $\overline{\Omega} \times \overline{\Omega}$ si hanno solo quando $x = y \in \partial\Omega$. Possiamo riscrivere l'espressione trovata come:

$$v_\Omega(x, y) = \frac{R}{4\pi\sqrt{R^4 + ||x||^2||y||^2 - 2R^2 x \cdot y}} \, .$$

Questa funzione, per $x, y \in \Omega$, è C^∞ dove non si annulla il denominatore, cioè ovunque su $\overline{\Omega} \times \overline{\Omega}$ escludendo i punti (x, y) con $x = y \in \partial\Omega$.
Infine, si controlla facilmente che anche la seconda condizione richista per avere una funzione di Green è soddisfatta: se $y \in \partial\Omega$ e $x \in \Omega$, allora $G_\Omega(x, y) = 0$. Infatti, in questo caso se $\mathbf{n}_x$ e $\mathbf{n}_y$ sono rispettivamente il versore di x e quello di y e tenendo conto che $||y|| = R$ equivale a $y \in \partial\Omega$, si trova da (4.11):

$$\begin{aligned} G_\Omega(x, y)|_{y\in\partial\Omega} &= -\frac{1}{4\pi|| \, ||x||\mathbf{n}_x - R\mathbf{n}_y||} + \frac{1}{4\pi||R\mathbf{n}_x - ||x||\mathbf{n}_y||} \\ &= -\frac{1}{4\pi|| \, ||x||\mathbf{n}_x - ||y||\mathbf{n}_y||} + \frac{1}{4\pi|| \, ||y||\mathbf{n}_x - ||x||\mathbf{n}_y||} \\ = -\frac{1}{4\pi\sqrt{||x||^2 + ||y||^2 - 2||x||||y||\mathbf{n}_x \cdot \mathbf{n}_y}} &+ \frac{1}{4\pi\sqrt{||y||^2 + ||x||^2 - 2||y||||x||\mathbf{n}_x \cdot \mathbf{n}_y}} = 0 \, . \end{aligned}$$

Tutte le condizioni richieste sono dunque soddisfatte: *la (4.11) è una funzione di Green per la palla di raggio R centrata nell'origine in $\mathbb{R}^3$.*
Valutiamo il nucleo di Poisson associato $N_\Omega(x, y)$. Utilizziamo coordinate polari sferiche con l'asse z diretto lungo il vettore uscente dall'origine e diretto verso il punto x, partendo da

$$G_\Omega(x, y) = -\frac{1}{4\pi||x - y||} + \frac{1}{4\pi||R\mathbf{n}_x - ||x||y/R||}$$

si trova in coordinate polari, dove $||y||$ è la coordinata radiale:

$$G_\Omega(x, y) = -\frac{1}{4\pi\sqrt{||x||^2 + ||y||^2 - 2||x||||y||\cos\theta}} + \frac{1}{4\pi\sqrt{R^2 + ||x||^2||y||^2/R^2 - 2||x||||y||\cos\theta}} \, . \tag{4.14}$$

Quindi dobbiamo valutare:

$$N_\Omega(x, y) = \mathbf{n}_y \cdot \nabla_y G_\Omega(x, y)|_{\partial\Omega} = \left.\frac{\partial}{\partial||y||} G_\Omega(x, y)\right|_{||y||=R}$$

Nel calcolo si tenga conto di due fatti: (i) dopo aver calcolato le derivate, si può porre $||y|| = R$ (ii) i denominatori dei due addendi a secondo membro in (4.14) coincidono per $||y|| = R$ come notato sopra, coincideranno, per $||y|| = R$ anche i denominatore delle *derivate* di tali espressioni, dato che che altro non sono che i precedenti denominatori elevati alla terza potenza.
Il calcolo produce immediatamente:

$$N_\Omega(x,y) = \left.\frac{R^2 - ||x||^2}{4\pi R||x-y||^3}\right|_{y\in\partial\Omega} \qquad ||x|| < R \tag{4.15}$$

(Per il problema di Dirichlet esterno alla sfera si può ragionare analogamente scambiando il ruolo di x e y e si ottiene, si tenga conto che ora il versore uscente da $\partial\Omega$ punta verso l'origine:

$$N_\Omega(x,y) = \left.\frac{||x||^2 - R^2}{4\pi R||x-y||^3}\right|_{y\in\partial\Omega} \qquad ||x|| > R\,.$$

Non diremo altro sul problema esterno.)
Le ipotesi del teorema 4.1 sono soddisfatte, la condizione (i) è vera dato che $v_\Omega \in C^\infty(\overline{\Omega}\times\overline{\Omega}\setminus\Delta)$ come provato precedentemente, la condizione (ii) (cioè la (4.6)) è anch'essa valida anche se non ne daremo la dimostrazione qui. Pertanto siamo in grado di determinare tutte le soluzioni del problema di Dirichlet in Ω con funzione sorgente di classe $C^2_0(\Omega)$ e dato di Dirichlet $\psi \in C^0(\partial\Omega)$, quando Ω è la palla aperta in $\mathbb{R}^3$ di raggio $R > 0$ centrata nell'origine. In particolare, se $\varphi \in C^2(\Omega)\cap C^0(\overline{\Omega})$ è armonica in Ω (cioè non c'è la sorgente f) e si riduce alla funzione continua ψ su $\partial\Omega$, vale:

$$\varphi(x) = \frac{R^2 - ||x||^2}{4\pi R}\oint_{||y||=R}\frac{\psi(y)}{||x-y||^3}\,dS(y)\,.$$

4.2.3 La funzione di Green nel cerchio in $\mathbb{R}^2$.

Ora Ω è il cerchio in $\mathbb{R}^2$ centrato nell'origine e di raggio $\mathbb{R}$. Vogliamo determinare una funzione di Green $G_\Omega(x,y)$ per tale insieme. In questo caso siamo nel piano, ma il ragionamente è molto simile a quello precedente sviluppato nello spazio, cambiano solo i calcoli. Cerchiamo una distribuzione di cariche all'esterno di $\bar{\Omega}$ che annulli il potenziale sulla circonferenza $\partial\Omega$.

Osservazioni 4.3. Dal punto di vista fisico non possiamo pensare tali cariche come carieche elettriche, visto che siamo in due dimensioni e che il potenziale dovuto ad una carica si deve pensare come logaritmico. A parte questo dettaglio, fondamentale dal punto di vista fisico, ma inessenzaile dal punto di vista matematico, la procedura è la stessa che nel caso tridimensionale.

Dalla definizione 3.1, ci aspettiamo una funzione di Green della forma della funzione armonica in $y \in \Omega\setminus\{x\}$ data da, per ogni x fissato in Ω:

$$G_\Omega(x,y) = \frac{1}{2\pi}\log||x-y|| - \frac{1}{2\pi}\log(q_x||x'(x)-y||)\,,$$

in modo tale che:

$$G_\Omega(x,y) = \frac{1}{2\pi}\log||x-y|| - \frac{1}{2\pi}\log(q_x||x'(x)-y||) = 0 \qquad \forall y \in \partial\Omega.$$

Dobbiamo determinare la carica immagine q_x e la posizione di essa x' in funzione di x.
Esattamente come nel caso precedente, per la simmetria del problema ci aspettiamo che $x' = \lambda(x)x$, con $|\lambda|\,||x|| \geq R$ dato che x' deve essere fuori da $\bar{\Omega}$. Se il metodo funziona si deve avere l'annullamento del potenziale totale su $\partial\Omega$, in particolare nei due punti y_1 e y_2 intersezione della retta congiungente x e x' con $\partial\Omega$:

$$\begin{cases} \frac{1}{2\pi}\log(R-||x||) - \frac{1}{2\pi}\log(q_x(||x'(x)||-R)) = 0 \\ \frac{1}{2\pi}\log(R+||x||) - \frac{1}{2\pi}\log(q_x(||x'(x)||+R)) = 0 \end{cases} \tag{4.16}$$

La soluzione è di nuovo:

$$\begin{cases} ||x||||x'|| = R^2 \\ q_x = ||x||/R \end{cases} \tag{4.17}$$

In particolare $|\lambda| = R^2/||x||^2$. Scegliendo $\lambda > 0$ si trova:

$$G_\Omega(x,y) = \frac{1}{2\pi}\log\frac{||x-y||}{||Rx/||x||-||x||y/R||}\,.$$

Questa funzione è di classe C^∞ su $\overline{\Omega}\times\overline{\Omega}$ se si escludono i punti (x,y) con $x=y$. Inoltre la funzione $v_\Omega(x,y) = G_\Omega(x,y) - G_n(x,y)$ è di classe C^∞ su $\overline{\Omega}\times\overline{\Omega}$ se si escludono i punti (x,y) con $x = y \in \partial\Omega$. Si controlla facilmente che per $||y|| = R$, $G_\Omega(x,y)$ si annulla.
Ragionando come nel problema precedente si può calcolare $N_\Omega(x,y)$, ottenendo:

$$N_\Omega(x,y) = \left.\frac{R^2-x^2}{2\pi R||x-y||^2}\right|_{y\in\partial\Omega} \qquad ||x|| < R$$

(Per il problema esterno si ottiene analogamente un nucleo di Poisson:

$$N_\Omega(x,y) = \left.\frac{x^2-R^2}{2\pi R||x-y||^2}\right|_{y\in\partial\Omega} \qquad ||x|| > R\,,$$

non ci occuperemo oltre del problema esterno.)
Si può verificare che la condizione (4.6) è effettivamente verificata per il nucleo di Poisson trovato, pertanto siamo in grado di determinare tutte le soluzioni del problema di Dirichlet in Ω con funzione sorgente di classe $C_0^2(\Omega)$ e dato di Dirichlet $\psi \in C^0(\partial\Omega)$, quando Ω è il disco aperto in $\mathbb{R}^2$ di raggio $R > 0$ centrato nell'origine. In particolare, se $\varphi \in C^2(\Omega)\cap C^0(\overline{\Omega})$ è armonica in Ω (cioè non c'è la sorgente f) e si riduce alla funzione continua ψ su $\partial\Omega$, vale:

$$\varphi(x) = \frac{R^2-||x||^2}{2\pi R}\oint_{||y||=R}\frac{\psi(y)}{||x-y||^2}\,dS(y)\,.$$

4.2.4 La funzione di Green in un semispazio di $\mathbb{R}^3$.

Consideriamo la seguente regione

$$\Omega = \{(x^1, x^2, x^3) \in \mathbb{R}^3 \,|x^3 > 0\} \qquad \text{con, quindi:} \quad \partial\Omega = \{(x^1, x^2, x^3) \in \mathbb{R}^3 \,|x^3 = 0\}$$

Vogliamo determinare una funzione di Green per questa regione illimitata. Possiamo vedere questa regione come il limite per $n \to +\infty$ delle regioni Ω_n date dall'intersezione di Ω e delle palle di raggio n centrate nell'origine. Non ci occuperemo della questione in dettaglio, diremo solo che i precedenti teoremi si generalizzano a questo caso quando si lavora con soluzioni dell'equazione di Poisson che decadono rapidamente a zero (con le lor derivate prime) all'infinito.
Dalla definizione 3.1, sappiamo che in $\mathbb{R}^3$:

$$G_3(x, y) = \frac{-1}{4\pi||x - y||} \, .$$

Dobbiamo allora cercare una distribuzione di cariche (nel semispazio $\{x^3 < 0\}$) tale da annullare il potenziale G_3 su $\partial\Omega$. È sufficiente porre una carica unitaria in $x' = (x^1, x^2, -x^3)$, dove $x = (x^1, x^2, x^3)$. Otteniamo dunque:

$$G_\Omega(x, y) = \frac{-1}{4\pi}\Big(\frac{1}{||x - y||} - \frac{1}{||x' - y||}\Big)$$

$$= \frac{-1}{4\pi}\Big(\frac{1}{\sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}} - \frac{1}{\sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 + y^3)^2}}\Big)$$

Il nucleo di Poisson $N_\Omega(x, y)$ risulta essere, tenendo conto che $\mathbf{n} = -\,\mathbf{e}_3$:

$$N_\Omega(x, y) = -\frac{\partial}{\partial y^3} G_\Omega(x, y)\Big|_{y^3=0}$$

$$= \frac{1}{4\pi}\frac{\partial}{\partial y^3}\Big(\frac{1}{\sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}} - \frac{1}{\sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 + y^3)^2}}\Big)\Big|_{y^3=0} =$$

$$= \frac{1}{4\pi}\frac{2x^3}{||x - y||^3}\Big|_{y^3=0} = \frac{1}{2\pi}\frac{x \cdot \mathbf{e}_3}{||x - y||^3}\Big|_{y^3=0} \, .$$

Consideriamo ora una funzione *continua e limitata* $\psi = \psi(y^1, y^2)$ definita sul piano $y^3 = 0$. Per tale funzione ha senso l'integrale, valutato per $x \in \Omega$ (cioè $x^3 > 0$):

$$\varphi(x) = \frac{x^3}{2\pi}\oint_{y^3=0} \frac{\psi(y)}{||x - y||^3}\, dS(y) \, . \tag{4.18}$$

Esplicitamente, passando in coordinate polari piane sul piano $x^3 = 0$ e scegliendo l'origine nel punto (x^1, x^2):

$$\varphi(x^1, x^2, x^3) = \frac{x^3}{2\pi}\int_0^{2\pi} d\theta \int_0^{+\infty} dr \frac{r\psi(r, \theta)}{(r^2 + (x^3)^2)^{3/2}} \, .$$

È evidente che la limitatezza di ψ assicura l'assoluta convergenza dell'integrale. Il calcolo diretto del laplaciano del primo membro, passando le derivate sotto il segno di integrale (si provi per esercizio che è possibile), provano che la funzione φ è effettivamente $C^2(\Omega)$ ed armonica in tale insieme. Inoltre si prova, essenzialmente verificando la proposizione 4.2, che vale la condizione:

$$\lim_{x \to x_0 \in \partial\Omega} \frac{x^3}{2\pi} \int_{y^3=0} \frac{\psi(y)}{||x-y||^3}\, dS(y) = \psi(x_0)$$

e pertanto (4.18) produce una soluzione del problema di Dirichlet in Ω, senza sorgente f e con dato al bordo ψ. Lasciamo per esercizio la formulazione di un corrispondente teorema di unicità delle soluzioni, tenendo conto del decadimento all'infinito delle funzioni φ individuate in (4.18).

4.3 *Calcolo del nucleo di Poisson per il problema del cerchio in $\mathbb{R}^2$ tramite l'analisi di Fourier

In quest'ultima sezione presentiamo un metodo alternativo a quello delle cariche immagine che permette, sotto opportune ipotesi sul dato al bordo f, di ricavare la soluzione del problema di Dirichlet per il cerchio in $\mathbb{R}^2$. Gli elementi della teoria delle serie di Fourier che usiamo qui saranno discussi nel capitolo 6. Si pù posticipare la lettura di questa sezione a quella della sezione rilevante nel capitolo 6.
Consideriamo dunque il seguente problema di Dirichlet

$$\begin{cases} \Delta\phi = 0 & \phi \in C^2(\Omega) \cap C^0(\overline{\Omega}) \\ \phi|_{\partial\Omega} = f & f \in C^0(\partial\Omega) \text{ assegnata,} \end{cases} \tag{4.19}$$

dove Ω è il cerchio in $\mathbb{R}^2$ di raggio R: $\Omega = \{(x^1, x^2) \in \mathbb{R}^2,\ (x^1)^2 + (x^2)^2 < R^2\}$. Per la simmetria del problema è conveniente introdurre un sistema di coordinate polari ρ, θ in $\mathbb{R}^2$, con $x^1 = \rho\cos\theta, x^2 = \rho\sin\theta$. L'equazione di Laplace assume la seguente forma:

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho\frac{\partial}{\partial\rho}\phi) + \frac{1}{\rho^2}\frac{\partial^2}{\partial\theta^2}\phi = 0\,. \tag{4.20}$$

Per risolverla cominciamo a cercare delle soluzioni particolari della forma:

$$\phi(\rho, \theta) = \psi(\rho)\chi(\theta),$$

con ψ e χ differenziabili due volte con continuità nel loro dominio. Restringendo le possibili soluzioni a tale classe di funzioni, l'equazione (4.20) diviene:

$$\Big(\frac{\partial^2}{\partial\rho^2}\psi + \frac{1}{\rho}\frac{\partial}{\partial\rho}\psi\Big)\chi + \frac{\psi}{\rho^2}\frac{\partial^2}{\partial\theta^2}\chi = 0\,,$$

ovvero, nei punti in cui le due funzioni ψ e χ non si annullano:

$$\Big(\frac{\partial^2}{\partial\rho^2}\psi + \frac{1}{\rho}\frac{\partial}{\partial\rho}\psi\Big)\frac{\rho^2}{\psi} = -\frac{1}{\chi}\frac{\partial^2}{\partial\theta^2}\chi\,.$$

Al primo membro troviamo una funzione dipendente solo dalla variabile ρ, mentre al secondo membro troviamo una funzione dipendente solo dalla variabile θ. L'uguaglianza può dunque essere verificata se e solo se entrambi i membri sono uguali ad una costante λ (indipendente da ρ, θ). Per continuità ci aspettiamo che valga ovunque:

$$\begin{cases} \rho^2\frac{\partial^2}{\partial\rho^2}\psi + \rho\frac{\partial}{\partial\rho}\psi = \lambda\psi \\ \frac{\partial^2}{\partial\theta^2}\chi = -\lambda\chi\,. \end{cases} \tag{4.21}$$

Sicuramente, per ogni fissata costante λ, le soluzioni di questo sistema risolvono l'equazione iniziale. Studiamo pertanto le soluzioni di tale sistema. Iniziamo a considerare la seconda equazione:

$$\frac{\partial^2}{\partial\theta^2}\chi(\theta) + \lambda\chi(\theta) = 0, \tag{4.22}$$

che è un'equazione differenziale ordinaria lineare del secondo ordine a coefficienti costanti di cui ben nota la soluzione generale. Dato che vogliamo, ovviamente che la soluzione soddisfi la condizione di monodromia $\chi(0) = \chi(2\pi)$, è necessario che $\mathbb{R} \ni \lambda \geq 0$, in modo tale che le soluzioni di (4.22) siano funzioni periodiche. In tal caso la soluzione generale è data da

$$\chi(\theta) = A\cos(\sqrt{\lambda}\theta) + B\sin(\sqrt{\lambda}\theta)\,.$$

Inoltre $\sqrt{\lambda}$ deve essere un numero intero, ovvero $\lambda = n^2$, $n \in \mathbb{N}$ per fornire la giusta periodicità che assicuri $\chi(0) = \chi(2\pi)$. Il caso $n = 0$ corrisponde a soluzioni costanti.
Consideriamo ora la seconda equazione, che un'equazione lineare:

$$\rho^2\frac{\partial^2}{\partial\rho^2}\psi + \rho\frac{\partial}{\partial\rho}\psi = n^2\psi, \tag{4.23}$$

Una classe di soluzioni linearmente indipendenti è data da:

$$\psi_n(\rho) = \begin{cases} 1, \log\rho, \\ \rho^n, \rho^{-n} \qquad n \geq 1 \end{cases}$$

Dobbiamo scartare la funzione $\log\rho$ per $n = 0$ e la funzione ρ^{-n} per $n \geq 1$ in quanto sono singolari in 0, mentre noi stiamo cercando delle soluzioni dell'equazione di Laplace che siano regolari in tutto il cerchio interno. Per linearità , una soluzione ovunque regolare dell'equazione (4.20) può avere la forma

$$\psi(\rho,\theta) = \sum_n \rho^n(A_n\cos(n\theta) + B_n\sin(n\theta)), \tag{4.24}$$

dove A_n, B_n sono coefficienti arbitrari e n varia in qualche sottoinsieme finito (ma arbitrariamente grande) dell'insieme degli interi positivi o nulli. I coefficienti A_n e B_n, in linea di principio si determinano imponendo la condizione al bordo $\psi(R,\theta) = f$:

$$\begin{cases} A_0 = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta \\ A_n = \frac{1}{R^n\pi}\int_0^{2\pi} f(\theta)\cos(n\theta)d\theta \\ B_n = \frac{1}{R^n\pi}\int_0^{2\pi} f(\theta)\sin(n\theta)d\theta \end{cases}$$

È chiaro che affinché si possa sempre risolvere il problema ci aspettiamo che l'insieme di variabilità di n sia tutto $\mathbb{N}$ e non solo un sottoinsieme finito. Questo fatto però pone il problema della convergenza della serie (4.24). Ulteriormente bisogna anche dimostrare che la serie (4.24) converge ad una soluzione del problema: questo fatto non è ovvio, mentre è ovvio, per linearità quando n varia su un insieme finito.
Formalmente dunque la soluzione del problema di Dirichlet interno è data dalla *serie*

$$\phi(\rho,\theta) = \frac{\alpha_0}{2} + \sum_{n=1}^{+\infty}(\rho/R)^n(\alpha_n\cos(n\theta) + \beta_n\sin(n\theta)),$$

dove α_n, β_n sono i coefficienti di Fourier della funzione f:

$$\alpha_0 = \frac{1}{\pi}\int_0^{2\pi} f(\theta)d\theta, \quad \alpha_n = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\cos(n\theta)d\theta, \quad \beta_n = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin(n\theta)d\theta.$$

Dobbiamo ora dimostrare che la funzione cosıcostruita è effettivamente soluzione del problema di Dirichlet, ovvero che è di classe $C^2(\Omega)\cap C^0(\overline{\Omega})$, che è armonica in Ω e che coincide con f su $\partial\Omega$. Prima di tutto dimostriamo l'armonicità di ϕ in Ω. Dobbiamo in particolare mostrare che è giustificato derivare sotto il segno di serie. Di fatto, nell'ipotesi che la funzione f sia limitata, i suoi coefficienti di Fourier α_n, β_n sono limitati, sia ha cioè che

$$|\alpha_n| < M,\ |\beta_n| < M, \qquad \forall n \in \mathbb{N}.$$

Ne segue quindi la seguente diseguaglinza per i termini della serie di Fourier di ϕ:

$$|(\rho/R)^n(\alpha_n\cos(n\theta) + \beta_n\sin(n\theta))| < |\rho/R|^n 2M$$

e dunque nei punti interni al cerchio, ovvero per $\rho > R$, sia ha $|\rho/R| < 1$ e dunque la serie è assolutamente convergente.
Possiamo ripetere lo stesso ragionamento anche per dimostrare la convergenza della serie delle derivate. Derivando k volte ogni termine della serie rispetto alla variabile θ otteniamo

$$\sum_{n\geq 1}(\rho/R)^n(\alpha_n n^k\cos(n\theta + k\frac{\pi}{2}) + \beta_n n^k\sin(n\theta + k\frac{\pi}{2})).$$

Tale serie è maggiorata dalla serie convergente $2M\sum_{n\geq 1}(\rho/R)^n n^k$ (se $\rho < R$), è giustificato dunque derivare sotto il segno di serie infinite volte rispetto a θ. Per quanto riguarda la derivazione rispetto alla variabile ρ abbiamo:

$$\frac{\partial^k}{\partial\rho^k}\phi(\rho,\theta) = \sum_{n\geq k} R^{-k}\frac{n!}{n-k!}(\rho/R)^{n-k}(\alpha_n\cos(n\theta)+\beta_n\sin(n\theta)).$$

É immediato verificare che la serie delle derivare è assolutamente convergente, in quanto maggiorata da $2M\sum_{n\geq k} R^{-k}\frac{n!}{n-k!}(\rho/R)^{n-k}$, dunque è possibile derivare sotto il segno di serie infinite volte rispetto a ρ.
Resta da verificare la continuità della soluzione cosıcostruita sul bordo Ω. Per dimostrare questa proprietà dobbiamo imporre delle ipotesi aggiuntive su f. Sappiamo infatti dalla teoria delle serie di Fourier che se f è continua e di classe C^1 a tratti su $\partial\Omega$, allora i suoi coefficienti di Fourier α_n, β_n soddisfano le seguenti diseguaglianze:

$$\sum_n |\alpha_n| < \infty, \qquad \sum_n |\beta_n| < \infty.$$

La funzione ϕ è , per costruzione, il limite di una successione di funzioni continue, infatti:

$$\phi(\rho,\theta) = \frac{\alpha_0}{2} + \lim_{N\to\infty}\sum_{n=1}^{N}\phi_n(\rho,\theta), \quad \phi_n(\rho,\theta) = (\rho/R)^n(\alpha_n\cos(n\theta)+\beta_n\sin(n\theta)).$$

Inoltre, $\forall\rho\in[0,R], \theta\in[0,2\pi]$, abbiamo $|\phi_n(\rho,\theta)| < |\alpha_n|+|\beta_n|$, quindi, utilizzando il teorema che assicura che il limite uniforme di una successione di funzioni continue è una funzione continua, abbiamo che $\phi(\rho,\theta)$ è una funzione continua in $\overline{\Omega}$. In particolare sul bordo di Ω:

$$\phi(R,\theta) = \frac{\alpha_0}{2} + \sum_{n=1}^{+\infty}(\alpha_n\cos(n\theta)+\beta_n\sin(n\theta)) = f(\theta)\ ,$$

dato che α_n, β_n sono i coefficienti di Fourier della funzione f.
Calcoliamo infine il nucleo di Poisson.

$$
\begin{aligned}
\phi(\rho,\theta) &= \frac{\alpha_0}{2} + \sum_{n\geq 1}(\rho/R)^n(\alpha_n\cos(n\theta)+\beta_n\sin(n\theta)) \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta + \sum_{n\geq 1}(\rho/R)^n\Big(\frac{1}{\pi}\int_0^{2\pi} f(\theta')\cos(n\theta')\cos(n\theta)d\theta' + \\
&\qquad + \frac{1}{\pi}\int_0^{2\pi} f(\theta')\sin(n\theta')\sin(n\theta)d\theta'\Big) \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta + \frac{1}{\pi}\sum_{n\geq 1}(\rho/R)^n\int_0^{2\pi} f(\theta')\cos n(\theta-\theta')d\theta' \\
&= \frac{1}{\pi}\int_0^{2\pi} f(\theta')\Big(\frac{1}{2}+\sum_{n\geq 1}(\rho/R)^n\cos n(\theta-\theta')\Big)d\theta' \\
&= \frac{1}{\pi}\int_0^{2\pi} f(\theta')\Big(\frac{1}{2}+\sum_{n\geq 1}(\rho/R)^n\frac{e^{in(\theta-\theta')}+e^{-in(\theta-\theta')}}{2}\Big)d\theta' \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\theta')\Big(1+\sum_{n\geq 1}\big(\frac{\rho}{R}\,e^{i(\theta-\theta')}\big)^n+\sum_{n\geq 1}\big(\frac{\rho}{R}\,e^{-i(\theta-\theta')}\big)^n\Big)d\theta' \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\theta')\Big(1+\frac{\frac{\rho}{R}\,e^{i(\theta-\theta')}}{1-\frac{\rho}{R}\,e^{i(\theta-\theta')}}+\frac{\frac{\rho}{R}\,e^{-i(\theta-\theta')}}{1-\frac{\rho}{R}\,e^{-i(\theta-\theta')}}\Big)d\theta' \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\theta')\frac{1-(\rho/R)^2}{1-2(\rho/R)\cos(\theta-\theta')+(\rho/R)^2}d\theta' \\
&= \int_0^{2\pi} f(\theta')N_\Omega(\rho,\theta,R,\theta')Rd\theta'
\end{aligned}
$$

Dove $N_\Omega(\rho,\theta,R,\theta') = \frac{1}{2\pi R}\frac{1-(\rho/R)^2}{1-2(\rho/R)\cos(\theta-\theta')+(\rho/R)^2}$ è il nucleo di Poisson, già calcolato in precedenza tramite il metodo della cariche immagine.

Capitolo 5

Equazioni iperboliche: alcuni risultati generali elementari per le equazioni di D'Alembert e di Klein-Gordon in $\mathbb{R} \times \mathbb{R}^n$.

In questo capitolo ci occuperemo di alcuni fatti generali riguardanti due equazioni del secondo ordine di tipo iperbolico: l'equazione di D'Alembert e quella di Klein-Gordon. La prima:

$$-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} + \Delta_{\mathbf{x}}\varphi = 0 \tag{5.1}$$

con $\varphi = \varphi(t, \mathbf{x})$, $(t, \mathbf{x}) \in \mathbb{R}\times\mathbb{R}^n$ è ben nota dalla fisica classica e descrive nello spaziotempo $\mathbb{R}\times\mathbb{R}^n$, in una certa approssimazione, tutti i fenomeni di propagazione ondosa/elastica in mezzi estesi in $\mathbb{R}^n$ (tipicamente $n = 1, 2, 3$). La costante c è la velocità di propagazione delle perturbazioni descritte dal campo φ, che dipende dal tipo di mezzo e di perturbazione. La seconda:

$$-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} + \Delta_{\mathbf{x}}\varphi - \mu^2\varphi = 0, \tag{5.2}$$

dove $\mu = \frac{mc^2}{\hbar}$ con $\varphi = \varphi(t, \mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ è un'equazione che nasce nella fisica moderna e descrive (se $\mathbb{R}^n = \mathbb{R}^4$ pensato come spaziotempo della relatività speciale e c è la velocità della luce) l'equazione di evoluzione relativistica per campo associato a particelle quantistiche di massa $m > 0$ e prive di spin e carica. Nel caso $m = 0$, ovviamente la forma della seconda equazione si riduce alla forma della prima. Le due equazioni sopra scritte possono essere leggermente modificate introducendo un *termine di sorgente* dato da una funzione nota $\rho = \rho(t, \mathbf{x})$:

$$-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} + \Delta_{\mathbf{x}}\varphi = \rho \tag{5.3}$$

e

$$-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} + \Delta_{\mathbf{x}}\varphi - \mu^2\varphi = \rho. \tag{5.4}$$

Nota importante. Nel seguito del capitolo Δ indicherà sempre e solo il laplaciano rispetto alle coordinate spaziali $\mathbf{x}$. Sopra abbiamo indicato tale operatore con $\Delta_{\mathbf{x}}$, ma d'ora in poi ometteremo l'indice $\mathbf{x}$.

5.1 L'equazione di D'Alembert come equazione della corda vibrante.

L'equazione di D'Alembert descrive, in prima approssimazione, tutti i fenomeni di propagazione odulatoria classici in mezzi estesi. A titolo di esempio, vogliamo mostrare come l'equazione di D'Alembert descriva le onde trasversali che si propagano lungo una corda elastica tesa. Consideriamo una corda orizzontale, a riposo descritta da $y = 0$ in un sistema di coordinate x, y solidale con un sistema di riferimento inerziale, con y verticale. Sia $\lambda > 0$, costante (nel tempo e nel punto della corda) la densità lineare di massa della corda lungo x e sia $T > 0$ il valore costante (nel tempo e nel punto della corda) della tensione della corda. Supponiamo che la corda, al tempo $t = 0$, venga deformata in una funzione $y = y(x)$ con $|y(x)|$ molto piccolo, e che poi venga lasciata libera (sempre con estremi fissati). A causa dell'elasticità del mezzo, accade che la configurazione della corda varierà nel tempo e sarà descritta da una funzione $y = y(t, x)$. Vogliamo ricavare, dalle leggi della dinamica, l'equazione a cui deve soddisfare questa funzione assumendo che la tensione T e la densità λ rimangano costanti e che le deformazioni trasversali siano piccole. Consideriamo un punto x_0 e quindi un pezzo di corda relativo all'intervallo $[x_0-h, x_0+h]$. Su tale porzione di corda agisce la tensione ai due estremi: $\mathbf{T}(x_0+h)$ e $\mathbf{T}(x_0-h)$. Entrambi i vettori saranno uscenti dalla porzione di corda e saranno in ogni punto tangenti alla corda. Si osservi che quindi le componenti lungo l'asse x di tali vettori hanno segno opposto. In prima approssimazione l'accelerazione nella direzione $\mathbf{e}_y$ della porzione di corda è $\frac{\partial^2 y}{\partial t}$, mentre la massa della porzione di corda è $2h\lambda$. La seconda equazione della dinamica afferma allora che deve valere, in prima approssimazione:

$$2h\lambda \frac{\partial^2 y}{\partial t^2} = (\mathbf{T}(x_0 + h) + \mathbf{T}(x_0 - h)) \cdot \mathbf{e}_y \,, \tag{5.5}$$

Il secondo membro si può riscrivere come:

$$(\mathbf{T}(x_0 + h) + \mathbf{T}(x_0 - h)) \cdot \mathbf{e}_y = T(\sin\alpha(x_0 + h) - \sin\alpha(x_0 - h)) \,,$$

dove $\alpha(x_0+h)$ e $\alpha(x_0-h)$ sono gli angoli che $\mathbf{T}(x_0+h)$ e $-\mathbf{T}(x_0-h)$ individuano rispetto a $\mathbf{e}_x$ e quindi, approssimando $\sin\alpha$ con $\tan\alpha$ tenendo conto che lavoriamo con piccoli $|y|$ e tenendo conto che: $\tan\alpha(x_0 + h) = \frac{\partial y}{\partial x}|_{x_0+h}$ e $\tan\alpha(x_0 - h) = \frac{\partial y}{\partial x}|_{x_0-h}$, (5.5) può essere riscritta come:

$$\frac{\lambda}{T}\frac{\partial^2 y}{\partial t^2} = \frac{\frac{\partial y}{\partial x}|_{x_0+h} - \frac{\partial y}{\partial x}|_{x_0-h}}{2h} \,.$$

In realtà l'identità trovata è solo approssimata. Tuttavia, nel limite per $h \to 0$, ci si aspetta che diventi rigorosamente valida. In tal caso, si trova l'equazione:

$$\frac{\lambda}{T}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2} \,.$$

Questa è l'equazione di D'Alembert in $\mathbb{R}^2$ per le perturbazioni ondose trasversali della corda:

$$-\frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}+\frac{\partial^2 y}{\partial x^2}=0\,,$$

in cui la *velocità di propagazione delle perturbazioni* v (vedremo più avanti il significato di tale nome) è data da:

$$v=\sqrt{\frac{T}{\lambda}}\,.$$

Osservazioni 5.1. Nel caso in cui sulla corda agisca anche la forza di gravità , sulla porzione di corda usata per ottenere l'equazione di D'Alembert agisce anche la forza verticale $-\lambda 2hg\mathbf{e}_y$. In questo caso, ripetendo il ragionamento fatto sopra, l'equazione finale che si ottiene è quella di D'Alembert con sorgente:

$$-\frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}+\frac{\partial^2 y}{\partial x^2}=\frac{\lambda}{T}g\,.$$

In generale, se sulla corda agisce qualche densità di forza per unità di massa individuata dalla funzione $f=f(t,x)$ nella direzione verticale, l'equazione che si ottiene alla fine è :

$$-\frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}+\frac{\partial^2 y}{\partial x^2}=-\frac{\lambda}{T}f(t,x)\,.$$

5.2 Condizioni iniziali ed al contorno.

I problemi tipici che si incontrano lavorando con equazioni iperboliche come (5.3) e (5.4) sono generalmente del seguente tipo.
Si cerca $\varphi\in C^2((\alpha,\beta)\times\overline{D})$ che soddisfi (5.3) oppure (5.4) in $(\alpha,\beta)\times\overline{D}$ per qualche $\rho\in C^0((\alpha,\beta)\times\overline{D})$ assegnata, dove:

(a) $(\alpha,\beta)\ni 0$

(b) $D\subset\mathbb{R}^n$ è un aperto, non vuoto, (non necessariamente connesso) con $\overline{D}$ compatto e ∂D regolare orientabile.

Vengono quindi assegnate *condizioni iniziali* e *condizioni al bordo* sulla funzione φ.
Le **condizioni iniziali** corrispondono alla coppia di richieste:

$$\varphi(0,\mathbf{x})=\varphi_0(\mathbf{x})\,,\quad \left.\frac{\partial\varphi}{\partial t}\right|_{(0,\mathbf{x})}=\varphi_1(\mathbf{x})\,,\quad \forall\mathbf{x}\in\overline{D},\text{ con }\varphi_0\in C^2(\overline{D})\text{ e }\varphi_1\in C^1(\overline{D})\text{ assegnate.}\tag{5.6}$$

Le **condizioni al bordo**, riferite all'insieme $S:=(\alpha,\beta)\times\partial D$ con vettore normale uscente $\mathbf{n}$, possono essere di tre tipi distinti:

(i) $\varphi\restriction_S = \psi$ con $\psi \in C^2(S)$ funzione assegnata tale che $\psi(0, \mathbf{x}) = \varphi_0(\mathbf{x})$;

(ii) $\mathbf{n} \cdot \nabla\varphi\restriction_S = \psi$ con $\psi \in C^1(S)$ funzione assegnata tale che $\psi(0, \mathbf{x}) = \mathbf{n} \cdot \nabla\varphi_0\restriction_S (\mathbf{x})$;

(iii) $a\varphi\restriction_S + b\mathbf{n} \cdot \nabla\varphi\restriction_S = \psi$ con $a, b \in \mathbb{R}$ costanti assegnate tali che $ab \neq 0$ e $\psi \in C^1(S)$ funzione assegnata tale che $\psi(0, \mathbf{x}) = a\varphi_0\restriction_S + b\mathbf{n} \cdot \nabla\varphi_0\restriction_S$.

Osservazioni 5.2.
(1) Le condizioni dette si possono notevolmente indebolire per esempio assumendo più debolmente che $\varphi \in C^2((0, \beta) \times D) \cap C^1((0, \beta) \times \overline{D})$ (e che soddisfi in tale insieme (5.3) oppure (5.4) per qualche $\rho \in C^0((\alpha, \beta) \times D)$), con $\varphi_0 \in C^1(\overline{D})$ e $\varphi_1 \in C^0(\overline{D})$, e $\psi \in C^1(S)$ in (ii) e $C^0(S)$ in (i) e (iii). In questo caso bisogna assumere più precise ipotesi di regolarità sul dominio D al fine di avere teoremi di esistenza ed unicità .
(2) Si possono considerare casi in cui D non è limitato e sono assegnate condizioni iniziali. In questo caso le condizioni al contorno, che sono importanti per i teoremi di esistenza ed unicità sono, in generale, rimpiazzate da condizioni sull'andamento all'infinito spaziale (cioè per $|\mathbf{x}| \to +\infty$ a t fissato) per il campo φ incognito. Nel caso in cui $D = \mathbb{R}$ e $(\alpha, \beta) = \mathbb{R}$, per l'equazione di D'Alembert non è necessario fissare alcun dato al contorno, come vedremo più avanti, per avere un teorema di esistenza ed unicità .
(3) Esaminaimo il significato delle condizioni al contorno nel caso di una corda orizzontale, di lunghezza fissata, vibrante trasversalmente. Nel caso di condizioni al contorno di tipo (i) la funzione ψ definita sul bordo S si riduce ad una coppia di funzioni $u = u(t)$ e $v = v(t)$, definite sui due estremi della corda, che stabiliscono come oscilla la corda ai suoi estremi al variare del tempo. Le condizioni al contorno di tipo (ii), per la corda vibrante corrispondono a fissare l'andamento temporale della componente verticale della forza che agisce sulla corda agli estremi. Le condizioni al contorno di tipo (iii) corrispondono a fissare una relazione (che dipende dal tempo) tra ciascuna forza che agisce ad ogni estremo e la deformazione della corda nello stesso estremo.

5.3 Bilancio energetico e teoremi di unicità.

5.3.1 Densità di energia ed equazione di continuità.

Consideriamo una funzione φ di classe $C^2(\Omega)$ dove $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ è un aperto sul quale la funzione soddisfa l'equazione di Klein-Gordon (5.4) e quindi in particolare l'equazione di D'Alembert (5.3) nel caso $\mu = 0$. Definiamo su Ω la funzione $E \in C^1(\Omega)$:

$$E(t, \mathbf{x}) := \frac{1}{2} \left[\frac{1}{c^2} \left(\frac{\partial\varphi(t, \mathbf{x})}{\partial t} \right)^2 + \nabla\varphi(t, \mathbf{x}) \cdot \nabla\varphi(t, \mathbf{x}) + \mu^2 \varphi(t, \mathbf{x})^2 \right] . \tag{5.7}$$

Chiameremo la funzione E **densità di energia** di φ. Questa funzione è di fondamentale importanza in matematica oltre che in fisica in quanto consente di provare dei teoremi di unicità per le soluzioni delle equazioni considerate.

Osservazioni 5.3.
(1) In realtà E descrive effettivamente la densità di energia associata a campo φ nel caso in cui esso sia il campo quantistico di Klein-Gordon. Negli altri casi, la grandezza E non ha sempre il significato di densità di energia anche se lo ha in certi casi importanti, per esempio quando φ descrive le deformazioni longitudinali di una sbarra elastica e l'equazione considerata è quella di D'Alembert piuttosto che quella di Klein-Gordon. In tal caso E è davvero la densità di energia elastica del mezzo continuo.
(2) Si osservi che $E(t, \mathbf{x}) \geq 0$ ovunque è definita e questo fatto sarà di cruciale importanza tra poco.

Proposizione 5.1. *Si consideri una funzione φ di classe $C^2(\Omega)$ dove $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ è un aperto sul quale la funzione soddisfa l'equazione di Klein-Gordon (5.4) e quindi in particolare l'equazione di D'Alembert (5.3) nel caso $\mu = 0$. La densità di energia E di φ soddisfa:*

$$\frac{\partial}{\partial t} E(t, \mathbf{x}) = \nabla \cdot \left(\frac{\partial \varphi}{\partial t} \nabla \varphi \right) - \frac{\partial \varphi}{\partial t} \rho(t, \mathbf{x}) \,, \quad \forall (t, \mathbf{x}) \in \Omega \,. \tag{5.8}$$

$\diamondsuit$

Dimostrazione. Per computo diretto, dalla definizione di E:

$$\frac{\partial}{\partial t} E = \frac{\partial \varphi}{\partial t} \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} + \nabla \varphi \cdot \frac{\partial \nabla \varphi}{\partial t} + \mu^2 \varphi \frac{\partial \varphi}{\partial t} \,.$$

Dato che, dall'equazione di Klein-Gordon con sorgente:

$$\frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = \Delta \varphi - \mu^2 \varphi - \rho \,,$$

sostituendo nell'espressione trovata sopra per la derivata temporale di E, abbiamo:

$$\frac{\partial}{\partial t} E = \frac{\partial \varphi}{\partial t} (\Delta \varphi - \mu^2 \varphi - \rho) + \nabla \varphi \cdot \nabla \frac{\partial \varphi}{\partial t} + \mu^2 \varphi \frac{\partial \varphi}{\partial t} \,,$$

dove abbiamo anche scambiato l'ordine di due derivate essendo la funzione φ di classe C^2. Il risultato ottenuto si può riscrivere:

$$\frac{\partial}{\partial t} E = \frac{\partial \varphi}{\partial t} \Delta \varphi + \nabla \varphi \cdot \nabla \frac{\partial \varphi}{\partial t} - \frac{\partial \varphi}{\partial t} \rho \,,$$

e cioè :

$$\frac{\partial}{\partial t} E(t, \mathbf{x}) = \nabla \cdot \left(\frac{\partial \varphi}{\partial t} \nabla \varphi \right) - \frac{\partial \varphi}{\partial t} \rho(t, \mathbf{x}) \,.$$

$\square$

Osservazioni 5.4. In riferimento al teorema precedente, in assenza della sorgente ρ, l'equazione

$$\frac{\partial}{\partial t}E(t,\mathbf{x}) = \nabla \cdot \left(\frac{\partial \varphi}{\partial t}\nabla\varphi\right)$$

può essere riscritta in termini di un'*equazione di continuità*:

$$\frac{\partial}{\partial t}E(t,\mathbf{x}) + \nabla \cdot \mathbf{J}_E = 0\,,$$

dove $\mathbf{J}_E := -\frac{\partial\varphi}{\partial t}\nabla\varphi$. Vediamo il significato fisico di tale equazione. Fissiamo un insieme $[t_1,t_2]\times\overline{V}\subset\Omega$, dove $V\subset\mathbb{R}^n$ è un aperto non vuoto a chiusura compatta il cui bordo è una superficie regolare orientabile e $t_1<t_2$. Un tale insieme esiste nell'inotno di ogni punto di Ω, dato che questo è aperto e che i cilindri aperti $(t_1,t_2)\times V$ sono una base della topologia di $\mathbb{R}\times\mathbb{R}^n$, prendendo, per esempio, gli insiemi V come palle aperte di $\mathbb{R}^n$. *Consideriamo il caso in cui non ci sia la sorgente ρ.* Se integriamo l'equazione (5.8) sul volume V otteniamo:

$$\int_V \frac{\partial}{\partial t}E(t,\mathbf{x})\,d^nx = \int_V \nabla\cdot\left(\frac{\partial\varphi}{\partial t}\nabla\varphi\right)\,d^nx\,.$$

La derivata parziale nel tempo può essere portata fuori dall'integrale, dato che $\frac{\partial}{\partial t}E$ e continua e quindi è limitata sul compatto $[t_1,t_2]\times\overline{V}$ e V ha misura finita (pari a quella di $\overline{V}$)[1]. In questo modo, l'equazione trovata può essere riscritta:

$$\frac{d}{dt}\int_V E(t,\mathbf{x})\,d^nx = \oint_{+\partial V}\left(\frac{\partial\varphi}{\partial t}\nabla\varphi\right)\cdot\mathbf{n}\,dS(x)\,.$$

Questa è , a tutti gli effetti, un'equazione di conservazione (o bilancio) della grandezza che si ottiene integrando E su un volume. L'identità trovata dice che la variazione per unità di tempo dell'energia totale presente nel volume V è pari al flusso di energia entrante che passa attraverso la superficie che circonda V stesso. In questo senso $\mathbf{J}_E = -\frac{\partial\varphi}{\partial t}\nabla\varphi$ si intepreta come la densità di corrente di energia o altrimenti detta il campo di flusso di energia.

5.3.2 Teoremi di unicità.

Possiamo ora eneunciare e provare un teorema di unicità per l'equazione di Klein-Gordon con sorgente (5.4) che include, come caso particolare l'equazione di D'Alembert con sorgente (5.3).

[1]Infatti, dato che E è continua e quindi limitata su $\overline{V}$ per ogni fissato t, è sicuramente (assolutamente) integrabile secondo Lebesgue su tale insieme per ogni valore del parametro t, inoltre $\frac{\partial}{\partial t}E$ una funzione continua congiuntamente nelle due variabili ed è dunque limitata, uniformemente in t, da qualche costante $M\geq 0$ sul compatto $[t_1,t_2]\times\overline{V}$. Dato che la funzione costante $V\ni\mathbf{x}\mapsto M$ (pensata come funzione della sola $\mathbf{x}$) è non negativa ed integrabile su V, avendo quest'ultimo misura finita, siamo nelle ipotesi di poter calcolare la derivata di $t\mapsto\int_V\frac{\partial}{\partial t}E(t,\mathbf{x})\,d^nx$, per ogni $t\in(t_1,t_2)$ passando la derivata in t sotto il segno di integrale (vedi la sezione B.2 in appendice).

Teorema 5.1. *Sia $(\alpha, \beta) \ni 0$ e $D \subset \mathbb{R}^n$ un aperto, non vuoto, con $\overline{D}$ compatto e ∂D regolare orientabile. Si consideri il problema di determinare $\varphi \in C^2((\alpha, \beta) \times \overline{D})$ che soddisfi l'equazione differenziale di Klein-Gordon con sorgente:*

$$-\frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} + \Delta\varphi - \mu^2\varphi = \rho\,,$$

(incluso il caso di D'Alembert $\mu = 0$) dove la costante $\mu \geq 0$ e la funzione $\rho \in C^0((\alpha, \beta) \times \overline{D})$ sono assegnate. Supponendo ulteriormente che siano state imposte condizioni iniziali:

$$\varphi(0, \mathbf{x}) = \varphi_0(\mathbf{x})\,, \quad \left.\frac{\partial \varphi}{\partial t}\right|_{(0,\mathbf{x})} = \varphi_1(\mathbf{x})\,, \quad \forall \mathbf{x} \in \overline{D},\ \textit{con } \varphi_0 \in C^2(\overline{D}) \textit{ e } \varphi_1 \in C^1(\overline{D}) \textit{ assegnate,}$$

e condizioni al contorno di tipo (i) oppure (ii) oppure (iii), con la funzione ψ assgnata come in 5.2. Infine, se si assegnano condizioni al contorno di tipo (iii), le costanti a e b sono supposte soddisfare $ab > 0$ (e non solo $ab \neq 0$).
Se esiste una soluzione questa è unica. $\diamondsuit$

Dimostrazione. Siano φ_1 e φ_2 due soluzioni dello stesso problema di sopra. La funzione $\phi := \varphi_1 - \varphi_2 \in C^2((\alpha, \beta) \times \overline{D})$ risolve allora l'equazione senza sorgente

$$-\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} + \Delta\varphi - \mu^2\phi = 0\,,$$

con condizioni iniziali:

$$\phi(0, \mathbf{x}) = 0\,, \quad \left.\frac{\partial \phi}{\partial t}\right|_{(0,\mathbf{x})} = 0\,, \quad \forall \mathbf{x} \in \overline{D},$$

e condizioni al contorno rispettivamente:(i) $\phi\restriction_S = 0$, oppure (ii) $\mathbf{n} \cdot \nabla\phi\restriction_S = 0$, oppure (iii) $a\phi\restriction_S + b\mathbf{n} \cdot \nabla\phi\restriction_S = 0$ con $ab > 0$, dove $S := (\alpha, \beta) \times \partial D$.
Ragionando esattamente come nelle osservazioni 5.4 arriviamo a concludere che:

$$\frac{d}{dt}\int_D E(t, \mathbf{x})\, d^n x = \oint_{+\partial D} \left(\frac{\partial \phi}{\partial t}\nabla\phi\right) \cdot \mathbf{n}\, dS(x)\,.$$

Si osservi che il secondo membro è una funzione continua di t come si prova subito dal teorema della convergenza dominata notando che ∂D ha misura finita e che la funzione integranda è congiuntamente continua in tutte le variabili (vedi la sezione B.2 in appendice). Concludiamo che, per ogni $T \in (\alpha, \beta)$:

$$\int_D E(T, \mathbf{x})\, d^n x = \int_0^T dt \oint_{+\partial D} \left(\frac{\partial \phi}{\partial t}\nabla\phi\right) \cdot \mathbf{n}\, dS(x)\,. \tag{5.9}$$

Dove abbiamo tenuto conto del fatto che, nelle nostre ipotesi E per il campo ϕ si annulla a $t = 0$ e quindi $\int_D E(0, \mathbf{x}) d^n x = 0$. Nel caso di condizioni al contorno di tipo (i) e (ii) il secondo membro di (5.9) è evidentemente nullo. Nel caso di condizioni al contorno di tipo (iii) si ottiene lo stesso

risultato con un pò più di fatica come proveremo alla fine. Concludiamo che, nelle nostre ipotesi, per ogni tipo di condizione al contorno e per ogni $T \in (\alpha, \beta)$, vale $\int_D E(T, \mathbf{x})\, d^n x = 0$ e quindi la funzione $E \geq 0$ deve essere quasi ovunque nulla. Essendo continua deve essere ovunque nulla. In definiva abbiamo ottenuto che, per ogni $(t, \mathbf{x}) \in (\alpha, \beta) \times \overline{D}$:

$$E(t, \mathbf{x}) := \frac{1}{2}\left[\frac{1}{c^2}\left(\frac{\partial \phi(t, \mathbf{x})}{\partial t}\right)^2 + \nabla\phi(t, \mathbf{x}) \cdot \nabla\phi(t, \mathbf{x}) + \mu^2 \phi(t, \mathbf{x})^2\right] = 0\,.$$

Dato che si tratta di una somma di addendi non negativi ogni addendo deve essere nullo separatamente. Se $\mu > 0$ concludiamo che $\phi = 0$ ovunque e quindi $\varphi_1 = \varphi_2$ su $(\alpha, \beta) \times \overline{D}$. Lo stesso risultato si ottiene se $\mu = 0$ osservando che, in virtù di quanto ottenuto sopra, le derivate temporali di ϕ devono annullarsi. Concludiamo (applicando il teorema di Lagrange) che per ogni fissato $\mathbf{x} \in \overline{D}$, $\phi(t, \mathbf{x}) = \phi(0, \mathbf{x})$ per ogni $t \in (\alpha, \beta)$. Ma $\phi(0, \mathbf{x}) = 0$ nelle nostre ipotesi. In definitiva $\varphi_1 = \varphi_2$ vale su $(\alpha, \beta) \times \overline{D}$ e quindi il teorema di unicità è provato.
Per concludere la dimostrazione proviamo che il secondo membro di (5.9) è nullo anche per condizioni al bordo di tipo (iii). Dato che $\nabla\phi \cdot \mathbf{n} = -\frac{a}{b}\phi$, il secondo membro può ancora essere scritto,

$$-\int_0^T dt \oint_{+\partial D} \frac{a}{b}\frac{\partial \phi}{\partial t}\phi\, dS(x) = -\frac{a}{2b}\int_0^T dt \oint_{+\partial D} \frac{\partial \phi^2}{\partial t}\, dS(x) = -\frac{a}{2b}\int_0^T dt \frac{d}{dt}\oint_{+\partial D} \phi^2\, dS(x)\,.$$

Dove abbiamo usato note conseguenze del teorema della convergenza dominata (vedi la sezione B.2 in appendice). In definitiva, dato che $\phi^2(0, \mathbf{x}) = 0$ su $\overline{D}$:

$$\int_D E(T, \mathbf{x})\, d^n x = -\frac{a}{2b}\int_0^T dt \frac{d}{dt}\oint_{+\partial D} \phi^2\, dS(x) = -\frac{a}{2b}\oint_{+\partial D} \phi^2(T, x)\, dS(x)\,.$$

Si osservi che se $ab > 0$ significa che a e b hanno lo stesso segno e pertanto:

$$-\frac{a}{2b}\oint_{+\partial D} \phi^2(T, x)\, dS(x) \leq 0\,.$$

D'altra parte, dato che $E \geq 0$ abbiamo anche che $\int_D E(T, \mathbf{x})\, d^n x \geq 0$. Di conseguenza l'identità ottenuta:

$$\int_D E(T, \mathbf{x})\, d^n x = -\frac{a}{2b}\oint_{+\partial D} \varphi^2(T, x)\, dS(x)\,,$$

implica che: $\int_D E(T, \mathbf{x})\, d^n x = 0$. $\square$

Osservazioni 5.5.
(1) Con una procedura di limite ed eseguendo in ordine diverso alcune delle integrazioni fatte nella dimostrazione di sopra, il risultato presentato nel teorema si può estendere al caso in cui si richiede più debolmente $\varphi \in C^2((0, \beta) \times D) \cap C^1((0, \beta) \times \overline{D})$, con $\varphi_0 \in C^1(\overline{D})$ e $\varphi_1 \in C^0(\overline{D})$, e $\psi \in C^1(S)$ in (ii) e $C^0(S)$ in (i) e (iii). In questa situazione però è necessario assumere che

il volume D sia più regolare e che sia ottenibile (in un preciso senso che non chiariremo qui) come limite di una successione di domini $D_1 \subset \cdots \subset D_n \subset D_{n+1} \subset \cdots \subset D$ in modo tale che $Vol(\partial D_n) \leq Vol(\partial D)$.
(2) Lavorando su tutto lo spazio $\mathbb{R}^n$, si può dimostrare, e noi lo faremo per l'equazione di D'Alembert sulla retta reale, che se al tempo $t = 0$ i dati iniziali sono a supporto compatto, allora la soluzione $\varphi = \varphi(t, \mathbf{x})$, dell'equazione di Klein-Gordon/D'Alembert senza sorgente su $(\alpha, \beta) \times \mathbb{R}^n$, ha supporto compatto quando ristretta ad ogni insieme $[\alpha', \beta'] \times \mathbb{R}^n$, con $[\alpha', \beta'] \subset (\alpha, \beta)$. Questo risultato non è per nulla ovvio, per esempio non vale per equazioni paraboliche oppure per l'equazione di Schrödinger. In base a tale risultato il seguente teorema di unicità non risulta essere inutile.

Teorema 5.2. *Sia $(\alpha, \beta) \ni 0$ Si consideri il problema di determinare $\varphi \in C^2((\alpha, \beta) \times \mathbb{R}^n)$ che soddisfi l'equazione differenziale di Klein-Gordon con sorgente:*

$$-\frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} + \Delta\varphi - \mu^2\varphi = 0\,,$$

(incluso il caso di D'Alembert $\mu = 0$) dove la costante $\mu \geq 0$ è assegnata. Supponendo ulteriormente che siano state imposte condizioni iniziali:

$$\varphi(0, \mathbf{x}) = \varphi_0(\mathbf{x})\,, \quad \left.\frac{\partial \varphi}{\partial t}\right|_{(0,\mathbf{x})} = \varphi_1(\mathbf{x})\,, \quad \forall \mathbf{x} \in \overline{D},\ \textit{con}\ \varphi_0 \in C^2_0(\overline{D})\ \textit{e}\ \varphi_1 \in C^1_0(\overline{D})\ \textit{assegnate.}$$

Se esiste una soluzione φ tale che ha supporto compatto quando ristretta ad ogni insieme $[\alpha', \beta'] \times \mathbb{R}^n$, con $[\alpha', \beta'] \subset (\alpha, \beta)$, tale soluzione è unica. $\diamondsuit$

Dimostrazione. Siano φ_1 e φ_2 due soluzioni del problema. Consideriamo $\phi := \varphi_1 - \varphi_2 \in C^2((\alpha, \beta) \times \mathbb{R}^n)$. Questa è ancora una soluzione del problema, con dati iniziali nulli ed ha supporto compatto quando ristretta ad ogni insieme $[\alpha', \beta'] \times \mathbb{R}^n$, con $[\alpha', \beta'] \subset (\alpha, \beta)$. Fissato $[\alpha', \beta'] \subset (\alpha, \beta)$, dato che il supporto di ϕ ristretta a tale insieme è compatto e quindi limitato, consideriamo una palla chiusa di raggio finito e centrata nell'origine, $B \subset \mathbb{R}^n$ in modo tale che $[\alpha', \beta'] \times B$ includa il supporto di ϕ. Consideriamo poi una seconda palla aperta di raggio finito e centrata nell'origine, $D \subset \mathbb{R}^n$ che includa la palla chiusa B. Per costruzione, per ogni $T \in [\alpha', \beta']$, la funzione $\mathbb{R}^n \ni \mathbf{x} \mapsto \phi(T, \mathbf{x})$ si annulla su $D \cup \partial D$, ma anche nella corona sferica aperta $D \setminus B$. Di conseguenza, si annulla con tutte le sue derivate (spaziali e temporali) fino a secondo ordine su ∂D. Lavorando come nella dimostrazione del teorema precedente abbiamo che, per ogni $T \in [\alpha', \beta']$:

$$\int_D E(T, \mathbf{x})\, d^n x = \int_0^T dt \oint_{+\partial D} \left(\frac{\partial \phi}{\partial t}\nabla\phi\right) \cdot \mathbf{n}\, dS(x) = 0\,,$$

dato che su $[0, T] \times \partial D$ la funzione ϕ e le sue derivate sono nulle. Ragionando come nel caso del teorema precedente si ha $\phi(T, \mathbf{x}) = 0$ per ogni $T \in (\alpha, \beta)$ e $\mathbf{x} \in \overline{D}$, ma qundi anche fuori da

$\overline{D}$ dato che fuori da tale insieme ϕ si annulla per ipotesi. Di conseguenza: $\varphi_1(t, \mathbf{x}) = \varphi_2(t, \mathbf{x})$ ovunque su $(\alpha, \beta) \times \mathbb{R}^n$. $\square$

Osservazioni 5.6. È interessante notare che, nelle ipotesi del teorema, scegliendo cioè D abbastanza grande in modo tale che ∂D non intersechi mai il supporto di $\varphi(t, \mathbf{x})$ per $t \in (\alpha', \beta')$, abbiamo che $\mathcal{E} := \int_D E(t, \mathbf{x})\, d^n x = \int_{\mathbb{R}^n} E(t, \mathbf{x})\, d^n x$. In questo modo abbiamo una nozione di *energia totale* associata al campo φ e tale energia è conservata nel tempo essendo, come è provato nel teorema $\frac{d\mathcal{E}}{dt} = 0$. Il valore di $\mathcal{E}$ dipende ovviamente dalla soluzione φ considerata.

Capitolo 6

Equazione di D'Alembert e di Klein-Gordon in $\mathbb{R} \times \mathbb{R}$.

In questo capitolo studieremo il problema dell'equazione di D'Alembert sul dominio spaziale dato da tutto $\mathbb{R}$ in assenza di condizioni al contorno. Successivamente discuteremo alcuni semplici risultati per l'equazione di D'Alembert e Klein-Gordon con un dominio spaziale dato da un segmento con l'aggiunta di condizioni al contorno, facendo uso di elementari teoremi della teoria della serie di Fourier.

6.1 Equazione di D'Alembert sulla retta reale senza condizioni al contorno.

Consideriamo l'equazione di D'Alembert in $\mathbb{R}^2$ e quindi con x che varia su tutta la retta reale. Benché si tratti di un caso molto particolare, è possibile in questo caso, scrivere esplicitamente la soluzione dell'equazione di D'Alembert. Inoltre molte delle proprietà di queste soluzioni hanno validità molto generale anche in dimensione maggiore ed in varietà ambiente (spazitempo) curve.

6.1.1 Assenza di sorgenti, formula di D'Alembert, domini di dipendenza.

Per prima cosa ci occupiamo del problema con soli dati iniziali ed in assenza di sorgenti (non ci sono dati al bordo in questo caso):

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\varphi}{\partial t^2} + \dfrac{\partial^2\varphi}{\partial x^2} = 0\,, \quad \varphi \in C^2(\mathbb{R}^2)\,, \\ \varphi(0,x) = \phi_0(x) \quad \forall x \in \mathbb{R}\,, \\ \dfrac{\partial\varphi}{\partial t}(0,x) = \phi_1(x) \quad \forall x \in \mathbb{R}\,, \end{cases} \tag{6.1}$$

dove $\phi_0 \in C^2(\mathbb{R})$ e $\phi_1 \in C^1(\mathbb{R})$ sono funzioni assegnate. Dimostreremo un teorema di esistenza ed unicità per il problema (6.1), dando esplicitamente l'espressione della soluzione in funzione

dei dati iniziali. Successivamente, in un'osservazione, mostreremo anche che il problema è ben posto nel senso di Hadamard.
Per risolvere l'equazione differenziale di D'Alembert:

$$-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}+\frac{\partial^2\varphi}{\partial x^2}=0\,, \tag{6.2}$$

facciamo il cambiamento di coordinate $v := (x-ct)/2$ e $w = (x+ct)/2$ che si inverte in $x = v+w$ e $t = (w-v)/c$ e pertanto definisce una funzione biettiva C^∞ da $\mathbb{R}^2$ in $\mathbb{R}^2$ con inversa C^∞. Con questa scelta risulta:

$$\frac{\partial}{\partial v}=\frac{\partial}{\partial x}-\frac{1}{c}\frac{\partial}{\partial t}\,,\quad \frac{\partial}{\partial v}=\frac{\partial}{\partial x}+\frac{1}{c}\frac{\partial}{\partial t}$$

e quindi si ha, per ogni funzione $\varphi \in C^2(\mathbb{R}^2)$:

$$\frac{\partial}{\partial v}\frac{\partial}{\partial w}\varphi(t(v,w),x(v,w))=\left(\frac{\partial}{\partial x}-\frac{1}{c}\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x}+\frac{1}{c}\frac{\partial}{\partial t}\right)\varphi(t,x)=\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\frac{\partial^2}{\partial x^2}\right)\varphi(t,x)\,.$$

Concludiamo che: $\varphi \in C^2(\mathbb{R}^2)$ risolve (6.2) se e solo se la funzione $C^2(\mathbb{R}^2)$, definita come $\psi(v,w) := \varphi(t(v,w),x(v,w))$, risolve

$$\frac{\partial^2\psi}{\partial v\partial w}=0\,. \tag{6.3}$$

Abbiamo allora il seguenti due lemmi.

Lemma 6.1. *Sia $\Omega \in \mathbb{R}^2$ un aperto connesso per segmenti paralleli all'asse x (cioè , per ogni coppia di punti in Ω il segmento parallelo all'asse x che li congiunge è tutti incluso in Ω). Se $\phi : \Omega \to \mathbb{R}$ è ovunque derivabile nella variabile x e soddisfa $\frac{\partial\phi}{\partial x} = 0$ ovunque su Ω, allora, su tutto Ω, $\psi(x,y) = F(y)$ per qualche funzione F della sola variabile y.* $\diamondsuit$

Dimostrazione. Si considerino due punti $(x,y),(x',y) \in \Omega$, con y fissato arbitrariamente. Il teorema di Lagrange per la funzione $s \mapsto \phi(s,y)$ può essere applicato sul segmento chiuso parallelo all'asse x che connette (x,y) e (x',y), dato che tale segmento è tutti incluso nel dominio della funzione ϕ e che la funzione $s \mapsto \phi(s,y)$ è derivabile su tale segmento per ipotesi. Si ottiene allora $\phi(x,y) - \phi(x',y) = (x-x')\frac{\partial\phi}{\partial x}|_{(\xi,y)} = 0$, cioè $\phi(x,y) = \phi(x',y)$. Indichiamo allora con $F(y)$ il valore comune che ψ assume sui punti in Ω appartenenti alla retta parallela all'asse x e tracciata alla generica quota y. Per costruzione, vale $\psi(x,y) = F(y)$ per ogni $(x,y) \in \Omega$. $\square$

Osservazioni 6.1. Il risultato è meno banale di quello che si potrebbe credere a prima vista, ed è per questo che lo abbiamo dimostrato esplicitamente. Infatti, se Ω non è connesso per segmenti paralleli all'asse x, la condizione $\frac{\partial\phi}{\partial x} = 0$ ovunque su Ω *non* assicura che si possa scrivere $\psi(x,y) = F(y)$ per qualche funzione F della sola variabile y! Si consideri infatti l'aperto $\Omega = \mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 \mid x = 0\,,\quad y \geq 0\}$ e su di esso la funzione $\phi = \phi(x,y)$ definita come segue.

(i) $\phi(x,y) = 0$ se $y < 0$,

(ii) $\phi(x,y) = 0$ se $x > 0$ e $y \geq 0$,

(iii) $\phi(x,y) = h(y)$ se $y \geq 0$ e $x < 0$, dove h è una qualsiasi (ma fissata) funzione $C^1([0,+\infty))$ che vale 0 per $y \in [0,1/3]$ e 1 per $y \in [2/3,+\infty)$.
La funzione ϕ costruita in questo modo è in $C^1(\Omega)$ e soddisfa $\frac{\partial \phi}{\partial x} = 0$ ovunque su Ω, ma non è possibile scrivere $\phi(x,y) = F(y)$ per qualche funzione F della sola variabile y: se ciò fosse possibile avremmo $1 = \phi(-1,1) = F(1) = \phi(1,1) = 0$.

Lemma 6.2. *La funzione $\varphi \in C^2(\mathbb{R}^2)$ risolve l'equazione (6.2) se e solo se è della forma*

$$\varphi(v,w) = f(x-ct) + g(x+ct)\,, \quad \textit{per ogni } (u,v) \in \mathbb{R}^2, \tag{6.4}$$

dove $f, g \in C^2(\mathbb{R})$. $\diamondsuit$

Dimostrazione. Per quanto detto prima dell'enunciato del lemma 6.1, definita la funzione in $C^2(\mathbb{R}^2)$ data da $\psi(v,w) := \varphi(t(v,w), x(v,w))$, è sufficiente dimostrare che le soluzioni di (6.3) sono tutte e sole della forma $\psi(v,w) = k(v) + h(w)$ dove $k, h \in C^2(\mathbb{R})$ e quindi definire $f(x-ct) := k((x-ct)/2)$ e $g(x+ct) := h((x+ct)/2)$. Dimostriamo quanto detto. Se $\psi \in C^2(\mathbb{R}^2)$ soddisfa la (6.3), poniamo $G(u,w) := \frac{\partial \psi}{\partial w}$. Valendo $\frac{\partial G(v,w)}{\partial v} = 0$, per $(v,w) \in \mathbb{R}^2$ che è sicuramente connesso per segmenti paralleli all'asse v, per il lemma 6.1 concludiamo che $\frac{\partial \psi}{\partial w} = F(w)$ per una certa funzione F. Tale funzione deve essere C^1, e quindi integrabile, dato che $\psi \in C^2(\mathbb{R}^2)$. Dato che $\mathbb{R}^2$ è anche connesso per segmenti paralleli all'asse w, possiamo allora scrivere, per v, w_0 fissati:

$$\int_{w_0}^{w} \frac{\partial \psi}{\partial w}(v,w')dw' = \int_{w_0}^{w} F(w')dw'\,.$$

Da cui:

$$\psi(v,w) = \psi(v,w_0) + \int_{w_0}^{w} F(w')dw'\,,$$

che possiamo riscrivere:

$$\psi(v,w) = k(v) + h(w)\,,$$

dove $k(v) := \psi(v,w_0)$ e $h(w) := \int_{w_0}^{w} F(w')dw'$. Le funzioni k e h risultano essere funzioni $C^2(\mathbb{R})$ per costruzione. Viceversa, se $\psi(v,w) = k(v) + h(w)$ per ogni $(u,v) \in \mathbb{R}^2$ con $k, h \in C^2(\mathbb{R})$, allora $\psi \in C^2(\mathbb{R}^2)$ e risolve (6.3), come si verifica immediatamente. $\square$

Dato che ora abbiamo la classe completa delle soluzioni dell'equazione (6.2), non ci resta che verificare se esistano, in tale classe, delle soluzioni che soddisfino anche le condizioni inziali del problema (6.1). Arriviamo in tal modo al seguente teorema di esistenza ed unicità di D'Alembert.

Teorema 6.1. *Esiste ed è unica la soluzione φ del problema (6.1) per ogni scelta delle condizioni iniziali $\phi_0 \in C^2(\mathbb{R})$ e $\phi_1 \in C^1(\mathbb{R})$. Tale soluzione si esprime tramite la* **formula di D'Alembert***:*

$$\varphi(t,x) = \frac{1}{2}\left[\phi_0(x-ct) + \phi_0(x+ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} \phi_1(\xi)\, d\xi\,. \tag{6.5}$$

$\diamondsuit$

Dimostrazione. Sappiamo dal lemma 6.2 che, se esiste, la soluzione deve avere forma $\varphi(t,x) = f(x-ct)+g(x+ct)$, dove $f,g \in C^2(\mathbb{R})$. Vogliamo determinare f e g in funzione delle condizioni iniziali. Per $t=0$ deve allora risultare $\phi_0(x) = f(x)+g(x)$ e quindi $\phi_0'(x) = f'(x)+g'(x)$. Dato che vale anche $\phi_1(x) = -cf'(x)+cg'(x)$, ricaviamo subito: $f'(x) = \frac{1}{2c}(c\phi_0'(x)-\phi_1(x))$ e $g'(x) = \frac{1}{2c}(c\phi_0'(x)+\phi_1(x))$. Possiamo integrare queste espressioni ottenendo, se a,b sono costati reali,

$$f(x) = a + \frac{1}{2}\phi_0(x) - \frac{1}{2c}\int_0^x \phi_1(\xi)d\xi\,, \quad g(x) = b + \frac{1}{2}\phi_0(x) + \frac{1}{2c}\int_0^x \phi_1(\xi)d\xi\,.$$

Di conseguenza, se esiste una soluzione al problema è nella classe di funzioni, parametrizzata dalle costanti $A \in \mathbb{R}$:

$$\varphi(t,x) = A + \frac{1}{2}\phi_0(x-ct) - \frac{1}{2c}\int_0^{x-ct} \phi_1(\xi)d\xi + \frac{1}{2}\phi_0(x+ct) + \frac{1}{2c}\int_0^{x+ct} \phi_1(\xi)d\xi\,.$$

Si osservi ogni funzione di tale classe è C^2 per costruzione e soddisfa necessariamente (6.2) per ogni scelta di $A \in \mathbb{R}$, dato che è proprio della forma richiesta nel lemma 6.2. La prima condizione iniziale è soddisfatta solo se $A=0$, valendo $\varphi(0,x) = A+\phi_0(x)$, e la seconda condizione iniziale è sempre soddisfatta, valendo: $\frac{\partial\varphi}{\partial t}(0,x) = \phi_1(x)$. In definitiva l'unica soluzione al problema (6.1) è la funzione della classe di sopra con $A=0$. Possiamo riscrivere la soluzione come:

$$\varphi(t,x) = \frac{1}{2}\phi_0(x-ct) + \frac{1}{2c}\int_{x-ct}^0 \phi_1(\xi)d\xi + \frac{1}{2}\phi_0(x+ct) + \frac{1}{2c}\int_0^{x+ct} \phi_1(\xi)d\xi\,,$$

e quindi:

$$\varphi(t,x) = \frac{1}{2}\left[\phi_0(x-ct)+\phi_0(x+ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} \phi_1(\xi)\,d\xi\,.$$

$\square$

Osservazioni 6.2.
(1) La forma generale della soluzione dell'equazione di D'Alembert ha comunque una struttura della forma:

$$\varphi(t,x) = f(x-ct)+g(x+ct)\,.$$

Il primo addendo a secondo membro rappresenta un profilo d'onda che procede da sinistra verso destra traslando senza deformarsi, alla velocità c (infatti, in un intervallo di tempo Δt, il profilo trasla di un intervallo di spazio $\Delta x = c\Delta t$). Questo tipo di onda è detta **onda progressiva**. Il primo addendo a secondo membro rappresenta un profilo d'onda che procede da destra verso sinistra traslando senza deformarsi, alla velocità c. Questo tipo di onda è detta **onda regressiva**. In questo senso la costante c che appare nell'equazione di D'Alembert rappresenta la velocità di propagazione delle perturbazioni soluzioni dell'equazione. In dimensione spaziale maggiore di 1, la situazione è analoga, ma si assiste anche ad una deformazione del profilo della perturbazione;

in nogni caso si riesce a provare che la costante c ha ancora lo stesso significato fisico, dopo avere introdotto la nozione di *velocità di fase*, della quale qui non ci occuperemo.
(2) Consideriamo il problema (6.1) e la sua soluzione espressa dalla formula di D'alembert (6.5). Se $(a, b) \subset \mathbb{R}$ è limitato, si definisce in $\mathbb{R}^2$ il **dominio di dipendenza futuro** $D_+(a, b)$ come l'insieme chiuso dato dal triangolo di base $[a, b]$ sull'asse $t = 0$ e vertice nel semipiano $t > 0$ individuato dall'intersezione delle due rette che partono da a e b rispettivamente ed hanno inclinazione $1/c$ e $-1/c$ rispettivamente. Tale vertice ha coordinate $x_+ = (a+b)/2$ e $t_+ = (b-a)/(2c)$. Si definisce analogamente il **dominio di dipendenza passato** $D_-(a, b)$ come l'insieme chiuso dato dal triangolo di base $[a, b]$ sull'asse $t = 0$ e vertice nel semipiano $t < 0$ individuato dall'intersezione delle due rette che partono da a e b rispettivamente ed hanno inclinazione $-1/c$ e $1/c$ rispettivamente. Tale vertice ha coordinate $x_- = (a + b)/2$ e $t_- = -(b - a)/(2c)$. Il **dominio di dipendenza** $D(a, b)$ è , per definizione l'unione di $D_+(a, b)$ e $D_-(a, b)$. Si osservi che le rette di inclinazione $\pm 1/c$, che individuano il bordo di $D(a, b)$, sono *rette caratteristiche* per l'equazione di D'Alembert.
Se si considera un punto $(t_0, x_0) \in D_+(a, b)$, la formula di D'alembert per in campo φ valutato in (t_0, x_0), mostra che il valore $\varphi(t_0, x_0)$ dipende solo dal valore di ϕ_0 e ϕ_1 in $[a, b]$. Più precisamente, i valori rilevanti di ϕ_0 e ϕ_1 sono quelli che cadono nel sottointervallo $[x_0 - ct_0, x_0 + ct_0] \subset [a, b]$. Tale sottointervallo si ottiene intersecando con l'asse $t = 0$ le due rette caratteristiche emanate, verso il passato, da (t_0, x_0). Un discorso analogo si può fare per i punto in $D_-(a, b)$.
La formula di D'Alembert implica quindi che, all'interno di $D(a, b)$, la funzione φ sia completamente determinata dalle due condizioni iniziali ristrette ad $[a, b]$, nel senso che, se alteriamo tali condizioni iniziali *fuori* da $[a, b]$, la soluzione φ non risulta essere alterata dentro $D(a, b)$.
L'esistenza di domini di dipendenza con le proprietà dette è comune alla teoria di tutte le equazioni differenziali a derivate parziali del secondo ordine di tipo iperbolico su varietà differenziabili Lorentziane, cioè su spazitempo (generalmente curvi), quando la forma quadratica dell'equazione è data dalla stessa metrica dello spaziotempo. Si tratta di uno dei punti di partenza per sviluppare la *teoria della causalità* in teoria dei campi in ambiente relativistico generale.
(3) La formula di D'Alembert implica che il problema iperbolico (6.1) sia ben posto nel senso di Hadamard. Sappiamo già che la soluzione esiste ed è unica, dobbiamo quindi studiare la dipendenza continua dai dati iniziali. L'ambiente naturale in cui studiare questo problema è un dominio di dipendenza. Consideriamo due set di condizioni iniziali ϕ_0, ϕ_1 e $\widetilde{\phi}_0, \widetilde{\phi}_1$, indichiamo con φ e $\widetilde{\varphi}$ le corrispondenti soluzioni dell'equazione di D'Alembert, fissiamo un intervallo $[a, b] \subset \mathbb{R}$ e l'associato dominio di dipendenza $D(a, b)$. Dalla formula di D'Alembert segue immediatamente che, se $(t, x) \in D(a, b)$

$$|\varphi(t,x)-\widetilde{\varphi(t,x)}| \leq \frac{1}{2}\sup_{\xi\in[a,b]}|\phi_0(\xi)-\widetilde{\phi}_0(\xi)|+\frac{1}{2}\sup_{\xi\in[a,b]}|\phi_0(\xi)-\widetilde{\phi}_0(\xi)|+\frac{1}{2c}\sup_{\xi\in[a,b]}|\phi_1(\xi)-\widetilde{\phi}_1(\xi)|\int_{x-ct}^{x+ct}d\xi.$$

L'ultimo integrale vale t e quindi è maggiorato da $T_{[a,b]}$ pari all'altezza del triangolo $D_+(a, b)$. In definitva abbiamo trovato che, se $|| \cdot ||_{\infty\, A}$ indica la norma dell'estremo superiore calcolata restringendo il dominio delle funzioni all'insieme A,

$$||\varphi - \widetilde{\varphi}||_{\infty\, D(a,b)} \leq ||\phi_0 - \widetilde{\phi}_0||_{\infty\, [a,b]} + T_{[a,b]}||\phi_1 - \widetilde{\phi}_1||_{\infty\, [a,b]}\,. \tag{6.6}$$

Se deriviamo entrambi i membri della formula di D'Alembert nella variabile t otteniamo che

$$\partial_t \varphi(t,x) = \frac{c}{2}\left[-\phi_0'(x-ct) + \phi_0'(x+ct)\right] + \frac{1}{2}(\phi_1(x+ct) + \phi_1(x-ct)) \ .$$

In conseguenza di quanto trovato abbiamo che:

$$\left|\left|\frac{\partial \varphi}{\partial t} - \frac{\partial \widetilde{\varphi}}{\partial t}\right|\right|_{\infty\, D(a,b)} \leq c||\phi_0' - \widetilde{\phi}_0'||_{\infty\,[a,b]} + ||\phi_1 - \widetilde{\phi}_1||_{\infty\,[a,b]} \ .$$

In modo analogo abbiamo anche che:

$$\left|\left|\frac{\partial \varphi}{\partial x} - \frac{\partial \widetilde{\varphi}}{\partial x}\right|\right|_{\infty\, D(a,b)} \leq ||\phi_0' - \widetilde{\phi}_0'||_{\infty\,[a,b]} + \frac{1}{c}||\phi_1 - \widetilde{\phi}_1||_{\infty\,[a,b]} \ .$$

Valgono, e si ottengono con la stessa procedura, delle disuguaglianze per le derivate seconde:

$$\left|\left|\frac{\partial^2 \varphi}{\partial x^2} - \frac{\partial^2 \widetilde{\varphi}}{\partial x^2}\right|\right|_{\infty\, D(a,b)} \leq ||\phi_0'' - \widetilde{\phi}_0''||_{\infty\,[a,b]} + \frac{1}{c^2}||\phi_1' - \widetilde{\phi}_1'||_{\infty\,[a,b]} \ ,$$

$$\left|\left|\frac{\partial^2 \varphi}{\partial t^2} - \frac{\partial^2 \widetilde{\varphi}}{\partial t^2}\right|\right|_{\infty\, D(a,b)} \leq c^2||\phi_0'' - \widetilde{\phi}_0''||_{\infty\,[a,b]} + c||\phi_1' - \widetilde{\phi}_1'||_{\infty\,[a,b]} \ ,$$

$$\left|\left|\frac{\partial^2 \varphi}{\partial t \partial x} - \frac{\partial^2 \widetilde{\varphi}}{\partial t \partial x}\right|\right|_{\infty\, D(a,b)} \leq c||\phi_0'' - \widetilde{\phi}_0''||_{\infty\,[a,b]} + ||\phi_1' - \widetilde{\phi}_1'||_{\infty\,[a,b]} \ .$$

Queste relazioni mostrano come, prendendo condizioni iniziali vicine fino ad un certo ordine di differenziabilità , si ottengono soluzioni vicine fino all'ordine di differenziabilità considerato. Questo è proprio il senso della dipendenza continua dai dati iniziali proposta da Hadamard. Questa proprietà si generalizza a equazioni differenziali di tipo imperbolico in dimensione ed ambienti molto più generali.

(4) La formula di D'Alembert definisce una funzione φ su $\mathbb{R}^2$ anche se le due funzioni ϕ_0 e ϕ_1 non sono C^2 in qualche punto isolato di $\mathbb{R}$ attorno al quale ϕ_1 sia comunque integrabile. Perché esista φ definita dal secondo membro della formula di D'Alembert è , a rigore, sufficiente che ϕ_1 sia integrabile. Si vede facilmente che se x_0 è uno dei punti isolati di singolarità di ϕ_0 o ϕ_1, il secondo membro della formula di D'Alembert è una funzione ovunque C^2 che soddisfa l'equazione di D'alembert e le condizioni iniziale, eccetto che sulle rette caratteristiche che escono dal punto $(0, x_0)$ (e sulle rette analoghe che escono dagli altri punti isolati di singolarità). In questo senso *le singolarità delle condizioni iniziali si propagano lungo le curve caratteristiche.* Questo fatto è piuttosto generale e vale per equazioni differanziali di tipo iperbolico in dimensione ed ambienti molto più generali.

L'osservazione (3) di sopra ha un'importante conseguenza precedentemente preannunciata.
Dato $[a,b] \subset \mathbb{R}$, pensato come retta a $t = 0$ in $\mathbb{R} \times \mathbb{R}$, definiamo lo **sviluppo causale** di $[a,b]$, indicato con $J(a,b) \subset \mathbb{R}$, come l'insieme dei punti di $\mathbb{R} \times \mathbb{R}$ che possono essere raggiunti da una

retta di pendenza $\leq -1/c$ oppure $\geq 1/c$ emanata da $[a,b]$.
$J(a,b)$ risulta essere l'unione dei due coni infiniti, uno di vertice con coordinate $x_+ = (a+b)/2$ (cioè il punto medio di (a,b)) e $t_+ = (b-a)/(2c)$, emanato verso il passato, e l'altro di vertice di coordinate $x_- = x_+$ e $t_- = -t_+$ emanato verso il futuro. Si osservi ancora che $\mathbb{R}^2 \setminus J(a,b)$ è l'unione di tutti i domini di dipendenza $D(c,d)$ con $c > b$ oppure $d < a$.

Teorema 6.2. *Se nel problema (6.1) le condizioni iniziali sono scelte a supporto compatto: $\phi_0 \in C^2_0(\mathbb{R})$ e $\phi_1 \in C^1_0(\mathbb{R})$, e $[a,b] \supset supp\phi_0 \cup supp\phi_1$, allora la soluzione φ del problema è nulla fuori da $J(a,b)$. Di conseguenza, per ogni fissato $[\alpha,\beta]$ con $[\alpha,\beta] \ni 0$:*
(a) *il supporto della soluzione φ ristretta a $[\alpha,\beta] \times \mathbb{R}$ è compatto;*
(b) *per ogni fissato $t \in [\alpha,\beta]$, il supporto di $\mathbb{R} \ni \mathbf{x} \mapsto \varphi(t,\mathbf{x})$ è compatto in $\mathbb{R}$.* $\square$

Dimostrazione. Dalla (6.6), scegliendo $\widetilde{\varphi}$ come la funzione ovunque nulla (che quindi risolve il problema con dati iniziali ovunque nulli), troviamo:

$$||\varphi||_{\infty\, D(a',b')} \leq ||\phi_0||_{\infty\, [a',b']} + T_{[a',b']}||\phi_1||_{\infty\, [a',b']}\,.$$

Fissiamo ora un qualsiasi punto (t_0,x_0) fuori da $J(a,b)$ con $t_0 \geq 0$. Per definizione di $J(a,b)$, se $a' = x_0 - ct_0$ e $b' = x_0 + ct_0$, allora $[a',b']$ non interseca mai $[a,b]$. Dato che $[a,b]$ contiene i supporti di ϕ_0 e ϕ_1, tali funzioni sono nulle in $[a',b']$. Concludiamo che

$$0 \leq ||\varphi||_{\infty\, D(a',b')} \leq ||\phi_0||_{\infty\, [a',b']} + T_{[a',b']}||\phi_1||_{\infty\, [a',b']} = 0 + 0 = 0$$

e quindi, in particolare, dato che $(t_0,x_0) \in D(a',b')$, $\varphi(t_0,x_0) = 0$. Fissiamo infine il compatto $[\alpha,\beta] \subset \mathbb{R}$ con $[\alpha,\beta] \ni 0$. Tenendo conto della forma di $J(a,b)$ che è dato dall'unione di due coni come precisato sopra, segue subito che il supporto di φ ristretta alla regione chiusa tra le due rette $t = \alpha$ e $t = \beta$, è contenuto nell'unione dei due trapezi chiusi di base minore in comune data da $[a,b]$ e basi maggiori individuate dalla porzione delle rette $t = \alpha$ e $t = \beta$ che cadono in $J(a,b)$. Tale insieme è evidentemente limitato, pertanto il supporto di φ ristretta alla regione chiusa tra le due rette $t = \alpha$ e $t = \beta$, che è un insieme chiuso per definizione, è anch'esso compatto. Il supporto di $\mathbb{R} \ni \mathbf{x} \mapsto \varphi(t,\mathbf{x})$, con $t \in [\alpha,\beta]$, è un chiuso sottoinsieme di un compatto ed è pertanto anch'esso compatto. $\square$

Osservazioni 6.3. Come già osservato precedentemente, ma ora possiamo essere più chiari, le due proprietà (a) e (b) sono valide anche in $\mathbb{R} \times \mathbb{R}^n$ per le soluzioni dell'equazione di Klein-Gordon e d'Alembert quando i dati iniziali hanno supporto compatto (e tale fatto vale in maniera molto generale per soluzioni di equazioni iperboliche su uno spaziotempo curvo con la proprietà della "globale iperbolicità"). Nel caso generale però, la dimostrazione di (a) e (b) è molto più complicata.

6.1.2 Equazione di D'Alembert con sorgente.

Consideriamo ora il problema con sorgente, data dalla funzione f:

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\varphi}{\partial t^2} + \dfrac{\partial^2\varphi}{\partial x^2} = f(t,x)\,, \quad \varphi \in C^2(\mathbb{R}^2)\,, \\ \varphi(0,x) = \phi_0(x) \quad \forall x \in \mathbb{R}\,, \\ \dfrac{\partial\varphi}{\partial t}(0,x) = \phi_1(x) \quad \forall x \in \mathbb{R}\,, \end{cases} \tag{6.7}$$

dove $\phi_0 \in C^2(\mathbb{R})$ e $\phi_1 \in C^1(\mathbb{R})$, $f \in C^0(\mathbb{R}^2)$ sono funzioni assegnate.

Abbiamo un primo risultato, abbastanza semplice, che riguarda l'unicità della soluzione.

Teorema 6.3. *Se esiste una soluzione al problema (6.7) con fissati dati $\phi_0 \in C^2(\mathbb{R})$ e $\phi_1 \in C^1(\mathbb{R})$, $f \in C^0(\mathbb{R}^2)$, allora è unica.* $\diamondsuit$

Dimostrazione. Se φ_1 e φ_2 risolvono il problema (6.7) con gli stessi dati $\phi_0 \in C^2(\mathbb{R})$ e $\phi_1 \in C^1(\mathbb{R})$, $f \in C^0(\mathbb{R}^2)$ allora $\varphi := \varphi_1 - \varphi_2$ risolve il problema (6.1) con condizioni iniziali nulle. In base al teorema 6.1 deve essere $\varphi(t,x) = 0$ ovunque, dato che la soluzione ovunque nulla risolve il problema posto ed è l'unica a farlo. Pertanto $\varphi_1(t,x) = \varphi_2(t,x)$ per ogni $(t,x) \in \mathbb{R}^2$. $\square$

Passiamo a dimostrare un teorema di esistenza nel caso in cui $f \in C^1(\mathbb{R}^2)$. Possiamo decomporre la funzione φ in due parti $\varphi = \phi + \Phi$, in cui ϕ soddisfa il problema omogeneo

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\phi}{\partial t^2} + \dfrac{\partial^2\phi}{\partial x^2} = 0\,, \quad \varphi \in C^2(\mathbb{R}^2)\,, \\ \varphi(0,x) = \phi_0(x) \quad \forall x \in \mathbb{R}\,, \\ \dfrac{\partial\phi}{\partial t}(0,x) = \phi_1(x) \quad \forall x \in \mathbb{R}\,, \end{cases} \tag{6.8}$$

mentre Φ soddisfa il problema con sorgente, ma con dati iniziali nulli

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\Phi}{\partial t^2} + \dfrac{\partial^2\Phi}{\partial x^2} = f(t,x)\,, \quad \varphi \in C^2(\mathbb{R}^2)\,, \\ \Phi(0,x) = 0 \quad \forall x \in \mathbb{R}\,, \\ \dfrac{\partial\Phi}{\partial t}(0,x) = 0 \quad \forall x \in \mathbb{R}\,. \end{cases} \tag{6.9}$$

Dovrebbe essere ovvio che $\varphi = \phi + \Phi$ soddisfa (6.7) se le due funzioni hanno le proprietà richieste. La funzione ϕ esiste sicuramente in base al teorema 6.1. Mostriamo ora che esiste anche una funzione Φ che risolve (6.9). Consideriamo infatti:

$$\Phi(t,x) := -\frac{c}{2}\int_0^t dt \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\xi\, f(\tau,\xi)\,. \tag{6.10}$$

Si osservi che l'integrale può essere riscritto (anche se $t \leq 0$) come l'integrale doppio di Riemann (o Lebesgue)

$$\Phi(t,x) := -\frac{c}{2}\int_{A(t,x)} f(\tau,\xi)\, d\tau d\xi \ ,$$

dove il dominio d'integrazione $A(t,x)$ nel piano (τ,ξ) è un compatto essendo dato dal triangolo di base $[x-ct, x+ct]$ sull'asse $\tau = 0$ e vertice (x,t), e quindi la funzione continua f è dunque integrabile su tale dominio. Si noti che quindi $A(t,x) = D_+(x-ct, x+ct)$ se $t \geq 0$, oppure $A(t,x) = D_-(x-ct, x+ct)$ se $t \leq 0$.
Vale, applicando i soliti teoremi di passaggio della derivata sotto il segno di integrale ove necessario,

$$-\partial_t\left(-\frac{c}{2}\int_0^t d\tau \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\xi f(\tau,\xi)\right) = \frac{c}{2}\int_{x-c(t-t)}^{x+c(t-t)} d\xi f(\tau,\xi) + \frac{c}{2}\int_0^t dt\partial_t \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\xi f(\tau,\xi) \ ,$$

e quindi

$$-\partial_t\Phi(t,x) = 0 + \frac{c^2}{2}\int_0^t d\tau\, [f(\tau, x+c(t-\tau)) + f(\tau, x-c(t-\tau))] \ . \tag{6.11}$$

Passando alla derivata seconda, ed usando esplicitamente il fatto che $f \in C^1(\mathbb{R}^2)$,

$$-\frac{2}{c^2}\partial_t^2\Phi(t,x) = \partial_t \int_0^t d\tau\, [f(\tau, x+c(t-\tau)) + f(\tau, x-c(t-\tau))]$$

$$= [f(t, x+c(t-t)) + f(t, x-c(t-t))] + c\int_0^t d\tau\, [\partial_x f(\tau, x+c(t-\tau)) - \partial_x f(\tau, x-c(t-\tau))] \ .$$

Abbiamo ottenuto:

$$-\frac{1}{c^2}\partial_t^2\Phi(t,x) = f(t,x) + \frac{c}{2}\int_0^t d\tau\, [\partial_x f(\tau, x+c(t-\tau)) - \partial_x f(\tau, x-c(t-\tau))] \ . \tag{6.12}$$

Similmente:

$$\partial_x\Phi(t,x) = -\partial_x\frac{c}{2}\int_0^t d\tau \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\xi f(\tau,\xi) = -\frac{c}{2}\int_0^t d\tau\, [f(\tau, x+c(t-\tau)) - f(\tau, x-c(t-\tau))] \ ,$$

e quindi:

$$\partial_x^2\Phi(t,x) = -\frac{c}{2}\int_0^t d\tau\, [\partial_x f(\tau, x+c(t-\tau)) - \partial_x f(\tau, x-c(t-\tau))] \ . \tag{6.13}$$

Sommando membro a membro (6.12) e (6.13) otteniamo:

$$-\frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2} + \frac{\partial^2\Phi}{\partial x^2} = f \ .$$

Da (6.10) e (6.11) abbiamo immediatamente che Φ soddisfa anche le condizioni iniziali nulle del problema (6.9) come volevamo. Si osservi che la funzione Φ risulta essere $C^2(\mathbb{R}^2)$, le derivate

seconde in x e t sono state calcolate sopra e si prova facilmente che sono continue, le derivate miste si calcolano analogamente e forniscono:

$$\partial_x \partial_t \Phi(t,x) = \partial_t \partial_x \Phi(t,x) = -\frac{c^2}{2} \int_0^t d\tau \left[\partial_x f(\tau, x + c(t-\tau)) + \partial_x f(\tau, x - c(t-\tau)) \right] ,$$

che è una funzione continua se $f \in C^1(\mathbb{R}^2)$.
Abbiamo provato il seguente teorema.

Teorema 6.4. *Si consideri il problema (6.7) dove $\phi_0 \in C^2(\mathbb{R})$ e $\phi_1 \in C^1(\mathbb{R})$, $f \in C^1(\mathbb{R}^2)$ sono funzioni assegnate. Esiste ed è unica la soluzione φ di tale problema e si esprime come:*

$$\varphi(t,x) = \frac{1}{2} \left[\phi_0(x-ct) + \phi_0(x+ct) \right] + \frac{1}{2c} \int_{x-ct}^{x+ct} \phi_1(\xi)\, d\xi - \frac{c}{2} \int_0^t dt \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\xi f(\tau, \xi) .$$

□

Osservazioni 6.4. La formula (6.10), come già osservato può essere scritta come

$$\Phi(t,x) := -\frac{c}{2} \int_{A(t,x)} f(\tau, \xi)\, d\tau d\xi ,$$

dove $A(t,x)$ è un dominio compatto dato da $D_+(x-ct, x+ct)$ se $t \geq 0$, oppure $D_-(x-ct, x+ct)$ se $t \leq 0$. Possiamo riscrivere la formula che determina Φ come

$$\Phi(t,x) := \int_{\mathbb{R}^2} G(t,x|\tau,\xi) f(\tau,\xi)\, d\tau d\xi ,$$

dove $G(t,x|\tau,\xi)$ non è altro che la funzione caratteristica, nel piano (τ,ξ), di $D_+(x-ct, x+ct)$ se $t \geq 0$, oppure $D_-(x-ct, x+ct)$ se $t \leq 0$, moltiplicata per il fattore $-(c/2)sign(t)$. Tale funzione (in realtà è più propriamente pensabile come una funzione generalizzata o distribuzione) si chiama **funzione di Green** (con condizioni di annullamento sulla superficie $t = 0$) dell'operatore di D'Alembert su $\mathbb{R}^2$:

$$\Box := -\frac{1}{c^2} \partial_t^2 + \partial_x^2 .$$

Le funzioni di Green per le equazioni iperboliche possono essere definite (con vari dati iniziali) anche in dimensione maggiore ed in ambienti più generali. Esse giocano un ruolo importante negli sviluppi della teoria specie nelle teorie relativistiche (come dimostrato da Riesz, Hadamard e Leray, Hörmander).

6.2 Dalla separazione delle variabili alla serie di Fourier.

Consideriamo ora il problema di dover risolvere l'equazione di D'Alembert senza sorgente per la funzione $\varphi = \varphi(t,x)$ quando il dominio spaziale è un intervallo $[-L/2, L/2] \subset \mathbb{R}$ e $t \in (\alpha, \beta) \ni 0$,

nella situazione in cui, oltre a condizioni iniziali a $t = 0$, sono imposte *condizioni al contorno di periodicità* :

$$\varphi(t, -L/2) = \varphi(t, L/2), \quad \frac{\partial \varphi}{\partial x}(t, -L/2) = \frac{\partial \varphi}{\partial x}(t, L/2), \quad \frac{\partial^2 \varphi}{\partial x^2}(t, -L/2) = \frac{\partial^2 \varphi}{\partial x^2}(t, L/2) \quad \forall t \in (\alpha, \beta) \tag{6.14}$$

La terza condizione segue dalla prima e dall'equazione differenziale stessa. Dato che l'equazione è :

$$\frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = \frac{\partial^2 \varphi}{\partial x^2} ,$$

possiamo tentare di risolverla, con la procedura detta di *separazione delle variabili*, assumendo una forma particolare delle soluzioni del tipo

$$\varphi(t, x) = f(t)g(x) .$$

Inserendo nell'equazione di sopra si arriva subito all'identità , che vale quando le funzioni f e g non si annullano,

$$\frac{1}{c^2 f(t)}\frac{\partial^2 f(t)}{\partial t^2} = \frac{1}{g(x)}\frac{\partial^2 g(x)}{\partial x^2} .$$

Dato che i due membri dell'identità ottenuta sono funzione di due variabili diverse, i due membri devono essere funzioni costanti separatamente. Otteniamo in tal modo le due equazioni, per qualche costante $E \in \mathbb{R}$:

$$\frac{d^2 f(t)}{dt^2} = c^2 E f(t) , \quad \frac{d^2 g(x)}{dx^2} = E g(x) .$$

La seconda equazione fornisce la classe di soluzioni

$$g_E(x) := C_+(E)e^{\sqrt{E}x} + C_-(E)e^{-\sqrt{E}x} . \tag{6.15}$$

Tuttavia dobbiamo ancora imporre la condizione di periodicità su $\varphi(t, x) = f(t)g(x)$ che, nel caso in esame è equivalente alla richiesta che la funzione g_E soddisfi

$$g_E(-L/2) = g_E(L/2) .$$

Se $E > 0$ in (6.15), la condizione scritta sopra non è mai soddisfatta. Nel caso in cui $E \leq 0$, gli esponenti diventano complessi:

$$\pm i\sqrt{-E}x , \quad x \in [-L/2, L/2] ,$$

e pertanto le funzioni g_E sono periodiche. Affinché risultino essere periodiche sul segmento di lunghezza L (non importa quali siano i suoi estremi, ciò vale per $[-L/2, L/2]$ come per $[0, L]$ o altro), è necessario e sufficiente che $\sqrt{-E}L/(2\pi)$ sia un numero naturale. Quindi deve essere $E = -(2\pi n/L)^2$ con $n \in \mathbb{N}$ arbitrario. In questo modo si trova subito che, etichettando le funzioni g_E con l'indice $n \in \mathbb{N}$ invece che E, esse possono solo essere del tipo:

$$g_0(x) := C_0 , \quad g_n(x) := C_+(n)e^{i\frac{2\pi n}{L}x} + C_-(n)e^{-i\frac{2\pi n}{L}x} .$$

Per tenere conto dei due segni degli esponenti è conveniente usare un unico esponenziale e fare variare n in $\mathbb{Z}$ invece che in $\mathbb{N}$. Abbiamo allora che le funzioni g_n ammissibili, hanno tutte la forma:

$$g_0(x) := C_0 \,, \quad g_n(x) := C_n e^{i\frac{2\pi n}{L}x} \,, \quad n \in \mathbb{Z} \setminus \{0\} \,. \tag{6.16}$$

L'equazione per la funzione f_E, che ora indicheremo con f_n, è ora:

$$\frac{d^2 f_n(t)}{dt^2} = -c^2 \left(\frac{2\pi n}{L}\right)^2 f_n(t) \,,$$

che ha come risultato la classe di soluzioni:

$$f_0(t) := D_0' t + D_0 \,, \quad f_n(t) := D_n e^{i\frac{2\pi n}{L}ct} \,, \quad n \in \mathbb{Z} \setminus \{0\} \,. \tag{6.17}$$

Ognuna delle funzioni, con $A_0, A_0', A_n^{(\pm)} \in \mathbb{C}$:

$$\varphi_0(t) = A_0 t + A_0 \,, \quad \varphi_n(t,x)_\pm := A_n^{(\pm)} e^{i\frac{2\pi n}{L}x} e^{\pm i\frac{2\pi n}{L}ct} \,, \quad n \in \mathbb{Z} \setminus \{0\} \,, \tag{6.18}$$

è una possibile soluzione del'equazione di D'alembert in $\mathbb{R} \times [-L/2, L/2]$ con condizioni al contorno periodiche sul segmento $[-L/2, L/2]$. Anche se queste soluzioni sono complesse, possiamo sempre ridurci al caso reale prendendo delle combinazioni lineari di esse con coefficienti opportuni, ricordando che $e^{i\theta} + e^{-i\theta} = 2\cos\theta$ e $-i(e^{i\theta} - e^{-i\theta}) = 2\sin\theta$. Dato che stiamo lavorando con un'equazione differenziale lineare omogenea, combinazioni lineari di soluzioni saranno ancora soluzioni. Quest'ultima osservazione potrebbe essere utile anche per cercare di soddisfare le condizioni iniziali, cioè la forma che φ e la sua derivata temporale devono assumere all'istante $t = 0$. Tuttavia, è intuitivo pensare che se le condizioni iniziali sono assegnate in termini di funzioni arbitrarie, non sarà possibile trovare una combinazione linare *finita* di soluzioni della forma (6.18) che soddisfi anche tali condizioni iniziali. Si può pensare che ciò sia invece possibile considerando anche combinazioni linari *infinite*. Questa idea è quella che ha condotto Fourier ad inventare la teoria della serie omonima (lavorando però con un'equazione differente – ma con analoghe caratteristiche per quanto riguarda l'applicazione della teoria della serie di Fourier – l'equazione del calore). L'idea fondamentale è quella di sviluppare le funzioni periodiche f definite su un intervallo $[-L/2, L/2] \subset \mathbb{R}$ (ma l'approccio si generalizza su varietà toroidali compatte k-dimensionali) con una serie di funzioni i cui termini siano funzioni esponenziali $e^{i\frac{2\pi n}{L}x}$ con opportuni coefficienti complessi e con $n \in \mathbb{Z}$:

$$f(x) = \sum_{n\in\mathbb{Z}} C_n e^{i\frac{2\pi n}{L}x} \,.$$

Nel caso della nostra funzione φ soluzione periodica dell'equazione di D'Alembert, ci si aspetta che essa abbia una forma, che assicura automaticamente la periodicità in x di φ:

$$\varphi(t,x) = \sum_{n\in\mathbb{Z}} C_n(t) e^{i\frac{2\pi n}{L}x} \,.$$

La dipendenza temporale di φ (e quindi il fatto che φ soddisfi l'equazione di D'Alembert) si scarica tutta nei coefficienti complessi $C_n(t)$. Ci aspettiamo, da quanto visto sopra, che la forma di tali coefficienti sia proprio una combinazione lineare finita di funzioni di t del tipo di quelle in (6.17). Le infinite costanti arbitrare che appaiono in tutte queste combinazioni lineari dovranno anche essere fissate in modo tale da soddisfare le condizioni iniziali. Dopo aver enunciato alcuni risultati ben noti della teroria torneremo all'equazione di D'Alembert e di Klein-Gordon per vedere come si conclude il discorso cominciato sopra. proptrick

6.3 Alcuni risultati elementari sulla serie di Fourier.

Richiamiamo qui alcuni semplici risultati della teoria della serie di Fourier. Tutti questi argomenti saranno approfonditi in corsi avanzati di analisi.
Supponiamo che una funzione $f : [-L/2, L/2] \to \mathbb{C}$ si possa sviluppare in serie di Fourier, per il momento lavorando del tutto formalmente senza farci domande sul tipo di convergenza:

$$f(x) = \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \ . \tag{6.19}$$

Abbiamo introdotto il fattore $1/\sqrt{L}$ per pura convenienza. Vogliamo determinare la forma dei coefficinti $f_n \in \mathbb{C}$. Moltiplicando membro a membro per $\frac{e^{-i\frac{2\pi m}{L}x}}{\sqrt{L}}$ abbiamo:

$$f(x)\frac{e^{-i\frac{2\pi m}{L}x}}{\sqrt{2}} = \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \frac{e^{-i\frac{2\pi m}{L}x}}{\sqrt{L}} \ . \tag{6.20}$$

Tenendo infine conto delle relazioni

$$\frac{1}{L}\int_{-L/2}^{L/2} e^{-i\frac{2\pi(n-m)}{L}x} dx = \delta_{nm} \ ,$$

ed integrando i due membri di (6.20), ammettendo di poter scambiare il simbolo di integrale con quello di somma in (6.20) (questo è sicuramente possibile f è una combinazione lineare finita di esponenziali oppure se la serie converge uniformemente), giungiamo alla conclusione che:

$$\int_{-L/2}^{L/2} f(x)\frac{e^{-i\frac{2\pi m}{L}x}}{\sqrt{L}} dx = \sum_{n\in\mathbb{Z}} f_n \int_{-L/2}^{L/2} \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \frac{e^{-i\frac{2\pi m}{L}x}}{\sqrt{L}} dx = \sum_{n\in\mathbb{Z}} f_n \delta_{nm} = f_m \ .$$

Cambiando nome all'indice:

$$f_n = \int_{-L/2}^{L/2} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} f(x) dx \ . \tag{6.21}$$

I numeri complessi f_n, con $n \in \mathbb{Z}$, individuati da (6.21) quando esistono, sono detti **coefficienti di Fourier della funzione** f. Ora che abbiamo un candidato per i coefficienti di Fourier f_n, ci si può chiedere in quale senso la serie (6.19) converga.

6.3.1 La serie di Fourier nello spazio di Hilbert $L^2([-L/2, L/2], dx)$.

La teoria della serie di Fourier, a livello più astratto, viene sviluppata nell'insieme di funzioni $f : [-L/2, L/2] \to \mathbb{C}$ misurabili che siano **a quadrato sommabile**, cioè soddisfino:

$$\int_{[-L/2,L/2]} |f(x)|^2 dx < +\infty \,, \tag{6.22}$$

rispetto alla misura dx di Lebesgue. L'insieme di funzioni determinato in tal modo si indica con $\mathscr{L}^2([-L/2, L/2], dx)$. $\mathscr{L}^2([-L/2, L/2], dx)$ risulta essere uno spazio vettoriale complesso dotato della forma quadratica hermitiana:

$$(f|g) := \int_{[-L/2,L/2]} \overline{g(x)} f(x)\, dx \,, \quad \text{se } f, g \in \mathscr{L}^2([-L/2, L/2], dx).$$

Questo non è un prodotto scalare hermitiano unicamente per il fatto che $(f|f) = 0$ implica che $f(x) = 0$ quasi ovunque, ma non necessariamente ovunque. Si rimedia al problema identificando funzioni che differiscono tra di loro solo quando valutate su un (arbitrario) insieme di misura nulla in $[-L/2, L/2]$, e lavorando con classi (di equivalenza) di funzioni piuttosto che con funzioni. Lo spazio vettoriale con prodotto scalare hermitiano che si ottiene da $\mathscr{L}^2([-L/2, L/2], dx)$ quozientando rispetto alla relazione di equivalenza che identifica due funzioni se differiscono su un (qualsiasi) insieme di misura nulla, si indica con $L^2([-L/2, L/2], dx)$. Tale spazio vettoriale complesso risulta anche essere *completo* [3] nella topologia normata indotta dalla norma associata al prodotto scalare suddetto (dove ora, più propriamente f indica una classe di equivalenza di funzioni):

$$||f||_2 := \sqrt{\int_{[-L/2,L/2]} |f(x)|^2 dx} \,.$$

La competezza rende, per definizione, lo spazio vettoriale complesso $L^2([-L/2, L/2], dx)$ dotato del prodotto scalare $(\cdot|\cdot)$ uno **spazio di Hilbert** complesso.
Si osservi che la definizione di $\mathscr{L}^2([-L/2, L/2], dx)$ e $L^2([-L/2, L/2], dx)$ le loro proprietà generali sono indipendenti dalla teoria della serie di Fourier. In riferimento alla serie di Fourier, risulta [3] che vale il seguente teorema fondamentale.

Teorema 6.5. *In riferimento alla definizione (6.21) dei coefficienti di Fourier di una funzione a valori complessi $f : [-L/2, L/2] \to \mathbb{C}$ valgono i fatti seguenti.*
(a) *$f \in \mathscr{L}^2([-L/2, L/2], dx)$ se e solo se:*

$$\sum_{n\in\mathbb{Z}} |f_n|^2 < +\infty \,, \tag{6.23}$$

ed in tal caso vale:

$$\int_{[-L/2,L/2]} |f(x)|^2 dx = \sum_{n\in\mathbb{Z}} |f_n|^2 \,. \tag{6.24}$$

(b) *$f \in \mathscr{L}^2([-L/2, L/2], dx)$ se e solo se:*

$$\lim_{N\to+\infty} \int_{[-L/2,L/2]} \left| f(x) - \sum_{|n|\leq N} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \right|^2 dx \to 0 \,. \tag{6.25}$$

(c) *Se $f, g \in \mathscr{L}^2([-L/2, L/2], dx)$ e $\{f_n\}_{n\in\mathbb{Z}}$, $\{g_n\}_{n\in\mathbb{Z}}$ sono i rispettivi coefficienti di Fourier, allora:*

$$\sum_{n\in\mathbb{Z}} \overline{g}_n f_n = \int_{[-L/2,L/2]} \overline{g(x)} f(x)\, dx \tag{6.26}$$

dove la serie a primo membro converge assolutamente. $\diamondsuit$

Osservazioni 6.5.
(1) Dato che le serie numeriche considerate sopra siono assolutamente convergenti, non importa l'ordine con cui si esegue la somma della serie. Per esempio, in riferimento alla serie (6.26), è sufficiente numerare $\mathbb{Z}$ con un'arbitraria funzione biettiva $h : \mathbb{N} \to \mathbb{Z}$ e sommare la serie su $m \in \mathbb{Z}$

$$\sum_{m\in\mathbb{N}} \overline{g}_{h(n)} f_{h(n)} \,,$$

che si ottiene in tal modo. Il valore comune delle somme di tali serie che si ottengono comunque fissiamo la funzione biettiva h è , per definzione, il numero

$$\sum_{n\in\mathbb{Z}} \overline{g}_n f_n \,.$$

(2) A meno di non interpretare la serie di Fourier come serie di distribuzioni, il senso più generale con il quale si intende la convergenza della serie (6.19) è quello in (6.25). Questo tipo di convergenza, detta *convergenza (della serie) in $L^2([-L/2, L/2], dx)$* , è quello nella topologia normata indotta dalla norma $||\cdot||_2$ sopra definita. Si osservi che, come spiegato sopra, in questo caso la funzione f deve pensarsi come una classe di equivalenza di funzioni che differiscono su insiemi di misura nulla. La (6.25) si può dunque trascrivere come, se definiamo $e_n(x) := \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}$ quasi ovunque $x \in [-L/2, L/2]$:

$$\lim_{N\to+\infty} \left|\left| f - \sum_{|n|\leq N} f_n e_n \right|\right|_2 = 0 \,,$$

che si trova scritta frequentemente come, semplicemente:

$$f = \sum_{n\in\mathbb{Z}} f_n e_n \,. \tag{6.27}$$

È importante notare che, in generale, la convergenza in L^2 *non implica la convergenza puntuale della serie.* Per questo motivo non abbiamo scritto l'argomento di f ed e_n in (6.27), le quali

funzioni, tra l'altro, sono individuate a meno di insiemi di misura nulla.
Tenuto conto della (6.27), la (6.26) si interpreta come una versione infinita dell'ordinario prodotto scalare di $\mathbb{C}^n$, se $f, g \in L^2([-L/2, L/2], dx)$:

$$(f|g) = \sum_{n\in\mathbb{Z}} \overline{g}_n f_n \,,$$

dove gli infiniti elementi di base ortonormale dello spazio vettoriale $L^2([-L/2, L/2], dx)$ sono dati dagli infiniti elementi e_n. In particolare si ha, se $f \in L^2([-L/2, L/2], dx)$:

$$||f||_2^2 = \sum_{n\in\mathbb{Z}} |f_n|^2 \,.$$

6.3.2 Convergenza uniforme della serie di Fourier e derivazione sotto il simbolo di serie.

Dato che vogliamo sviluppare in serie di Fourier le soluzione dell'equazione di D'Alembert e Klein-Gordon, siamo più che altro interssati alla convergenza *puntuale* della serie di Fourier ed alla possibilità di *derivare sotto il segno di serie*. Mostriamo come si possano ottenere serie di Fourier con queste proprietà rafforzando le ipotesi di regolarità delle funzioni sviluppate in serie di Fourier. Ricordiamo che una funzione definita su $[a, b]$ si dice C^k **a tratti** su $[a, b]$ se si può decomporre $[a, b]$ come $[c_0, c_1] \cup [c_1, c_2] \cup \cdots \cup [c_m, b]$ con un numero finito di punti $a := c_0 < c_1 < \cdots < c_{m+1} =: b$, in modo tale che $f \restriction_{[c_l, c_{l+1}]} \in C^k([c_l, c_{l+1}])$ per $l = 0, \ldots, m$ (quindi, in particolare, esistono le derivate sinistre e destre fino all'ordine k anche sui bordi di $[c_l, c_{l+1}]$). Si osservi che la derivata k-esima di f (pensata come derivata destra o sinistra agli estremi di ogni sottointervallo $[c_k, c_{k+1}]$) può non essere continua su $[a, b]$ ma i valori che essa assume formano un insieme limitato.
Il primo risultato è stabilito nella seguente proposizione.

Proposizione 6.1. *Sia $N = 0, 1, \ldots$ fissato e $f : [-L/2, L/2] \to \mathbb{C}$ una funzione con i seguenti requisiti:*
(i) $f \in C^N([-L/2, L/2]; \mathbb{C})$,
(ii) f sia C^{N+1} a tratti su $[-L/2, L/2]$,
(iii) f sia periodica con tutte le sue derivate fino alla derivata N-esima inclusa.
Se f_n sono i coefficienti di Fourier di f dati da (6.21) e $f_n^{(k)}$ indica l'analogo coefficiente di Fourier della funzione $\frac{d^k f}{dx^k}$, allora vale qunato segue.
(a) *Per $k = 0, 1, \ldots, N + 1$, vale:*

$$f_n^{(k)} = \left(\frac{2\pi i}{L}\right)^k n^k f_n \quad \forall n \in \mathbb{Z} \,. \tag{6.28}$$

(b) *La serie di Fourier di f e delle sue derivate fino all'ordine $k = N + 1$ può essere derivata sotto il simbolo di serie (interpretando la convergenza delle serie nel senso di L^2), dato che*

risulta, per $k = 0, 1, \ldots, N+1$*:*

$$f_n^{(k)} \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} = \frac{d^k}{dx^k}\left(f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}\right) .$$

(c) *Per* $k = 0, 1, \ldots, N$ *vale:*

$$\sum_{n\in\mathbb{Z}} |n|^k |f_n| < +\infty . \tag{6.29}$$

$\diamondsuit$

Dimostrazione. (a) Fissiamo $k = 1, \ldots, N+1$. Dalla (6.21) abbiamo che:

$$f_n^{(k)} = \int_{-L/2}^{L/2} \frac{d^k f}{dx^k} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx = \int_{-L/2}^{L/2} \frac{d}{dx}\left(\frac{d^{k-1} f}{dx^{k-1}} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}}\right) dx - \int_{-L/2}^{L/2} \frac{d^{k-1} f}{dx^{k-1}} \frac{d}{dx} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx .$$

Il primo integrale a secondo membro risulta essere nullo dato che

$$[-L/2, L/2] \ni x \mapsto \frac{d^{k-1} f}{dx^{k-1}} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}}$$

è una funzione periodica su $[-L/2, L/2]$ per ipotesi. Abbiamo trovato che:

$$f_n^{(k)} = -\int_{-L/2}^{L/2} \frac{d^{k-1} f}{dx^{k-1}} \frac{d}{dx} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx .$$

Possiamo iterare $k-1$ volte la procedura, con lo stesso risultato, fino ad ottenere, alla fine

$$f_n^{(k)} = (-1)^k \int_{-L/2}^{L/2} f(x) \frac{d^k}{dx^k} \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx = (-1)^k \int_{-L/2}^{L/2} f(x) \left(-i\frac{2\pi n}{L}\right)^k \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx .$$

Abbiamo quindi trovato che

$$f_n^{(k)} = \left(i\frac{2\pi n}{L}\right)^k \int_{-L/2}^{L/2} f(x) \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx = \left(\frac{2\pi n i}{L}\right)^k \int_{-L/2}^{L/2} f(x) \frac{e^{-i\frac{2\pi n}{L}x}}{\sqrt{L}} dx = \left(\frac{2\pi i}{L}\right)^k n^k f_k ,$$

che è la (6.28).
(b) Il calcolo diretto mostra che

$$\frac{d^k}{dx^k}\left(f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}\right) = \left(\frac{2\pi i}{L}\right)^k n^k f_n e^{i\frac{2\pi n}{L}x} .$$

Da (a) abbiamo allora che, come enunciato nella tesi:

$$\frac{d^k}{dx^k}\left(f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}\right) = f_n^{(k)} \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} .$$

Questo risultato implica che si possa derivare sotto il segno di serie, interpretando però la convergenza nel senso di L^2, dato che lo sviluppo di Fourier di $\frac{d^k f}{dx^k}$ si scrive:

$$\frac{d^k f}{dx^k} = \sum_n f_n^{(k)} \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} = \sum_n \frac{d^k}{dx^k}\left(f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}\right) .$$

(c) Da (6.28) abbiamo anche che, se $k = 1\ldots, N+1$,

$$2\left|n^{k-1} f_n\right| = 2\left|n^k f_n\right| \frac{1}{n} = \left(\frac{L}{2\pi}\right)^k 2\frac{\left|f_n^{(k)}\right|}{n} \leq \left(\frac{L}{2\pi}\right)^k \left(\left|f_n^{(k)}\right|^2 + \frac{1}{n^2}\right) ,$$

dove abbiamo banalmente usato la disuguaglianza

$$0 \leq \left(\left|f_n^{(k)}\right| - \frac{1}{n}\right)^2 = \left|f_n^{(k)}\right|^2 + \frac{1}{n^2} - 2\frac{\left|f_n^{(k)}\right|}{n} .$$

Concludiamo che, se $k-1 = 1, 2, \ldots N$

$$\sum_{n\in\mathbb{Z}} \left|n^{k-1} f_n\right| \leq \frac{1}{2}\left(\frac{L}{2\pi}\right)^k \left(\sum_{n\in\mathbb{Z}} \left|f_n^{(k)}\right|^2 + \sum_{n\in\mathbb{Z}} \frac{1}{n^2}\right) < +\infty ,$$

ossia cambiando il nome di $k-1$ in k ed assumendo ora $k = 1, 2, \ldots N$:

$$\sum_{n\in\mathbb{Z}} \left|n^k f_n\right| \leq \frac{1}{2}\left(\frac{L}{2\pi}\right)^{k+1} \left(\sum_{n\in\mathbb{Z}} \left|f_n^{(k+1)}\right|^2 + \sum_{n\in\mathbb{Z}} \frac{1}{n^2}\right) < +\infty .$$

A commento del $< +\infty$, si osservi che la seconda serie a secondo membro converege come ben noto, mentre la prima serie a secondo membro converge per (a) del teorema (6.5), dato che ogni funzione $\frac{d^{k+1} f}{dx^{k+1}}$, per $k = 0, \ldots, N$ è limitata in valore assoluto da qualche numero $M < +\infty$ per ipotesi e quindi è in $L^2([-L/2, L/2], dx)$, avendo $[-L/2, L/2]$ misura finita:

$$\int_{[-L/2,L/2]} \left|\frac{d^{k+1} f}{dx^{k+1}}\right|^2 dx \leq \int_{[-L/2,L/2]} M^2 dx = M^2 L < +\infty .$$

□

Abbiamo poi il seguente utile risultato che discende dalla precedente proposizione.

Proposizione 6.2. *Sia $N = 0, 1, \ldots$ fissato e $f : [-L/2, L/2] \to \mathbb{C}$ una funzione con i seguenti requisiti:*
(i) $f \in C^N([-L/2, L/2]; \mathbb{C})$,
(ii) f sia C^{N+1} a tratti su $[-L/2, L/2]$,

(iii) f sia periodica con tutte le sue derivate fino alla derivata N-esima inclusa. Allora gli sviluppi di Fourier, per $k = 0, 1, \ldots, N$:

$$\frac{d^k f}{dx^k}(x) = \sum_{n\in\mathbb{Z}} f_n^{(k)} \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}$$

convergono puntualmente, assolutamente ed uniformemente su $[-L/2, L/2]$ (dove $f_n^{(k)}$ è l'n-esimo coefficiente di Fourier di $\frac{d^k f}{dx^k}$ (con $f_k^{(0)} := f_k$)). $\diamondsuit$

Dimostrazione. Nelle ipotesi fatte, prendendo $k = 0$ in (6.29), abbiamo che

$$\sum_{n\in\mathbb{Z}} |f_n| < +\infty \,,$$

pertanto la serie di funzioni per $x \in [-L/2, L/2]$,

$$\sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}}$$

è termine a termine dominata dalla serie di costanti convergente

$$\frac{1}{\sqrt{L}} \sum_{n\in\mathbb{Z}} |f_n| < +\infty \,,$$

dove abbiamo usato il fatto che $|e^{i\frac{2\pi n}{L}x}| = 1$. Come conseguenza di un ben noto teorema di Weierstrass, esisterà una funzione $g : [-L/2, L/2] \to \mathbb{C}$ con

$$g(x) = \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \,,$$

in cui la convergenza della serie è assoluta ed uniforme. Di conseguenza la convergenza vale anche nel senso di L^2, dato che:

$$\left\| g - \sum_{|n|\le N} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \right\|_2^2 \le \int_{-L/2}^{L/2} \left| g - \sum_{|n|\le N} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \right|^2 dx$$

ed il secondo membro è superiormente limitato da:

$$\sup_{x\in[-L/2,L/2]} \left| g - \sum_{|n|\le N} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \right|^2 \int_{-L/2}^{L/2} dx \to 0 \quad \text{se } N \to +\infty$$

a causa della convergenza uniforme della serie. Dato che la serie converge anche a f nel senso di L^2, deve essere

$$\sqrt{\int_{[-L/2,L/2]} |f(x)-g(x)|^2\,dx} = ||f-g||_2 = \left|\left| f - \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}\cdot}}{\sqrt{L}} + \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}\cdot}}{\sqrt{L}} - g \right|\right|_2 .$$

e quindi

$$\sqrt{\int_{[-L/2,L/2]} |f(x)-g(x)|^2\,dx} \leq \left|\left| f \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}\cdot}}{\sqrt{L}} \right|\right|_2 + \left|\left| \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}\cdot}}{\sqrt{L}} - g \right|\right|_2 = 0 .$$

Concludiamo che $f(x) = g(x)$ quasi ovunqunque. Data la continuità di f e g, dovrà essere $f(x) = g(x)$ ovunque su $[-L/2, L/2]$. Abbiamo ottenuto che, nel senso della convergenza uniforme, vale su $[-L/2, L/2]$:

$$f(x) = (g(x) =) \sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} .$$

Se $N \geq 1$, possiamo fare lo stesso ragionamento anche per la serie:

$$\frac{df}{dx} = \sum_{n\in\mathbb{Z}} f_n^{(1)} \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} ,$$

tenendo conto che, essendo per (6.28),

$$f_n^{(1)} = \frac{2\pi i}{L} n f_n ,$$

deve valere:

$$\sum_{n\in\mathbb{Z}} |f_n^{(1)}| = \frac{2\pi}{L} \sum_{n\in\mathbb{Z}} |n f_n| < +\infty ,$$

dove abbiamo applicato (6.29) ristretta al caso $k = 1$. In questo modo, seguendo la stessa strada seguita per la serie della funzione f, si prova che la serie di Fourier di $\frac{df}{dx}$ converge assolutamente ed uniformemente a $\frac{df}{dx}$. Si procede nello stesso modo, usando (6.28) e (6.29) per ogni ordine di derivazione k fino a $k = N$ (e non oltre dato che non è assicurato che valga (6.29) per $k = N+1$). □

Osservazioni 6.6. In realtà si può provare che la serie di Fourier converge puntualmente sotto ipotesi molto più deboli di quelle che abbiamo usato sopra (anche se questo non garantisce la convergenza uniforme). Si ha a tal proposito il seguente classico teorema di Dirichlet che citiamo senza dimostrazione.

Teorema 6.6. **(Teorema di Dirichlet)** *Sia $f : [-L/2, L/2] \to \mathbb{C}$ una funzione con i seguenti requisiti:*

(i) sia limitata,
(ii) sia continua eccetto un numero finito di punti $x_k \in (-L/2, L/2)$, $k = 1, \dots, p$ in cui esistono finiti il limite destro $f(x_k^+)$ ed il limite sinistro $f(x_k^-)$,
(iii) ammetta in ogni punto derivata destra e sinistra, usando nei punti di discontinuità il limite destro e sinistro per il calcolo del rapporto incrementale da destra e da sinistra.
Sotto queste ipotesi la serie di Fourier di f (6.19), con coefficienti di Fourier dati da (6.21), soddisfa,
(a) *per ogni $x \in (L/2, L/2) \setminus \{x_1, \dots, x_p\}$:*

$$\sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} = f(x)\,, \tag{6.30}$$

(b) *per ogni $k = 1, \dots, p$,*

$$\sum_{n\in\mathbb{Z}} f_n \frac{e^{i\frac{2\pi n}{L}x_k}}{\sqrt{L}} = \frac{f(x_k^-) + f(x_k^+)}{2}\,, \tag{6.31}$$

(c) *per $x = \pm L/2$ la serie converge a $\frac{f(-L/2)+f(L/2)}{2}$.* $\diamondsuit$

6.4 Il problema su $\mathbb{R} \times [-L/2, L/2]$ con condizioni al bordo periodiche.

Consideriamo ora il problema di determinare le soluzioni dell'equazione di Klein-Gordon o D'Alembert nell'insieme $\mathbb{R} \times [-L/2, L/2]$ una volta imposte condizioni iniziali e condizioni di periodicità ai bordi del compatto $[-L/2, L/2]$. L'esistenza di soluzioni sarà provata facendo uso della teoria della serie di Fourier sviluppata precedentemente in particolare la proposizione 6.1 ed il la proposizione 6.2.

6.4.1 Teorema di unicità.

Abbiamo un primo teorema di unicità.

Teorema 6.7. *Si consideri il seguente problema su $\mathbb{R} \times [-L/2, L/2]$ con $\mu \geq 0$ costante fissata,*

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\varphi}{\partial t^2} + \dfrac{\partial^2\varphi}{\partial x^2} - \mu^2\varphi = 0\,, \quad \varphi \in C^2(\mathbb{R} \times [-L/2, L/2], \mathbb{C})\,, \\ \varphi(t, -L/2) = \varphi(t, L/2)\,, \quad \dfrac{\partial\varphi}{\partial x}(t, -L/2) = \dfrac{\partial\varphi}{\partial x}(t, L/2) \quad \forall t \in \mathbb{R}\,, \\ \varphi(0, x) = \phi_0(x) \quad \forall x \in [-L/2, L/2]\,, \\ \dfrac{\partial\varphi}{\partial t}(0, x) = \phi_1(x) \quad \forall x \in [-L/2, L/2]\,, \end{cases} \tag{6.32}$$

dove $\phi_0 \in C^2([-L/2, L/2], \mathbb{C})$ *e* $\phi_1 \in C^1([-L/2, L/2], \mathbb{C})$ *sono funzioni assegnate che soddisfano le condizioni di periodicità*[1]*:*

$$\phi_0(-L/2) = \phi_0(L/2)\,, \quad \frac{\partial \phi_0}{\partial x}(-L/2) = \frac{\partial \phi_0}{\partial x}(L/2)\,, \quad \frac{\partial^2 \phi_0}{\partial x^2}(-L/2) = \frac{\partial^2 \phi_0}{\partial x^2}(L/2) \tag{6.33}$$

e

$$\phi_1(-L/2) = \phi_1(-L/2)\,, \quad \frac{\partial \phi_1}{\partial x}(-L/2) = \frac{\partial \phi_1}{\partial x}(L/2)\,. \tag{6.34}$$

Se esiste una soluzione al problema posto, essa è unica. In particolare, se i dati iniziali ϕ_0 *e* ϕ_1 *sono funzioni a valori reali, la soluzione* φ *è a valori reali.* $\diamondsuit$

Dimostrazione. Se una soluzione φ del problema, ammesso che esista, è complessa, possiamo sempre decomporla in parte reale ed immaginaria: $\varphi(t,x) = Re\varphi(t,x) + iIm\varphi(t,x)$. Data la natura reale dell'equazione di Klein-Gordon, avremo anche che la parte reale $Re\varphi$ e quella immaginaria $Im\varphi$, che sono funzioni *reali* con la stessa regolarità di φ, soddisfano la stessa equazione di Klein-Gordon separatamente. Inoltre soddisfano le condizioni al contorno di periodicità e si raccordano, separatamente, alle parti reali ed immaginarie dei dati iniziali per costruzione. In base a ciò è sufficiente provare il teorema di unicità nel caso di φ reale (cioè per la parte reale ed immaginaria di φ separatamente, quando φ è complessa). Assumiamo dunque di lavorare con funzioni reali soluzioni del problema considerato con dati iniziali reali. La dimostrazione della proprietà di unicità, è, escluso un punto, uguale a quella del teorema 5.1 ponendo $(\alpha, \beta) := \mathbb{R}$, $D := [-L/2, L/2]$. L'unica differenza è che ora, se ϕ è la differenza di due soluzioni del problema posto, l'identità

$$\int_0^T dt \oint_{+\partial D} \left(\frac{\partial \phi}{\partial t} \nabla \phi \right) \cdot \mathbf{n}\, dS(x) = 0$$

nella dimostrazione del teorema 5.1 si scrive ora nella forma semplificata:

$$\int_0^T dt \left(\frac{\partial \phi}{\partial t} \frac{\partial \phi}{\partial x}\bigg|_{L/2} - \frac{\partial \phi}{\partial t} \frac{\partial \phi}{\partial x}\bigg|_{-L/2} \right) = 0\,,$$

e questa identità vale banalmente in virtù delle condizioni di periodicità imposte sulle soluzioni del problema e quindi su ϕ.
Se la parte immaginaria dei dati iniziali è nulla, una soluzione del problema per la parte immaginaria di φ è la soluzione ovunque nulla. In base alla proprietà di unicità della soluzione, concludiamo che questa è l'unica soluzione e che quindi la parte immaginaria della soluzione (complessa a priori) φ è identicamente nulla. $\square$

[1] La terza delle condizioni in (6.33) deve essere imposta a causa delle condizioni di periodicità per φ scritte sopra e della forma dell'equazione differenziale stessa valutata a $t = 0$.

6.4.2 Esistenza delle soluzioni per dati iniziali sufficientemente regolari.

Passiamo ad un teorema di esistenza per il problema (6.32) con i vincoli (6.33) e (6.34). In realtà dovremo rinforzare le condizioni di regolarità sui dati iniziali per poter usare i risultati presentati prima relativi alla serie di Fourier. In riferimento al problema (6.32) con i vincoli (6.33) e (6.34), supponiamo che una soluzione φ esista e che sia sviluppabile in serie di Fourier per ogni tempo $t \in \mathbb{R}$. In tal caso avremo uno sviluppo del tipo:

$$\varphi(t,x) = \sum_{n\in\mathbb{Z}} C_n(t) \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \,. \tag{6.35}$$

Vogliamo trovare la forma delle funzioni del tempo C_n e poi le vogliamo fissare in modo tale da rispettare i dati iniziali. Si osservi che, ammesso che la serie converga puntualmente, le condizioni di periodicità in $\pm L/2$ sono già automaticamente rispettate, data la periodicità di ogni funzione: $x \mapsto e^{i\frac{2\pi n}{L}x}$. Assumendo che si possa derivare fino al secondo ordine sotto il segno di somma, risulta subito che:

$$-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} + \frac{\partial^2\varphi}{\partial x^2} - \mu^2\varphi = \sum_{n\in\mathbb{Z}} \frac{e^{i\frac{2\pi}{L}x}}{\sqrt{L}} \left\{ -\frac{1}{c^2}\frac{d^2C_n}{dt^2} - \left[\left(\frac{2\pi n}{L}\right)^2 + \mu^2\right] C_n \right\} = 0 \,.$$

Consideriamo allora il set infinito di equazioni:

$$\frac{d^2C_n}{dt^2} = -c^2\left[\left(\frac{2\pi n}{L}\right)^2 + \mu^2\right] C_n \,, \quad \forall n \in \mathbb{Z} \,. \tag{6.36}$$

Se i C_n le soddisfano, se il secondo membro di (6.35) converge puntualmente e se si possono passare sotto il segno di integrale le derivate di φ fino al secondo ordine dando luogo a funzioni continue (è richiesto che $\varphi \in C^2(\mathbb{R}\times[-L/2, L/2], \mathbb{C})$), allora il secondo membro di (6.35) definisce una soluzione dell'equazione di Klein-Gordon con le richieste condizioni di periodicità. (Queste ultime valgono in quanto le funzioni $e^{i2\pi nx/L}$ sono evidentemente periodiche con tutte le loro derivate di ogni ordine, pertanto, somme di tali funzioni saranno ancora periodiche. Nel caso di somme infinite, cioè serie, la perodicità varrà se le serie convergono puntualmente. Nel caso in esame, se la serie a secondo membro di (6.35) converge puntualmente e convergono le serie fino alle delle derivate prime in x (e si può scambiare l'operatore di derivata con il simbolo di serie), allora simao sicuri che, fino alla derivate prime in x è soddisfatto il vincolo di periodocità ai bordi di $[-L/2, L/2]$ per la soluzione.)
In generale il candidato per la soluzione φ data dalla serie a secondo membro di (6.35) sarà a valori complessi. Tuttavia se le condizioni iniziali sono rappresentate da funzioni reali e se la soluzione rispetta tali condizioni iniziali, il secondo membro di (6.35) definisce una funzione reale come garatito dal teorema di unicità sopra dimostrato.
Le soluzioni di (6.36) sono tutte e sole della forma

$$C_n(t) = C_n^{(+)} e^{-iE_n t} + C_{-n}^{(-)} e^{iE_n t} \,, \quad E_n := c\sqrt{\left(\frac{2\pi n}{L}\right)^2 + \mu^2} \,, \forall n \in \mathbb{Z} \,, \tag{6.37}$$

se $\mu > 0$, dove $C_n^{(\pm)} \in \mathbb{C}$ sono costanti arbitrarie. Se $\mu = 0$, abbiamo le soluzioni di sopra quando $n \neq 0$, con la differenza che ora:

$$C_0(t) = At + \sqrt{L}C_0 \,. \tag{6.38}$$

dove $A, C_0 \in \mathbb{C}$ sono costanti arbitrarie. In definitiva, un candidato soluzione per $\mu > 0$ è dato dalla popolare formula:

$$\varphi(t,x) = \sum_{n\in\mathbb{Z}} \left(\frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \right) \tag{6.39}$$

dove, nel secondo esponenziale abbiamo scambiato n con $-n$, dato che la somma opera su tutto $\mathbb{Z}$ e $E_n = E_{-n}$, ciò non altera il risultato, purché le due serie

$$\sum_{n\in\mathbb{Z}} \frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} \quad \text{e} \quad \sum_{n\in\mathbb{Z}} \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \tag{6.40}$$

convergano separatamente, cosa che proveremo tra poco. Nel caso $\mu = 0$, il candidato soluzione si deve modificare in:

$$\varphi(t,x) = At + \sum_{n\in\mathbb{Z}} \left(\frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \right) \,, \tag{6.41}$$

dove come prima, le due serie che si ottengono separando i due addendi nel termine generico della serie convergono separatamente, e $\sqrt{L}C_0 = C_0^{(+)} + C_0^{(-)}$. Si osservi che abbiamo trovato, nel caso $\mu = 0$, una forma di soluzione che è combinazione lineare di funzioni del tipo (6.18), come già discusso nella sezione 6.2. Sempre ammettendo che le due serie (6.40) convergano e che si possano derivare sotto il segno di somma nella variabile t, andiamo a valutare i coefficienti $C_n^{(\pm)}$ in modo da soddisfare le condizioni iniziali. Da (6.39) e ricordando che $\varphi(0,x) = \phi_0(x)$, $\frac{\partial\varphi}{\partial t}(0,x) = \phi_1(x)$ abbiamo che deve essere, dove scambiamo nuovamente n con $-n$ nel secondo esponenziale e teniamo conto che se le due serie in (6.40) convergono separatamente allora la somma di esse coincide con la somma della serie che ha come elementi la somma dei corrispondenti elementi delle due serie (e la stessa cosa accade per le serie delle derivate in t),

$$\phi_0(x) = \sum_{n\in\mathbb{Z}} \frac{C_n^{(+)} + C_{-n}^{(-)}}{\sqrt{L}} e^{i\frac{2\pi n}{L}x} \,, \quad \phi_1(x) = \sum_{n\in\mathbb{Z}} iE_n \frac{-C_n^{(+)} + C_{-n}^{(-)}}{\sqrt{L}} e^{i\frac{2\pi n}{L}x} \,. \tag{6.42}$$

La seconda formula ha un ulteriore addendo A a secondo membro se $\mu = 0$. Per ottenere la seconda identità abbiamo derivato in t sotto il segno di somma le due serie in (6.40) assumendo che ciò fosse possibile, e poi abbiamo posto $t = 0$ nel risultato. Si osservi che le due espressioni trovate non sono altro che gli sviluppi di Fourier di ϕ_0 e ϕ_1, i cui coefficienti di Fourier, rispettivamente, $\phi_{(0)\,n}$ e $\phi_{(1)\,n}$ sono stati scritti in funzione dei $C_{n\pm}$. Più precisamente

$$\phi_{(0)\,n} = C_n^{(+)} + C_{-n}^{(-)} \,, \quad \phi_{(1)\,n} = iE_n(-C_n^{(+)} + C_{-n}^{(-)}) \quad \forall n \in \mathbb{Z} \,, \quad \text{se } \mu > 0 \,,$$

oppure

$\phi_{(0)\,n} = C_n^{(+)} + C_{-n}^{(-)} \quad \forall n \in \mathbb{Z}, \quad \phi_{(1)\,n} = iE_n(-C_n^{(+)} + C_{-n}^{(-)}) \quad \forall n \in \mathbb{Z}\backslash\{0\}, \quad \phi_{(1)\,0} = A. \quad$ se $\mu > 0$,

Queste relazioni si invertono in:

$$C_n^{(+)} = \frac{1}{2}\phi_{(0)\,n} + \frac{i}{2E_n}\phi_{(1)\,n}\,, \quad C_n^{(-)} = \frac{1}{2}\phi_{(0)\,-n} - \frac{i}{2E_n}\phi_{(1)\,-n} \quad \forall n \in \mathbb{Z}\,, \tag{6.43}$$

se $\mu > 0$. Se $\mu = 0$ le identità di sopra valgono solo per $n \neq 0$, e vale anche:

$$C_0^{(+)} + C_0^{(-)} = \phi_{(0)\,0}\,, \quad A = \phi_{(1)\,0}\,. \tag{6.44}$$

Se $\mu = 0$, non è necessario conoscere $C_0^{(+)}$ e $C_0^{(-)}$ separatamente ai fini della costruzione della soluzione φ, dato che nel secondo membro di (6.39) *compare solo la loro somma $C_0^{(+)} + C_0^{(-)}$.*

Nota. Possiamo riassumere tutto come segue. Supponiamo che i dati iniziali ϕ_0 e ϕ_1 (assunti soddisfare (6.33) e (6.34)) siano sviluppabili in serie di Fourier convergenti puntualmente alle stesse ϕ_0 e ϕ_1. Consideriamo ancora la funzione φ definita dal secondo membro di (6.39) (o (6.41) se $\mu = 0$) dove i coefficienti $C_n^{(\pm)}$ soddisfano (6.43) (e (6.44) se $\mu = 0$). Se le due serie (6.40) in cui spezziamo la serie a secondo membro di (6.39) (o (6.41) se $\mu = 0$) convergono puntualmente e definiscono funzioni $C^2(\mathbb{R} \times [-L/2, L/2]; \mathbb{C})$ le cui derivate fino al secondo ordine possono essere calcolate derivando sotto il simbolo di somma, allora φ definito in (6.39) (o (6.41) se $\mu = 0$) è una soluzione del problema (6.32). Tutte queste richieste sono soddisfatte pur di assumere che i dati iniziali siano sufficientemente regolari come chiarito nel seguente teorema.

Teorema 6.8. *Si consideri il problema con condizioni al contorno periodiche (6.32) per la funzione $\varphi \in C^2(\mathbb{R} \times [-L/2, L/2])$ con $\mu \geq 0$ costante fissata.*
Se si assume che i dati iniziali soddisfano $\phi_0 \in C^2([-L/2, L/2], \mathbb{C})$ di classe C^3 a tratti su $[-L/2, L/2]$ e $\phi_1 \in C^1([-L/2, L/2], \mathbb{C})$ di classe C^2 a tratti su $[-L/2, L/2]$ e che valgano le condizioni di periodicità sui dati iniziali (6.33) e (6.34), allora esiste (ed è unica) una soluzione φ al problema. φ è data dalla serie convergente puntualmente:

$$\varphi(t, x) = \sum_{n \in \mathbb{Z}} \left(\frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \right)\,,$$

se $\mu > 0$, oppure:

$$\varphi(t, x) = At + \sum_{n \in \mathbb{Z}} \left(\frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \right)\,,$$

se $\mu = 0$. A secondo membro di entrambe le equazioni, $E_n := c\sqrt{\left(\frac{2\pi n}{L}\right)^2 + \mu^2}$ e i coefficienti $C_{n\pm}$ sono definiti da:

$$C_n^{(+)} = \frac{1}{2}\phi_{(0)\,n} + \frac{i}{2E_n}\phi_{(1)\,n}\,, \quad C_n^{(-)} = \frac{1}{2}\phi_{(0)\,-n} - \frac{i}{2E_n}\phi_{(1)\,-n} \quad \forall n \in \mathbb{Z}$$

se $\mu \neq 0$. Se $\mu = 0$ le identità di sopra valgono solo per $n \neq 0$, e vale:

$$C_0^{(+)} + C_0^{(-)} = \phi_{(0)\,0}\,, \quad A := \phi_{(0)\,1}\,,$$

infine, $\phi_{(0)\,n}$ e $\phi_{(1)\,n}$ sono, rispettivamente i coefficienti di Fourier dei dati iniziali ϕ_0 e ϕ_1. $\diamondsuit$

Dimostrazione. È sufficiente verificare che tutte le richieste formulate nella **Nota** scritta prima dell'eneunciato di questo teorema siano soddisfatte. Bisogna quindi verificare i due seguenti fatti:

(a) che ϕ_0 e ϕ_1 siano sviluppabili in serie di Fourier convergenti puntualmente alle stesse ϕ_0 e ϕ_1;

(b) che due serie (6.40) in cui spezziamo la serie a secondo membro di (6.39) (o (6.41) se $\mu = 0$) convergano puntualmente e definiscano funzioni $C^2(\mathbb{R} \times [-L/2, L/2]; \mathbb{C})$ le cui derivate fino al secondo ordine possono essere calcolate derivando sotto il simbolo di somma.

Ci restringeremo a lavorare per $\mu > 0$, dato che la dimostrazione per l'altro caso è banalmente simile.

Prima di tutto notiamo che (a) è vero dato che i dati iniziali sono $C^1([-L/2, L/2]; \mathbb{C})$ e quindi vale la proposizione 6.2. Non resta ora che provare (b). Nelle ipotesi fatte sulla regolarità di ϕ_0 e ϕ_1, abbiamo dalla proposizione 6.1 che

$$\sum_{n\in\mathbb{Z}} |n|^h |\phi_{(0)\,n}| < \infty \quad \text{se } h = 0, 1, 2, \qquad \sum_{n\in\mathbb{Z}} |n|^k |\phi_{(1)\,n}| < \infty \quad \text{se } k = 0, 1\,. \tag{6.45}$$

D'altra parte, usando la definizione (6.43), essendo $E_n := c\sqrt{\left(\frac{2\pi n}{L}\right)^2 + \mu^2}$, risulta che, per $|n|$ più grande di qualche fissato intero $M > 0$, $E_n > 1$ e quindi:

$$\frac{2}{\sqrt{L}} \left| C_n^{(\pm)} e^{\pm i\left(\frac{2\pi n}{L} x - E_n t\right)} \right| \leq |\phi_{(0)\,\pm n}| + \frac{1}{E_n} |\phi_{(1)\,\pm n}| \leq |\phi_{(0)\,\pm n}| + |\phi_{(1)\,\pm n}|\,.$$

Di conseguenza le serie di funzioni che appiono in (6.39) sono dominate dalle serie di costanti positive convergenti, per (6.45), $\sum_{n\in\mathbb{Z}} |\phi_{(0)\,\pm n}| + |\phi_{(1)\,\pm n}|$. In base al teorema di Weistrass le due serie (6.40) in cui spezziamo la serie a secondo membro di (6.39) e quindi la serie stessa a secondo membro di (6.39), converge assolutamente ed uniformemente ad una funzione continua φ su $\mathbb{R} \times [-L/2, L/2]$. Consideriamo ora la funzione definita in tal modo:

$$\varphi(t, x) = \sum_{n\in\mathbb{Z}} \frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L} x - E_n t\right)} + \sum_{n\in\mathbb{Z}} \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L} x - E_n t\right)}\,.$$

Consideriamo separatamente ognuna delle due serie a secondo membro. Possiamo derivare sotto il segno di serie rispetto alla variabile x (o t) se la serie delle derivate rispetto a x (o t) dei termini generici della serie iniziale converge uniformemente. Si osservi le derivate in x e t dei termini generici della serie definente φ sono funzioni continue. Se riusciamo a dominare la serie delle derivate in x e quella in t con serie di costanti convergenti, ragionando esattamente come prima usando il teorema di Weistrass, abbiamo non solo che φ è derivabile in x e t (e le

derivate si calcolano scambiando la serie con il simbolo di derivata corrispondente), ma anche che $\varphi \in C^1(\mathbb{R} \times [-L/2, L/2]; \mathbb{C})$. Infatti, in tal caso, le derivate di φ in x e t risulterebbero essere limiti di serie di funzioni continue convergenti uniformemente.
Le due serie delle derivate in x forniscono, a parte costanti moltiplicative comuni inessenziali,

$$\sum_{n\in\mathbb{Z}} \frac{nC_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} - \sum_{n\in\mathbb{Z}} \frac{nC_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} .$$

D'altra parte, usando la definizione (6.43), per $|n|$ più grande di qualche fissato intero $M' > 0$ vale anche $E_n \geq c|n|$, e quindi:

$$\frac{2}{\sqrt{L}} \left| nC_n^{(\pm)} e^{\pm i\left(\frac{2\pi n}{L}x - E_n t\right)} \right| \leq |n\phi_{(0)\,\pm n}| + \frac{|n|}{E_n}|\phi_{(1)\,\pm n}| \leq |n\phi_{(0)\,\pm n}| + \frac{1}{c}|\phi_{(1)\,\pm n}| \,.$$

Di conseguenza le due serie delle derivate in x (quella dei coefficienti $C_n^{(+)}$ e quella dei coefficienti $C_n^{(-)}$) sono dominate dalle due serie di costanti positive convergenti, per (6.45), date da $\sum_{n\in\mathbb{Z}} |n\phi_{(0)\,\pm n}| + \frac{1}{c}|\phi_{(1)\,\pm n}|$. In base al teorema di Weistrass la serie delle derivate in x converge assolutamente ed uniformemente ad una funzione continua φ su $\mathbb{R} \times [-L/2, L/2]$. Inoltre tale funzione deve coincidere con $\partial_x \varphi(t, x)$ dato che siamo nelle ipotesi di poter scambiare la derivata con il simbolo di somma nella serie che definisce φ.
Le serie delle derivate in t forniscono, a parte costanti moltiplicative comuni inessenziali,

$$\sum_{n\in\mathbb{Z}} \frac{E_n C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} - \sum_{n\in\mathbb{Z}} \frac{E_n C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} .$$

D'altra parte, usando la definizione (6.43), per $|n|$ più grande di qualche fissato intero $N > 0$ vale anche $E_n \leq c\sqrt{n^2 + 3n^2} = 2c|n|$, e quindi:

$$\frac{2}{\sqrt{L}} \left| E_n C_n^{(\pm)} e^{\pm i\left(\frac{2\pi n}{L}x - E_n t\right)} \right| \leq |E_n \phi_{(0)\,\pm n}| + |\phi_{(1)\,\pm n}| \leq 2c|n\phi_{(0)\,\pm n}| + |\phi_{(1)\,\pm n}| \,.$$

Di conseguenza le due serie delle derivate in t (quella dei coefficienti $C_n^{(+)}$ e quella dei coefficienti $C_n^{(-)}$) sono dominate dalle due serie di costanti positive convergenti, per (6.45), date da $\sum_{n\in\mathbb{Z}} 2c|n\phi_{(0)\,\pm n}| + |\phi_{(1)\,\pm n}|$. In base al teorema di Weistrass la serie delle derivate in t converge assolutamente ed uniformemente ad una funzione continua φ su $\mathbb{R} \times [-L/2, L/2]$. Inoltre tale funzione deve coincidere con $\partial_t \varphi(t, x)$ dato che siamo nelle ipotesi di poter scambiare la derivata con il simbolo di somma nella serie che definisce φ. La procedura può essere ripetuta per le derivate seconde, incluse quelle miste, si vede subito, in tal caso che le serie delle derivate seconde (di tipo fissato) sono comunque dominate da serie di costanti del tipo

$$\sum_{n\in\mathbb{Z}} A|n^2\phi_{(0)\,\pm n}| + B|n\phi_{(1)\,\pm n}| \,,$$

con $A, B > 0$ dipendenti dal tipo di derivate. Queste serie di costanti convergono per (6.45). Si conclude, con lo stesso ragionamento di sopra, che $\varphi \in C^2(\mathbb{R} \times [-L/2, L/2])$ e che la serie (6.39) che definisce φ si può derivare termine a termine fino alle derivate seconde. Questo è quanto volevamo e conclude la dimostrazione provando che

$$\begin{aligned}\varphi(t,x) &= \sum_{n\in\mathbb{Z}} \frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \sum_{n\in\mathbb{Z}} \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \\ &= \sum_{n\in\mathbb{Z}} \left(\frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \right)\end{aligned}$$

è una soluzione del problema considerato. □

Osservazioni 6.7. Cosideriamo la forma generale della soluzione per il problema con condizioni al bordo periodiche nella decomposizione della soluzione per il problema con condizioni periodiche in $[-L/2, L/2]$ per $\mu = 0$, cioè per l'equazione di D'Alembert:

$$\varphi(t,x) = \sum_{n\in\mathbb{Z}} \left(\frac{C_n^{(+)}}{\sqrt{L}} e^{i\left(\frac{2\pi n}{L}x - E_n t\right)} + \frac{C_n^{(-)}}{\sqrt{L}} e^{-i\left(\frac{2\pi n}{L}x - E_n t\right)} \right) ,$$

Trascuriamo pure il termine At che non ci interessa qui e teniamo conto del fatto che ora $E_n = c\frac{2\pi|n|}{L}$. Si subito vede che φ è una sovrapposizione di onde del tipo:

$$e^{\pm i\left(\frac{2\pi n}{L}x - iE_n t\right)} = \cos\left(\frac{2\pi n}{L}(x - ct)\right) \pm i \sin\left(\frac{2\pi n}{L}(x - ct)\right) \quad \text{con } n > 0,$$

$$e^{\pm i\left(\frac{2\pi n}{L}x - iE_n t\right)} = \cos\left(\frac{2\pi n}{L}(x + ct)\right) \pm i \sin\left(\frac{2\pi n}{L}(x + ct)\right) \quad \text{con } n < 0.$$

Queste soluzioni hanno la stessa forma delle soluzioni dell'equazione di D'Alembert in $\mathbb{R} \times \mathbb{R}$, cioè di soluzioni del tipo $f(x-ct)$ (onde progressive) oppure $g(x+ct)$ (onde regressive). Tuttavia ora, a parte la scelta di n, la loro forma funzionale è fissata: vedendole come funzioni reali, possono solo essere seni oppure coseni. Come nel caso della teoria in $\mathbb{R} \times \mathbb{R}$, c rappresenta la velocità di propagazione di tali profili, in questo caso si dice **velocità di fase** dell'onda φ_n. Consideriamo la soluzione $\varphi_n := \sin\left(\frac{2\pi n}{L}(x - ct)\right)$, per le altre analoghe soluzioni si possono fare discorsi analoghi. Se *fissiamo un punto* $x \in [-L/2, L/2]$ ed osserviamo, in quel punto, come oscilla φ_n, essa avrà un periodo di oscillazione $T_n = L/(nc)$. La **frequenza dell'onda** φ_n è , per definizione, l'inverso di tale periodo $\nu_n := nc/L$. Se invece, *a tempo fissato*, fotografiamo l'onda φ_n, essa sarà descritta da un sinusoide di periodo spaziale $\lambda_n = L/n$. Questo numero è detto **lunghezza d'onda** dell'onda φ_n. La lunghezza d'onda e la frequenza soddisfano la relazione, rispetto alla velocità di fase: $\lambda_n \nu_n = c$.

6.5 Il problema su $\mathbb{R} \times [-L/2, L/2]$ con condizioni al bordo di annullamento

Consideriamo ora il problema di determinare le soluzioni dell'equazione di Klein-Gordon o D'Alembert nell'insieme $\mathbb{R} \times [-L/2, L/2]$, una volta imposte condizioni iniziali e condizioni di annullamento ai bordi del compatto $[-L/2, L/2]$. Questo caso è fisicamente più interessante del precedente, dato che sistemi fisici comuni descritti dall'equazione di D'Alembert (es. le corde della chitarra), obbediscono a tali condizioni al contorno. L'esistenza di soluzioni sarà provata facendo uso della teoria della serie di Fourier sviluppata precedentemente.

6.5.1 Teorema di unicità.

Abbiamo un primo teorema di unicità.

Teorema 6.9. *Si consideri il seguente problema su $\mathbb{R} \times [-L/2, L/2]$ con $\mu \geq 0$ costante fissata,*

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\varphi}{\partial t^2} + \dfrac{\partial^2\varphi}{\partial x^2} - \mu^2\varphi = 0\,, \quad \varphi \in C^2(\mathbb{R} \times [-L/2, L/2], \mathbb{C})\,, \\ \varphi(t, L/2) = \varphi(t, -L/2) = 0 \quad \forall t \in \mathbb{R}\,, \\ \varphi(0, x) = \phi_0(x) \quad \forall x \in [-L/2, L/2]\,, \\ \dfrac{\partial\varphi}{\partial t}(0, x) = \phi_1(x) \quad \forall x \in [-L/2, L/2]\,, \end{cases} \tag{6.46}$$

dove $\phi_0 \in C^2([-L/2, L/2], \mathbb{C})$ e $\phi_1 \in C^1([-L/2, L/2], \mathbb{C})$ sono funzioni assegnate che soddisfano le condizioni di annullamento al bordo[2]*:*

$$\phi_0(-L/2) = \phi_0(L/2) = 0\,, \quad \frac{\partial^2\phi_0}{\partial x^2}(-L/2) = \frac{\partial^2\phi_0}{\partial x^2}(L/2) = 0 \tag{6.47}$$

e

$$\phi_1(-L/2) = \phi_1(-L/2) = 0\,. \tag{6.48}$$

Se esiste una soluzione al problema posto, essa è unica. In particolare, se i dati iniziali ϕ_0 e ϕ_1 sono funzioni a valori reali, la soluzione φ è a valori reali. $\diamondsuit$

Dimostrazione. Se una soluzione φ del problema, ammesso che esista, è complessa, possiamo sempre decomporla in parte reale ed immaginaria: $\varphi(t, x) = Re\varphi(t, x) + iIm\varphi(t, x)$. Data la natura reale dell'equazione di Klein-Gordon, avremo anche che la parte reale $Re\varphi$ e quella immaginaria $Im\varphi$, che sono funzioni *reali* con la stessa regolarità di φ, soddisfano la stessa equazione di Klein-Gordon separatamente. Inoltre soddisfano le condizioni al contorno di annullamento e si raccordano, separatamente, alle parti reali ed immaginarie dei dati iniziali per costruzione. In base a ciò è sufficiente provare il teorema di unicità nel caso di φ reale (cioè per la parte reale ed immaginaria di φ separatamente, quando φ è complessa). Assumiamo

[2] La seconda delle condizioni in (6.47) deve essere imposta a causa delle condizioni di annullamento al bordo di φ scritte sopra e della forma dell'equazione differenziale stessa valutata a $t = 0$.

dunque di lavorare con funzioni reali soluzioni del problema considerato con dati iniziali reali. La dimostrazione della proprietà di unicità, è, escluso un punto, uguale a quella del teorema 5.1 ponendo $(\alpha, \beta) := \mathbb{R}$, $D := [-L/2, L/2]$. L'unica differenza è che ora, se ϕ è la differenza di due soluzioni del problema posto, l'identità

$$\int_0^T dt \oint_{+\partial D} \left(\frac{\partial \phi}{\partial t} \nabla \phi \right) \cdot \mathbf{n} \, dS(x) = 0$$

nella dimostrazione del teorema 5.1 si scrive ora nella forma semplificata:

$$\int_0^T dt \left(\frac{\partial \phi}{\partial t} \frac{\partial \phi}{\partial x} \bigg|_{L/2} - \frac{\partial \phi}{\partial t} \frac{\partial \phi}{\partial x} \bigg|_{-L/2} \right) = 0 \, ,$$

e questa identità vale banalmente in virtù delle condizioni di annullamento al bordo imposte sulle soluzioni del problema e quindi su ϕ.
Se la parte immaginaria dei dati iniziali è nulla, una soluzione del problema per la parte immaginaria di φ è la soluzione ovunque nulla. In base alla proprietà di unicità della soluzione, concludiamo che questa è l'unica soluzione e che quindi la parte immaginaria della soluzione (complessa a priori) φ è identicamente nulla. $\square$

6.5.2 Esistenza delle soluzioni per dati iniziali sufficientemente regolari.

Passiamo ora ad un teorema di esistenza la cui dimostrazione sfrutta il teorema 6.8 di esistenza nel caso di condizioni al contorno periodiche ed un trucco piuttosto ingegnoso. Cominciamo con un lemma.

Lemma 6.3. *Nelle stesse ipotesi del teorema 6.7, se le condizioni iniziali ϕ_0 e ϕ_1, oltre a soddisfare le ipotesi del teorema, sono funzioni dispari, allora la soluzione φ, se esiste, è anch'essa una funzione dispari nella variabile x, cioè:*

$$\varphi(t, x) = -\varphi(t, -x) \, , \quad \forall (t, x) \in \mathbb{R} \times [-L/2, L/2] \, . \tag{6.49}$$

$\diamondsuit$

Dimostrazione. Sia φ la soluzione, ammesso che esista, del problema (6.32), con condizioni iniziali $\phi_0 \in C^2([-L/2, L/2], \mathbb{C})$ e $\phi_1 \in C^1([-L/2, L/2], \mathbb{C})$ date da funzioni dispari che soddisfano i vincoli al contorno (6.33) e (6.34). Consideriamo la funzione, definita su $\mathbb{R} \times [-L/2, L/2]$:

$$\Phi(t, x) := \varphi(t, x) + \varphi(t, -x) \, .$$

La soluzione φ è una funzione dispari nella variabile x se e solo se Φ è identicamente nulla. Dimostriamo che Φ è la funzione nulla se valgono le ipotesi del lemma. Si osservi che, per

costruzione $\Phi \in C^2(\mathbb{R} \times [-L/2, L/2], \mathbb{C})$. Inoltre, dato che nell'equazione di Klein-Gordon appaiono le derivate seconde nello spazio unicamente, la funzione $(t, x) \mapsto \varphi(t, -x)$ soddisferà l'equazione di Klein Gordon (dato che è soddisfatta da φ). Sommando membro a membro le due equazioni di Klein-Gordon per $\varphi(t, x)$ e $\varphi(t, -x)$, otteniamo che

$$-\frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} + \frac{\partial^2 \Phi}{\partial x^2} - \mu^2 \Phi = 0 \, .$$

Per costruzione la funzione Φ soddisfa anche

$$\Phi(0, x) = 0 \quad \frac{\partial \Phi}{\partial t}(0, x) = 0 \, , \quad \forall x \in [-L/2, L/2] \, ,$$

dato che le condizioni iniziali per φ sono per ipotesi delle funzioni dispari e quindi:

$$\Phi(0, x) = \varphi(0, x) + \varphi(0, -x) = \phi_0(x) + \phi_0(-x) = 0 \, ,$$

e anche

$$(\partial_t \Phi)(0, x) = (\partial_t \varphi)(0, x) + (\partial_t \varphi)(0, -x) = \phi_0(x) + \phi_0(-x) = 0 \, .$$

Infine Φ, sempre per costruzione, soddisfa le condizioni al contorno periodiche:

$$\Phi(t, L/2) = \varphi(t, L/2) + \varphi(t, -L/2) = \varphi(t, -L/2) + \varphi(t, L/2) = \Phi(t, -L/2) \, ,$$

e

$$\partial_t \Phi(t, L/2) = \partial_t \varphi(t, L/2) + \partial_t \varphi(t, -L/2) = \partial_t \varphi(t, -L/2) + \partial_t \varphi(t, L/2) = \partial_t \Phi(t, -L/2) \, .$$

In definitiva $\Phi \in C^2(\mathbb{R} \times [-L/2, L/2])$ soddisfa l'equazione di Klein-Gordon, con dati iniziali nulli e condizioni periodiche al bordo. Usando il teorema 6.9 concludiamo che questa deve essere l'unica soluzione del problema detto, ma allora deve coincidere con la soluzione banale data dalla funzione ovunque nulla, notando che la soluzione nulla risolve lo stesso problema (con le stesse condizioni iniziali ed al bordo). □

Possiamo ora enunciare e provare il teorema di esistenza. L'idea della dimostrazione è trasformare il problema con condizioni al bordo di annullamento in un problema con condizioni al bordo periodiche, ma definito su un dominio spaziale più grande. La soluzione determinata nel dominio più grande, che sappiamo esistere per il teorema 6.8, ristretta al dominio iniziale, sarà la soluzione del problema.

Teorema 6.10. *Si consideri il problema con condizioni al contorno periodiche (6.46) per la funzione $\varphi \in C^2(\mathbb{R} \times [-L/2, L/2])$ con $\mu \geq 0$ costante fissata.*
Se si assume che i dati iniziali soddisfano $\phi_0 \in C^2([-L/2, L/2], \mathbb{C})$ di classe C^3 a tratti su $[-L/2, L/2]$ e $\phi_1 \in C^1([-L/2, L/2], \mathbb{C})$ di classe C^2 a tratti su $[-L/2, L/2]$ e che valgano le condizioni di annullamento al bordo sui dati iniziali (6.47) e (6.48), allora esiste (ed è unica)

una soluzione φ al problema. $\diamondsuit$

Dimostrazione. Prima di tutto notiamo che tutti i risultati che abbiamo provato fino ad ora su $\mathbb{R} \times [-L/2, L/2]$, inclusi i teoremi di esistenza ed unicità in presenza di condizioni al bordo periodiche, valgono se si sostituisce $[-L/2, L/2]$ con un qualsiaso intervallo $[a, b]$, dove $a < b$. Anche le formule risolutive sono identiche con l'eccezione che il dominio d'integrazione $[-L/2, L/2]$ (per esempio nel calcolo dei coefficienti di Fourier) deve essere ovviamente rimpiazzato da $[a, b]$, ed il parametro L che appare, per esempio, negli esponenti deve essere sostituito con $b - a$. Con una traslazione di assi spaziali, portiamo il segmento $[-L/2, L/2]$ nel segmento $[0, L]$. Dimostreremo il teorema di esistenza in questo intervallo e poi torneremo sull'intervallo iniziale. La soluzione per l'intervallo $[-L/2, L/2]$ si otterrà banalmente con una traslazione di assi di $-L/2$.
Lavoriamo allora sul segmento $[0, L]$ sul quale definiamo le funzioni: $\tilde{\phi}_0(x) := \phi_0(x - L/2)$ e $\tilde{\phi}_1(x) := \phi_1(x - L/2)$ per ogni $x \in [0, L]$. Consideriamo poi il segmento di lunghezza doppia $[-L, L]$, estendiamo le funzioni $\tilde{\phi}_0$ e $\tilde{\phi}_1$ dal segmento $[0, L]$ a tutto il segmento $[-L, L]$ in modo tale che risultino funzioni dispari. Indichiamo le funzioni estese in questo modo con Φ_0 e Φ_1. Ora passeremo dal problema con condizioni al bordo di annullamento su $[0, L]$ ad un nuovo problema sul segmento allargato $[-L, L]$ con condizioni al bordo *periodiche* del quale conosciamo già un teorema di esistenza. La soluzione che otterremo in quel caso, ristretta al dominio originale, sarà la soluzione del nostro problema.
Date le proprietà delle funzioni ϕ_0 e ϕ_1, si ha facilemente che $\Phi_0 \in C^2([-L, L], \mathbb{C})$ (si noti che, riguardo alle derivate seconde, per il punto $x = 0$ vale la (6.47) che assicura che la derivata seconda in $x = 0$ esista e sia continua) ed è di classe C^3 a tratti su $[-L, L]$, $\Phi_1 \in C^1([-L, L], \mathbb{C})$ ed è di classe C^2 a tratti su $[-L, L]$. Infine sono soddisfatte le condizioni di periodicità al bordo di $[-L, L]$:

$$\Phi_0(-L) = \Phi_0(L) \ (= 0)\,, \quad \frac{\partial \Phi_0}{\partial x}(-L) = \frac{\partial \Phi_0}{\partial x}(L)\,, \quad \frac{\partial^2 \Phi_0}{\partial x^2}(-L) = \frac{\partial^2 \Phi_0}{\partial x^2}(L) \ (= 0)$$

e

$$\Phi_1(-L) = \Phi_1(L) \ (= 0)\,, \quad \frac{\partial \Phi_1}{\partial x}(-L) = \frac{\partial \Phi_1}{\partial x}(L)\,.$$

Si noti che le condizioni scritte sulle derivate prime sono conseguenza del fatto che le funzioni Φ_0 e Φ_1 sono funzioni dispari e quindi le loro derivate prime (in x) sono funzioni pari, le rimaneti condizioni sono anche conseguenza delle condizioni di annullamento al bordo delle funzioni ϕ_0 e ϕ_1. Possiamo allora invocare il teorema 6.8 che assicura l'esistenza di una funzione $\Phi \in C^2(\mathbb{R} \times [-L, L]; \mathbb{C})$ che soddisfi l'equazione di Klein-Gordon in tale insieme, che si raccordi con i dati iniziali Φ_0 e Φ_1 al tempo $t = 0$ e che soddisfi condizioni di periodicità

$$\Phi(t, -L) = \Phi(t, L)\,, \quad \frac{\partial \Phi}{\partial x}(t, -L) = \frac{\partial \Phi}{\partial x}(t, L) \quad \forall t \in \mathbb{R}\,.$$

La soluzione Φ è una funzione dispari in x per il lemma 6.3, dato che i dati iniziali sono funzioni dispari. Quindi, in particolare $\Phi(t, 0) = -\Phi(t, -0) = -\Phi(t, 0) = 0$ per ogni $t \in \mathbb{R}$ per il fatto che Φ è dispari. Inoltre essendo Φ dispari ma anche periodica su $[-L, L]$, deve vale

contemporaneamente $\Phi(t,L) = -\Phi(t,-L)$ e $\Phi(t,L) = \Phi(t,-L)$ e quindi $\Phi(t,L) = 0$ per ogni $t \in \mathbb{R}$.
Se allora definiamo $\phi(t,x) := \Phi\restriction_{\mathbb{R}\times[0,L]}(t,x)$, questa funzione soddisfa per costruzione:

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\phi}{\partial t^2} + \dfrac{\partial^2\phi}{\partial x^2} - \mu^2\phi = 0\,, \quad \varphi \in C^2(\mathbb{R}\times[0,L],\mathbb{C})\,, \\ \phi(t,0) = \varphi(t,L) = 0 \quad \forall t \in \mathbb{R}\,, \\ \phi(0,x) = \tilde{\phi}_0(x) \quad \forall x \in [0,L]\,, \\ \dfrac{\partial\phi}{\partial t}(0,x) = \tilde{\phi}_1(x) \quad \forall x \in [0,L]\,. \end{cases}$$

Di conseguenza, la funzione definita da $\varphi(t,x) := \phi(t,x+L/2)$ per ogni $(t,x) \in \mathbb{R}\times[-L/2,L/2]$ soddisfa

$$\begin{cases} -\dfrac{1}{c^2}\dfrac{\partial^2\varphi}{\partial t^2} + \dfrac{\partial^2\varphi}{\partial x^2} - \mu^2\varphi = 0\,, \quad \varphi \in C^2(\mathbb{R}\times[-L/2,L/2],\mathbb{C})\,, \\ \varphi(t,L/2) = \varphi(t,-L/2) = 0 \quad \forall t \in \mathbb{R}\,, \\ \varphi(0,x) = \phi_0(x) \quad \forall x \in [-L/2,L/2]\,, \\ \dfrac{\partial\varphi}{\partial t}(0,x) = \phi_1(x) \quad \forall x \in [-L/2,L/2]\,, \end{cases}$$

ed è quindi una soluzione del problema con condizioni al contorno periodiche (6.46) con dati iniziali ϕ_0 e ϕ_1. $\square$

Osservazioni 6.8.
(1) Studiamo la forma particolare dello sviluppo di Fourier della soluzione del problema considerato, *nel caso in cui il campo φ sia reale*, dato che si presta a qualche osservazione interessante dal punto di vista fisico, specialmente nel caso in cui $\mu = 0$, cioè per l'equazione di D'Alembert. Sotto opportune ipotesi di regolarità delle condizioni iniziali, la generica soluzione Φ del problema periodico su $\mathbb{R}\times[-L,L]$, come sappiamo dal teorema 6.8 è data dallo sviluppo di Fourier:

$$\Phi(t,x) = At + \sum_{n\in\mathbb{Z}} C_n^{(+)}\frac{e^{i\frac{\pi n}{L}x}}{\sqrt{2L}}e^{-iE_nt} + C_n^{(-)}\frac{e^{-i\frac{\pi n}{L}x}}{\sqrt{2L}}e^{iE_nt}\,,$$

dove il termine At può apparire solo se $\mu = 0$. Tuttavia, nel caso in esame, dato che richiederemo che Φ sia una funzione dispari di x, l'unica possibilità è $A = 0$ anche nel caso $\mu = 0$. Pertanto partiamo con la generica soluzione:

$$\Phi(t,x) = \sum_{n\in\mathbb{Z}} C_n^{(+)}\frac{e^{i\frac{\pi n}{L}x}}{\sqrt{2L}}e^{-iE_nt} + C_n^{(-)}\frac{e^{-i\frac{\pi n}{L}x}}{\sqrt{2L}}e^{iE_nt}\,, \tag{6.50}$$

dove $E_n = \sqrt{\left(\frac{\pi n}{L}\right)^2 + \mu}$. Dalla dimostrazione del teorema 6.10, sappiamo che la soluzione generica del problema con condizioni di annullamento al bordo su $[0,L]$ si ottiene restringendo la funzione Φ a $[0,L]$ sotto l'ipotesi che Φ sia *dispari*. Ma allora deve valere:

$$\Phi(t,x) = -\Phi(t,-x) = \sum_{n\in\mathbb{Z}} -C_n^{(+)}\frac{e^{-i\frac{\pi n}{L}x}}{\sqrt{2L}}e^{-iE_nt} - C_n^{(-)}\frac{e^{i\frac{\pi n}{L}x}}{\sqrt{2L}}e^{iE_nt}\,. \tag{6.51}$$

Sommando membro a membro con (6.50) e dividendo il risultato a metà si ottiene allora che:

$$\Phi(t,x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{Z}} iC_n^{(+)} e^{-iE_n t} \sin\left(\frac{\pi n x}{L}\right) - iC_n^{(-)} e^{iE_n t} \sin\left(\frac{\pi n x}{L}\right) . \tag{6.52}$$

Ora, tenendo conto del fatto che Φ è reale, possiamo ancora scrivere che:

$$\Phi(t,x) = \overline{\Phi(t,x)} = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{Z}} \overline{iC_n^{(+)}} e^{iE_n t} \sin\left(\frac{\pi n x}{L}\right) - \overline{iC_n^{(-)}} e^{-iE_n t} \sin\left(\frac{\pi n x}{L}\right) , \tag{6.53}$$

che, sommata membro a membro con (6.52), fornisce:

$$\Phi(t,x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{Z}} -Im\left(C_n^{(+)} e^{-iE_n t}\right) \sin\left(\frac{\pi n x}{L}\right) + Im\left(C_n^{(-)} e^{iE_n t}\right) \sin\left(\frac{\pi n x}{L}\right) .$$

Se infine definiamo $C_n^{(\pm)} = \alpha_n^{(\pm)} + i\beta_n^{(\pm)}$, e decomponiamo gli esponenziali complessi $e^{\pm iE_n t} = \cos(E_n t) + i\sin(E_n t)$, con un semplice calcolo, l'identità trovata si riduce a:

$$\Phi(t,x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{Z}} \left(\alpha_n^{(-)} - \alpha_n^{(+)}\right) \sin(E_n t) \sin\left(\frac{\pi n x}{L}\right) + \left(\beta_n^{(-)} - \beta_n^{(+)}\right) \cos(E_n t) \sin\left(\frac{\pi n x}{L}\right) .$$

Possiamo concludere che, per un' opportuna scelta delle costanti *reali* A_n e B_n etichettate su $\mathbb{N}$, la soluzione del problema con condizioni di annullamento al bordo di $[0,L]$ ha la struttura, se $(t,x) \in \mathbb{R} \times [0,L]$:

$$\varphi(t,x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{N}\setminus\{0\}} \left[A_n \sin(E_n t) + B_n \cos(E_n t)\right] \sin\left(\frac{\pi n x}{L}\right) . \tag{6.54}$$

Sopra abbiamo omesso i termini con $n = 0$ dato che non forniscono alcun contributo essendo $\sin 0 = 0$, inoltre abbiamo tenuto conto del fatto che $E_n = E_{-n}$, $\sin\left(\frac{\pi n x}{L}\right) = -\sin\left(\frac{-\pi n x}{L}\right)$ e questo consente di sommare sui naturali invece che sugli interi raccogliendo i coefficienti opportunamente.
Mostriamo ora come individuare i coefficienti A_n e B_n in funzione dei dati iniziali. Con la stessa procedura che abbiamo usato nella dimostrazione del teorema 6.8 si riesce facilmente a dimostrare che la serie di sopra converge uniformemente, si può derivare sotto il segno di somma nella vartiabile t ottenendo una serie che converge ancora uniformemente:

$$\frac{\partial}{\partial t}\varphi(t,x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{N}\setminus\{0\}} \left[A_n \cos(E_n t) - B_n \sin(E_n t)\right] E_n \sin\left(\frac{\pi n x}{L}\right) .$$

Specializzando le due serie per $t = 0$ si ha allora il legame con i dati iniziali:

$$\phi_0(x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{N}\setminus\{0\}} B_n \sin\left(\frac{\pi n x}{L}\right) , \quad \phi_1(x) = \frac{1}{\sqrt{2L}} \sum_{n\in\mathbb{N}\setminus\{0\}} E_n A_n \sin\left(\frac{\pi n x}{L}\right) .$$

Moltiplicando entrambe le espressioni per $\sin\left(\frac{\pi m x}{L}\right)$ ed integrando su $[0, L]$, passando il simbolo di integrale sotto quello di serie, dato che ciò è concesso per via della uniforme convergenza della serie, si trova infine:

$$B_n = \sqrt{\frac{8}{L}} \int_0^L \sin\left(\frac{\pi n x}{L}\right) \phi_0(x) dx \ , \quad A_n = \sqrt{\frac{8}{L E_n^2}} \int_0^L \sin\left(\frac{\pi n x}{L}\right) \phi_1(x) dx \quad \forall n \in \mathbb{N} \setminus \{0\} \ , \quad (6.55)$$

dove abbiamo tenuto conto dell'identità

$$\int_0^L \sin\left(\frac{\pi m x}{L}\right) \sin\left(\frac{\pi n x}{L}\right) \, dx = \frac{L}{2} \delta_{nm} \ , \quad \forall m, n \in \mathbb{N} \setminus \{0\} \ .$$

(2) Consideriamo esplicitamente il caso dell'equazione di D'Alembert, $\mu = 0$, per cui $E_n = c \left| \frac{\pi n}{L} \right|$. La decomposizione (6.54) di $\varphi(t, x)$ è interessante perché non è data in termini di onde regressive o progressive, come quelle che appaiono nella decomposizione della soluzione per il problema con condizioni periodiche in $[0, L]$:

$$e^{\pm i(\frac{2\pi n}{L} x - i E_n t)} = \cos\left(\frac{2\pi n}{L}(x - ct)\right) \pm i \sin\left(\frac{2\pi n}{L}(x - ct)\right) \quad \text{con } n > 0,$$

$$e^{\pm i(\frac{2\pi n}{L} x - i E_n t)} = \cos\left(\frac{2\pi n}{L}(x + ct)\right) \pm i \sin\left(\frac{2\pi n}{L}(x + ct)\right) \quad \text{se } n < 0.$$

Invece appaiono soluzioni dette **onde stazionarie** o **armoniche** (che non si vedono "propagare" nelle due direzioni come invece accade nel caso di condizioni periodiche):

$$A_n \sin\left(\frac{c\pi n}{L} t\right) \sin\left(\frac{\pi n x}{L}\right) + B_n \cos\left(\frac{c\pi n}{L} t\right) \sin\left(\frac{\pi n x}{L}\right) \ , \quad \forall n \in \mathbb{N} \setminus \{0\} \ .$$

A differenza delle funzioni precedenti, in queste, vi sono punti nello spazio, detti **nodi**, in cui *le onde si annullano per ogni valore del tempo.* Le posizioni $x_k^{(n)}$ dei noti sono ottenute risolvendo $\sin(\pi n x_k^{(n)}/L) = 0$ su $[0, L]$ ed ottenendo quindi, per ogni fissato $n \in \mathbb{N} \setminus \{0\}$, $x_k^{(n)} = \frac{k}{n} L$ per tutti i $k \in \mathbb{N}$ con $k \leq n$.

(3) Le corde degli strumenti musicali "a corda" vibrano trasversalmente soddisfacendo l'equazione di D'Alembert con condizioni di annullamento al bordo. La vibrazione è trasmessa all'aria ed è percepita dal nostro orecchio in termini di suono. L'onda sonora ha una struttura simile a quella dell'onda meccanica di deformazione delle corde che l'hanno prodotto, per tale motivo ci riferiremo a tale onda meccanica pur parlando del suono. La decomposizione in onde stazionarie o armoniche, descritta sopra, del moto di oscillazione di una corda che produce suono è il punto di partenza per la teoria fisica della musica prodotta da tali strumenti. Questo genere di onde oscillano temporalmente con un periodo $T_n = 2L/(cn)$ e quindi una frequenza $\nu_n = cn/(2L)$. La lunghezza d'onda spaziale è ancora data da $\lambda_n = 2L/n$. Le oscillazioni con

n fissato, le armoniche, vengono avverite dal nostro orecchio come toni puri (note pure)[3]. la nota LA di riferimento, per accordare gli strumenti corrisponde alla frequenza di 440 oscillazioni al secondo. Tuttavia è difficilissimo, con strumenti meccanici (non elettronici) produrre toni puri. Infatti, il suono che si ottiene pizzicando una corda (clavicembalo) oppure perquotendola (pianoforte), corrisponde ad una soluzione dell'equazione di D'Alembert con condizioni di annullamento al bordo, la cui forma e decomposizione in armoniche sinusoidali dipende dalle condizioni iniziali, cioè dalla procedura con la quale è stata fatta oscillare la corda. Il suono è praticamente sempre composto da molte armoniche secondo una certa distribuzione con coefficienti A_n e B_n che dipendono dalla procedura usata per mettere la corda in oscillazione. Il numero $A_n^2 + B_n^2$ è legato all'intensità (o energia) del suono, più precisamente all'intensità dell'armonica n-esima. Quando si cerca di suonare una precisa nota mettendo in oscillazione una certa corda in un certo modo, in realtà si produce un certa soluzione delle equazioni di D'Alembert tale che, decomponendola in armoniche, una certa frequenza prevale sulle altre: *cioè, l'armonica corrispondente ha coefficienti A_n e B_n più grandi di tutti gli altri analoghi coefficienti delle altre armoniche.* La frequenza dell'armonica che prevale definisce la nota suonata. Le rimanenti armoniche della decomposizione attenuate, ma sempre presenti e con intensità che dipendono fal tipo di strumento, producono il caratteristico *timbro* dello strumento, per il quale una stessa nota, suonata da un violino oppure da un pianoforte viene avvertita dal nostro orecchio come differente.

[3]In realtà il nostro orecchio sembra essere incapace di distiguere le singole frequenze (malgrado talvolta si sostenga che qualcuno sia abile a farlo e si parla, in tal caso, di "orecchio assoluto"). Il nostro orecchio è in realtà capace di distiguere solo i rapporti tra varie frequenze suonate contemporaneamente o a breve distanza temporale. Per esempio il rapporto di *un ottava* è quello di due frequenze una di valore doppio dell'altra. Nella teoria musicale ideata da Bach il rapporto tra due note consecutive e $2^{1/12}$ ed è questo il rapporto di frequenze tra le note di due tasti consecutivi in un pianoforte, includendo nell'ordine sia i tasti bianchi che quelli neri (che sono appunto, $7 + 5 = 12$ in tutto per ogni ottava).

Appendice A

Un accenno all'approccio moderno per il problema ellittico: soluzioni in senso debole e teoremi di regolarità ellittica.

In questa appendice tutti gli integrali che appaiono sono riferiti alla misura di Lebesgue.

Tutto l'approccio moderno allo studio delle soluzioni dell'equazione di Laplace e Poisson si basa sulla seguente definizione.

Definizione A.1. Sia $\Omega \subset \mathbb{R}^n$ un aperto non vuoto e sia $f \in \mathscr{L}^1_{loc}(\Omega)$ assegnata. Una funzione $\varphi : \Omega \to \mathbb{R}$ con $\varphi \in \mathscr{L}^1_{loc}(\Omega)$ è detta risolvere l'equazione di Poisson:

$$\Delta\varphi = f$$

in senso debole o, equivalentemente, **in senso distribuzionale**, se vale l'identità :

$$\int_\Omega \varphi \Delta g \, d^n x = \int_\Omega f g \, d^n x \, , \quad \text{per ogni funzione } g \in C_0^\infty(\Omega). \tag{A.1}$$

$\diamondsuit$

Chiariamo subito il significato di questa definizione. Prima di tutto vediamo perché si parla di soluzioni *in senso debole*. Supponiamo che $\varphi \in C^2(\Omega)$ risolva l'equazione di Poisson $\Delta\varphi = f$ in senso proprio. Usando l'integrazione per parti, se $g \in C_0^\infty(\Omega)$, si ha immediatamente che:

$$\int_\Omega f g \, d^n x = \int_\Omega (\Delta\varphi) g \, d^n x = \int_\Omega \nabla \cdot [(\nabla\varphi) g] \, d^n x - \int_\Omega (\nabla\varphi) \cdot \nabla g \, d^n x$$

$$= \int_\Omega \nabla \cdot [(\nabla\varphi)g]\, d^n x - \int_\Omega \nabla \cdot [\varphi \nabla g]\, d^n x + \int_\Omega \varphi \Delta g\, d^n x\,.$$

Tenendo conto che g si annulla (con tutte le derivate) prima di arrivare al bordo di Ω, i primi due integrali nell'ultima riga risultano essere nulli (per il secondo teorema fondamentale del calcolo oppure il teorema della divergenza). Pertanto rimane:

$$\int_\Omega fg\, d^n x = \int_\Omega f\Delta g\, d^n x\,.$$

Abbiamo in tal modo provato che: *le soluzioni in senso proprio sono anche soluzioni in senso debole.*

Non ci aspettiamo che valga il contrario per un motovo elementare: se φ soddisfa l'equazione di Poisson in senso proprio e pertanto vale l'identità l'identità (A.1), quest'ultima varrà anche se la funzione φ è ridefinita in modo da non essere più differenziabile su un insieme di misura nulla secondo Lebesgue, per esempio l'insieme dei punti di coordinate razionali in Ω. Tuttavia, se sappiamo che la soluzione in senso debole φ di $\Delta\varphi = f$, con $f \in C^0(\Omega)$, *è una funzione di classe* $C^2(\Omega)$, allora possiamo concludere che φ è anche una soluzione in senso proprio. La dimostrazione è abbastanza semplice. Partendo dalla (A.1) e tenendo conto che $\varphi \in C^2(\Omega)$ si ha immediatamente procedendo in senso inverso a quanto fatto sopra:

$$\int_\Omega (\Delta\varphi - f)g = 0\, d^n x\,, \quad \text{per ogni } g \in C_0^\infty(\Omega) \tag{A.2}$$

Se fosse $\Delta_x\phi - f(x) \neq 0$ nel punto $x \in \Omega$, per continuità , il segno di tale funzione dovrebbe mantenersi costante in un intorno di x. Supponiamo il segno sia positivo (l'altro caso si studia analogamente) sulla palla aperta $B(x)$ a chiusura compatta con $\overline{B(x)} \subset \Omega$. Stringendo tale palla se necessario, si avrebbe $\min_{\overline{B(x)}}(\Delta\phi - f) \geq k > 0$. Se ora $g \in C_0^\infty(\Omega)$ è tale che $\operatorname{supp} g \subset \overline{B(x)}$, $g \geq 0$ e $\int_\mathbb{R} g\, d^n x = 1$ (si possono costruire facilmente tali funzioni), allora:

$$\int_\Omega (\Delta\phi - f)g\, d^n x = \int_{\overline{B(x)}} (\Delta\phi - f)g\, d^n x \geq \int_{\overline{B(x)}} kg\, d^n x \geq k > 0\,.$$

La nozione di soluzione debole è legata ad una nozione più generale: quella di *derivata in senso debole.*

Definizione A.2. Se $\Omega \subset \mathbb{R}^n$ è un aperto non vuoto, si dice che $f : \Omega \to \mathbb{R}$ ($\mathbb{C}$) **ammette derivata**

$$h = \frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\partial x^{1\alpha_1} \cdots \partial x^{n\alpha_1 n}}$$

in senso debole, se $h \in \mathscr{L}^1_{loc}(\Omega)$ e vale:

$$\int_\Omega \frac{\partial^{\alpha_1+\cdots+\alpha_n} g}{\partial x^{1\alpha_1} \cdots \partial x^{n\alpha_1 n}} f\, d^n x = (-1)^{\alpha_1+\cdots+\alpha_n} \int_\Omega gh\, d^n x\,, \quad \text{per ogni } g \in C_0^\infty(\Omega)\,.$$

$\diamondsuit$

Procedendo come prima si verifica che se h è la derivata in senso proprio (o *forte*) dell'ordine detto di f allora è anche derivata in senso debole di f. Viceversa se h è la derivata in senso proprio dell'ordine detto di f e vale $f \in C^{\alpha_1+\cdots+\alpha_n}(\Omega)$ allora h è la derivata in senso proprio (o forte) dell'ordine detto di f.
Deve essere chiaro che le soluzioni in senso debole di $\Delta\varphi = f$ non sono altro che le soluzioni di tale equazione quando il laplaciano è inteso come operatore differenziale in senso debole.
La procedura moderna per risolvere l'equazione di Poisson (aggiungendo dati al bordo se assegnati) è decomposta in due passi:

(i) cercare, se esiste, una soluzione *in senso debole* dell'equazione considerata,

(ii) dimostrare, se possibile, che tale soluzione o una sua ri-definizione su insiemi di misura nulla, è soluzione anche in senso ordinario. Cioé , nella terminologia moderna, è soluzione in *senso forte.*

La tecnologia matematica per determinare le soluzioni in senso debole di una qualsiasi equazione differenziale a derivate parziali, non necessariamente quella di Poisson, è stata enormemente sviluppata e costituisce un ramo importantissimo dell'analisi funzionale moderna. Essa si basa sulla definizione e sull'uso delle proprietà di funzioni che appartengono ad opportuni spazi funzionali detti *spazi di Sobolev* [4], sui quali si rimanda a corsi più avanzati [5]. Diremo solo che vale la seguente definizione.

Definizione A.3. Se $\Omega \subset \mathbb{R}^n$ è un aperto non vuoto e $k = 0, 1, 2, \ldots$ è fissato, lo **spazio di Sobolev** $W^k(\Omega)$ è costituito dalle funzioni $f : \Omega \to \mathbb{C}$ per le quali esistono le derivate in senso debole fino all'ordine k, sono funzioni misurabili, e soddisfano:

$$\int_\Omega \left| \frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\partial x^{1\alpha_1} \cdots \partial x^{n\alpha_1 n}} \right|^2 d^n x < +\infty \quad \text{per } \alpha_1 = 0, 1, \ldots, \text{ con } i = 1, \ldots, n \text{ tali che } \textstyle\sum_i \alpha_i = k.$$

$\diamondsuit$

Nel caso specifco di equazioni lineari di tipo ellittico, esistono teoremi di regolarità che stabiliscono sotto quali ipotesi soluzioni deboli sono soluzioni in senso forte: sono i noti *teoremi di regolarità ellittica.* Tali teoremi sono stati estesi anche a casi più generali (equazioni ipoellittiche) in particolare dal matematico L. Hörmander. I teoremi fondamentali sono due: Il *Lemma di Sobolev* [4] ed il *teorema di regolarità ellittica di Friedrichs* [2]. Il primo, nella versione più elementare (la tesi vale infatti imponendo solo che le derivate *non miste*, fino all'ordine r, esistano in senso debole e siano funzioni a quadrato sommabile [4]), afferma quanto segue.

Teorema A.1. (**Lemma di Sobolev**.) *Se $f \in W^r(\Omega)$ con Ω aperto non vuoto in $\mathbb{R}^n$, allora f differisce su un insieme di misura nulla da una funzione che appartiene a $C^p(\Omega)$ con p dato dal più grande intero tale che $0 \le p < r + \frac{n}{2}$. (In particolare quindi, se $f \in W^r(\Omega)$ è continua,*

allora $f \in C^p(\Omega))$. $\diamondsuit$

Il secondo, nella versione più semplice (si generalizza infatti ad operatori ellittici di ordine superiore al secondo) si enuncia come segue.

Teorema A.2. (**Teorema di regolarità ellittica di Friedrichs**.) *Sia* $P\varphi = f$ *un'equazione differenziale alle derivate parziali, lineare del secondo ordine su* $\Omega \subset \mathbb{R}^n$ *aperto non vuoto, a coefficienti dati da funzioni di classe* $C^2(\Omega)$*, dove:*

$$P\varphi = \sum_{i,j=1}^{n} \frac{\partial}{\partial x^i}\left(A^{ij}(x)\frac{\partial}{\partial x^j}\varphi\right) .$$

Si supponga che la matrice caratteristica del sistema $A = A(x)$ *associata all'operatore* P *soddisfi la condizione di* **forte ellitticità** *, per qualche* $C > 0$*:*

$$\sum_{i,j=1}^{n} A^{ij}(x)y_i y_j \geq C||y||^2 , \quad \textit{per ogni } x \in \Omega \textit{ e ogni } y \in \mathbb{R}^n.$$

Se $\varphi : \Omega \to \mathbb{R}$ *risolve in senso debole l'equazione* $P\varphi = f$*, cioè :*

$$\int_\Omega \varphi P g \, d^n x = \int_\Omega f g \, d^n x , \quad \textit{per ogni } g \in C_0^\infty(\Omega)$$

e $f \in W^k(\Omega)$ *per qualche* $k = 0, 1, \ldots$ *fissato, allora* $\varphi \in W^{k+2}(\Omega)$. $\diamondsuit$

A titolo di esempio, supponiamo che $f \in C^\infty(\Omega)$ e che $\varphi \in W^k(\Omega)$, per qualche $k = 0, 1, \ldots$ fissato, sia soluzione in senso debole di $P\varphi = f$. Se le ipotesi del teorema di Friedrichs sono valide, allora $\varphi \in W^\infty(\Omega)$. A sua volta però il Lemma di Sobolev prova che, più fortemente, modificando φ su un insieme di misura nulla ed ottenendo φ', si ha che $\varphi' \in C^\infty(\Omega)$. Mostriamo che questa nuova funzione φ' è in realtà una *soluzione in senso forte dell'equazione* $P\varphi = f$. Infatti, per ipotesi vale:

$$\int_\Omega \varphi P g \, d^n x = \int_\Omega f g \, d^n x$$

per ogni $g \in C_0^\infty(\Omega)$. La ridefinizione di φ in φ' non altera l'identità scritta sopra visto che le due funzioni differiscono su un insieme di misura nulla. Possiamo allora usare la derivazione per parti ottenendo che:

$$\int_\Omega (P\varphi') g \, d^n x = \int_\Omega f g \, d^n x$$

per ogni $g \in C_0^\infty(\Omega)$. Procedendo come mostrato sopra, l'arbitrarietà di $g \in C_0^\infty(\Omega)$ implica che $P\varphi' = f$ sia valida in senso forte su Ω. Per cui la funzione φ, ridefinita come una funzione $\varphi' \in C^\infty(\Omega)$, grazie al Lemma di Sobolev, soddisfa in senso forte l'equazione differenziale.

Appendice B

Limite e derivazione sotto il segno integrale e di serie dalla teoria della misura.

In questa appendice dimostreremo alcuni teoremi che consentono di scambiare il simbolo di integrale e di serie con quello di derivata, facendo uso essenzialmente del teorema della convergenza dominata di Lebesgue. Il caso della serie sarà visto come sottocaso del caso integrale, in riferimento alla misura che conta i punti su $\mathbb{N}$. I teoremi che daremo sono quindi, nel caso dell'integrale, riferiti ad una generica misura positiva assegnata su uno spazio misurabile. Per comodità , riportiamo (senza dimostrazioni per altro elementari [1]) i tre teoremi classici di scambio del simbolo limite e derivata con quelli di serie ed integrale.

Teorema B.1. *Sia $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni continue, tutte definite sul compatto $K \subset \mathbb{R}^m$ ed a valori in $\mathbb{R}$. Se $\{f_n\}_{n\in\mathbb{N}}$ converge uniformemente per $n \to +\infty$, allora si ha, dove gli integrali indicano quelli valutati nel senso di Riemann,*

$$\lim_{n\to+\infty} \int_K f_n d^m x = \int_K \left(\lim_{n\to+\infty} f_n \right) d^m x \, .$$

$\diamondsuit$

Teorema B.2. *Sia $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni derivabili con continuità sull'intervallo chiuso $[a, b] \subset \mathbb{R}$. Se la serie associata alle f_n e quella associata alle derivate di tali funzioni convergono uniformemente su $[a, b]$, allora la somma della serie delle f_n è derivabile e vale:*

$$\frac{d}{dt} \sum_{n=0}^{+\infty} f_n(t) = \sum_{n=0}^{+\infty} \frac{df_n(t)}{dt} \, , \quad \textit{per ogni } t \in [a, b].$$

$\diamondsuit$

Teorema B.3. *Si consideri una classe di funzioni a valori in* $\mathbb{R}$, $\{f_t\}_{t\in A}$, *definite sul compatto* $K \subset \mathbb{R}^n$ *e dove* $A \subset \mathbb{R}$ *è un intervallo aperto. Se valgono le condizioni seguenti:*

(i) $K \ni x \mapsto f_t(x)$ *è continua sul compatto* K *per ogni* $t \in I$,

(ii) esiste

$$A \times K \ni (t, x) \mapsto \frac{\partial f_t(x)}{\partial t}$$

ed è continua (congiuntamente nelle due variabili), allora la funzione (tutti gli integrali sono è indifferentemente intesi nel senso di Riemann o Lebesgue):

$$I \ni t \mapsto \int_K f_t(x) d^n x$$

è di classe $C^1(I)$ *e vale l'identità :*

$$\frac{d}{dt} \int_K f_t(x) d^n x = \int_K \frac{\partial f_t(x)}{\partial t} d^n x \,. \tag{B.1}$$

$\diamondsuit$

Note.

(1) Nel seguito, quando ci riferiremo a serie *assolutamente convergenti* (cioè la serie dei valori assoluti converge ad un numero finito) ne indicheremo la somma con

$$\sum_{n\in\mathbb{N}} a_n \,,$$

dove non è specificato l'ordine con cui si esegue la somma. Ciò non è scorretto dato che le serie assolutamente convergenti possono essere riordinate a piacimento senza alterarne la somma [1].

(2) Nel seguito la misura dell'integrale di Lebesgue sarà ancora indicata con $d^n x$, che è lo stesso simbolo usato nell'integrale di Riemann. Questa notazione non genererà confusione in quanto nelle situazioni in cui compariranno entrambi gli integrali essi coincideranno in valore.

B.1 Teoremi della convergenza monotona e dominata.

Se (X, Σ, μ) è uno spazio misurabile, dove X è l'insieme ambiente, Σ una σ-algebra su X e $\mu : \Sigma \to [0, +\infty) \cup \{+\infty\}$ una misura positiva su X, lo strumento fondamentale per ottenere i teoremi di scambio tra simbolo di integrale e quello di limite/derivata è il ben noto *teorema della convergenza dominata di Lebesgue* [3]. Per completezza prima citiamo anche il cosiddetto *teorema della convergenza monotona* [3] dato che lo abbiamo usato nelle dispense. Nel seguito $\mathscr{L}^1(X, \Sigma, \mu)$ indicherà lo spazio delle funzioni misurabili su X integrabili rispetto a μ. Nel caso in cui $X = A \subset \mathbb{R}^n$ è Lebesgue-misurabile (cioè appartiene alla σ-algebra di lebesgue) e μ è la misura di Lebesgue $d^n x$ su $\mathbb{R}^n$, scriveremo semplicemente $\mathscr{L}(A)$.

Teorema B.4. (**Convergenza monotona**.) *In riferimento allo spazio misurabile* (X, Σ, μ), *sia* $\{f_n\}_{n\in\mathbb{N}}$ *una successione di funzioni definite su* x *che siano* Σ*-misurabili ed* μ*-integrabili. Se valgono le due condizioni:*

(i) $f_n(x) \in [0, +\infty) \cup \{+\infty\}$ *per ogni* $n \in \mathbb{N}$,

(ii) $f_n(x) \leq f_{n+1}(x)$ *quasi ovunque su* X *e per ogni* $n \in \mathbb{N}$,

allora, posto $f(x) := \lim_{n\to+\infty} f_n(x)$, *vale*

$$\int_X f d\mu = \lim_{n\to+\infty} \int_X f_n d\mu \,.$$

$\diamondsuit$

Passiamo al teorema della convergenza dominata.

Teorema B.5. (**Convergenza dominata**.) *In riferimento allo spazio misurabile* (X, Σ, μ), *sia* $\{f_n\}_{n\in\mathbb{N}}$ *una successione di funzioni definite su* x *che siano* Σ*-misurabili ed* μ*-integrabili. Se valgono le due condizioni:*

(i) esiste $f(x) := \lim_{n\to+\infty} f_n(x) \in \mathbb{C}$ *quasi ovunque rispetto a* μ *su* X,

(ii) esiste $g \in \mathscr{L}^1(X, \mu)$ *con* $g \geq 0$ *quasi ovunque su* X *e tale che:*

$$|f_n(x)| \leq g(x)\,, \quad \textit{quasi ovunque su } X, \textit{ per ogni } n \in \mathbb{N}\,,$$

allora valgono i seguenti fatti.

(a) $f \in \mathscr{L}^1(X, \Sigma, \mu)$,

(b) $\int_X |f| d\mu \leq \int_X g d\mu$,

(c) $\int_X |f_n - f| d\mu \to 0$ *per* $n \to +\infty$,

(d) $\int_X f_n d\mu \to \int_X f d\mu$ *per* $n \to +\infty$ *ovvero, in altre paroleunicitaperiodica*

$$\lim_{n\to+\infty} \int_X f_n d\mu = \int_X \left(\lim_{n\to+\infty} f_n \right) d\mu \,. \tag{B.2}$$

$\diamondsuit$

Osservazioni B.1.

(1) Il controesempio classico per il teorema della convergenza dominata è quello in cui si lavora in $\mathscr{L}^1(\mathbb{R})$ con le gaussiane di centro $n \in \mathbb{N}$: $f_n(x) := e^{-(x-n)^2}$. Vale, a causa dell'invarianza per traslazioni della misura di Lebesgue:

$$\int_{\mathbb{R}} e^{-(x-n)^2} dx = \int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}\,.$$

Pertanto

$$\lim_{n\to+\infty} \int_{\mathbb{R}} f_n(x) dx = \sqrt{\pi}\,.$$

D'altra parte, se $x \in \mathbb{R}$ è fissato, si ha immediatamente che

$$\lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} e^{-(x-n)^2} = 0\,.$$

Concludiamo che la (B.2) non può valere, dato che il primo membro varrebbe $\sqrt{\pi}$ nel caso in esame, mentre il secondo membro varrebbe 0.
La spiegazione del fatto che non si possa applicare il teorema della convergenza dominata è evidente. Non può esistere una funzione g che soddisfa le ipotesi: in ciascun punto dovrebbe maggiorare ogni gaussiana traslata arbitrariamente verso destra. Si può dimostrare che questo implica che g non possa essere integrabile. In realtà l'inesistenza di g segue immediatamente dal fatto che non vale (B.2) come abbiamo direttamente appurato.
(2) La non esistenza di una funzione g che soddisfi le ipotesi del teorema della convergenza dominata, *non* implica automaticamente che non valga la (B.2), visto che il teorema della convergenza dominata fornisce condizioni *sufficienti*, ma non necessarie affinché valga la (B.2).
(3) Il teorema della convergenza dominata di Lebesgue include il caso in cui si esaminano delle serie. In questo caso (X, Σ, μ) è costruito in questo modo: $X = \mathbb{N}$, Σ è $\mathscr{P}(\mathbb{N})$: l'insieme delle parti di $\mathbb{N}$, e $\mu = \delta$, la *misura che conta i punti*: $\delta(N) =$ numero di elementi di $N \subset \mathbb{N}$. Le funzioni misurabili sono le successioni $\{a(m)\}_{m\in\mathbb{N}} \subset \mathbb{C}$. Infine le funzioni integrabili sono le successioni tali che

$$\sum_{m\in\mathbb{N}} |a(m)| < +\infty\,.$$

In altre parole le funzioni integrabili non sono altro che le successioni che producono serie assolutamente convergenti. Si osservi che in tal caso, come ben noto, la somma della serie $\sum_{m\in\mathbb{N}} a(m)$ *non* dipende dall'ordinamento con cui si esegue la somma.
In questo caso, il teorema di Lebesgue fornisce condizioni sufficienti per poter scambiare il simbolo di somma con quello di limite,

$$\lim_{n\to+\infty} \sum_{m\in\mathbb{N}} a_n(m) = \sum_{m\in\mathbb{N}} \left(\lim_{n\to+\infty} a_n(m) \right)\,,$$

quando si ha *una classe* di successioni $\{\{a_n(m)\}_{m\in\mathbb{N}}\}_{n\in\mathbb{N}}$ per cui $a_n(m) \to a(m)$ se $n \to +\infty$.

Il teorema della convergenza dominata benché molto più generale (vale con ogni tipo di misura e lavora anche su domini di misura infinita), fornisce una dimostrazione alternativa del classico teorema B.1 riferito all'integrale di Riemann:

Proposizione B.1. *Sia $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni continue, tutte definite sul compatto $K \subset \mathbb{R}^m$ ed a valori in $\mathbb{R}$ (o $\mathbb{C}$). Se $\{f_n\}_{n\in\mathbb{N}}$ converge uniformemente per $n \to +\infty$, allora si ha (dove gli integrali indicano quelli valutati nel senso di Riemann o Lebesgue indifferentemente)*

$$\lim_{n\to+\infty} \int_K f_n d^m x = \int_K \left(\lim_{n\to+\infty} f_n \right) d^m x\,.$$

$\diamondsuit$

Dimostrazione. Essendo le funzioni continue su un compatto, l'integrale di esse secondo Riemann coincide con quello di Lebesgue [3]. Dato che la successione di funzioni continue converge uniformemente, il limite di tali funzioni sarà ancora una funzione continua $f : K \to \mathbb{R}$ (per cui integrabile secondo Riemann e Lebesgue e i due integrali coincideranno nuovamente). Sia $M = \max_K |f(x)|$, che esiste finito in virtù del fatto che K è compatto e f continua. In virtù della convergenza uniforme, se $\epsilon > 0$ esisterà $N_\epsilon \in \mathbb{N}$ tale che, se $n > N_\epsilon$

$$\max_K |f_n - f| \leq \epsilon \, .$$

Quindi in particolare:

$$-\epsilon - M < f(x) - \epsilon < f_n(x) < f(x) + \epsilon < M + \epsilon \, , \quad \text{per ogni } x \in K.$$

In particolare, per $n > N_\epsilon$:

$$|f_n(x)| < M + \epsilon \, , \quad \text{per ogni } x \in K.$$

Possiamo allora applicare il teorema convergenza dominata con $g(x) := M + \epsilon$ costantemente su K, provando la (B.2) che coincide con la nostra tesi. Si noti che g è per costruzione in $\mathscr{L}^1(K)$ dato che $\int_K |g| d^m x = (M + \epsilon) Vol(K)$ dove $Vol(K)$ è la misura di Lebesgue (coincidente con quella di Peano-Jordan-Riemann) di K che esiste ed è finita essendo K un compatto. $\square$

Per studiare il problema di scambiare il simbolo di derivata con quello di integrale e di serie abbiamo bisogno di una formulazione leggermente modificata del teorema della convergenza dominata.

Teorema B.6. (**Convergenza dominata 2**.) *Se si generalizzano le ipotesi del teorema B.5 rimpiazzando la successione $\{f_n\}_{n\in\mathbb{N}}$ con una famiglia di funzioni $\{f_t\}_{t\in A} \subset \mathscr{L}^1(X, \mu)$ dove $A \subset \mathbb{R}^m$ è un intorno aperto di $t_0 \in \mathbb{R}$, in modo tale che:*

(i) esiste $f(x) := \lim_{t\to t_0} f_t(x) \in \mathbb{C}$ quasi ovunque rispetto a μ su X,

(ii) esiste $g \in \mathscr{L}^1(X, \mu)$ con $g \geq 0$ quasi ovunque su X e tale che:

$$|f_t(x)| \leq g(x) \, , \quad \textit{quasi ovunque su } X, \textit{ per ogni } t \in A \, ,$$

gli enunciati (a), (b), (c) e (d) del teorema B.5 sono ancora validi sostituendo ovunque $\lim_{n\to+\infty}$ *con* $\lim_{t\to t_0}$.$\diamondsuit$

Dimostrazione. La tesi è immediata conseguenza del teorema della convergenza dominata e del noto risultato di analisi che afferma che: *una funzione tra due spazi metrici $f : X_1 \to X_2$ ammette limite $y \in X_2$ per $x \to x_0 \in X_1$ se e solo se ammette tale limite* per successioni*, ovvero, per ogni successione $\{x_n\}_{n\in\mathbb{N}} \subset X_1$ vale*

$$\lim_{n\to+\infty} f(x_n) = y \, .$$

Nel caso in esame $X_1 = A \subset \mathbb{R}^m$ e $X_2 = \mathbb{C}$ dotati delle distanze standard. $\square$

Questa formulazione del teorema della convergenza dominata ha diverse conseguenze immediate sulle serie di funzioni. A titolo di esempio citiamo il seguente corollario che si dimostra subito lavorando sullo spazio con misura $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \delta)$ già visto in un precedente esempio.

Proposizione B.2. *Sia $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni a valori in $\mathbb{C}$ (o $\mathbb{R}$) definite sull'insieme $A \subset \mathbb{R}^n$ tale che, per ogni $t \in A$ valga*

$$\sum_{n\in\mathbb{N}} |f_n(t)| < +\infty \, .$$

e, per ogni $n \in \mathbb{N}$ esiste finito

$$f_n = \lim_{t\to t_0} f_n(t)$$

dove t_0 è un punto di accumulazione di A (includendo valori infiniti come casi limite). Se esiste $\{g_n\}_{n\in\mathbb{N}}$ con $g_n \geq 0$ e $\sum_{n\in\mathbb{N}} g_n < +\infty$ tale che:

$$|f_n(t)| \leq g_n \, , \quad \textit{per ogni } t \in A \, ,$$

allora, per ogni $t_0 \in A$:

$$\lim_{t\to t_0} \sum_{n\in\mathbb{N}} f_n(t) = \sum_{n\in\mathbb{N}} \lim_{t\to t_0} f_n(t) = \sum_{n\in\mathbb{N}} f_n \, .$$

e

$$\sum_{n\in\mathbb{N}} |f_n| < +\infty \, .$$

$\diamondsuit$

B.2 Derivazione sotto il segno di integrale e di serie.

Possiamo allora enunciare e provare il teorema fondamentale riguardante la derivazione sotto il segno di integrale per una misura positiva generale.

Teorema B.7. (**Derivazione sotto il segno di integrale**.) *In riferimento allo spazio misurabile (X, Σ, μ), si consideri una famiglia di funzioni $\{f_t\}_{t\in A} \subset \mathscr{L}^1(X, \mu)$ dove $A \subset \mathbb{R}^m$ è un insieme aperto e $t = (t^1, \ldots, t^m)$. Se valgono le seguenti due condizioni:*

(i) per un certo valore k in $\{1, 2, \ldots, n\}$ esistono le derivate:

$$\frac{\partial h_t(x)}{\partial t^k} \, , \quad \textit{per ogni } x \in X \textit{ e } t \in A$$

(ii) esiste $g \in \mathscr{L}^1(X, \mu)$ con $g \geq 0$ quasi ovunque su X e tale che:

$$\left| \frac{\partial h_t(x)}{\partial t^k} \right| \leq g(x) \, , \quad \textit{quasi ovunque su } X, \textit{ per ogni } t \in A \, ,$$

allora valgono i seguenti fatti.
(a) $X \ni x \mapsto \frac{\partial h_t}{\partial t^k} \in \mathscr{L}^1(X, \Sigma, \mu)$,
(b) *si possono scambiare i simboli di integrale con quello di derivata per ogni* $t \in A$*:*

$$\frac{\partial}{\partial t^k} \int_X h_t(x) d\mu(x) = \int_X \frac{\partial h_t(x)}{\partial t^k} d\mu(x) \,. \tag{B.3}$$

Se infine:

(iii) per una fissata g *la condizione in (ii) vale contemporaneamente per tutti i valori di* $k = 1, 2, \ldots, m$*, quasi ovunque in* $x \in X$ *e tutte le funzioni (per ogni* $t \in A$ *fissato):*

$$A \ni t \mapsto \frac{\partial h_t(x)}{\partial t^k}$$

sono continue, allora
(c) *la funzione:*

$$A \ni t \mapsto \int_X h_t(x) d\mu(x)$$

è in $C^1(A)$*.* $\diamondsuit$

Dimostrazione. Notiamo che, per ogni $t \in A$, le funzioni $X \ni x \mapsto \frac{\partial h_t}{\partial t^k}$ sono sicuramente misurabili essendo limite (usando la definizione di derivata come limite del rapporto incrementale) di funzioni misurabili. Inoltre sono μ-integrabili dato che sono maggiorate, in valore assoluto, da una funzione integrabile per l'ipotesi (ii). Fissiamo $t_0 \in A$. Considerando il rapporto incrementale si ha, dove scriviamo, un po' impropriamente, $t_0 + \tau^k$ al posto di $(t_0^1, \ldots, t_0^{k-1}, t_0^k + \tau^k, t_0^{k+1}, \ldots, t_0^m)$:

$$\left.\frac{\partial}{\partial t^k}\right|_{t_0} \int_X h_t(x) d\mu(x) = \lim_{\tau^k \to 0} \int_X \frac{h_{t_0+\tau^k}(x) - h_{t_0}(x)}{\tau^k} d\mu(x) \,.$$

D'altra parte, per il teorema di Lagrange (restringendosi a lavorare in un intorno aperto e convesso di t_0) e tenendo conto dell'ipotesi (ii) abbiamo:

$$\left| \frac{h_{t_0+\tau^k}(x) - h_{t_0}(x)}{\tau^k} \right| = \left| \left.\frac{\partial h_t(x)}{\partial t}\right|_{t(\tau^k, x)} \right| \leq g(x) \,,$$

dove $t(\tau^k, x)$ è un punto che si trova tra t_0 e $(t_0^1, \ldots, t_0^{k-1}, t_0^k + \tau^k, t_0^{k+1}, \ldots, t_0^m)$ sul segmento che unisce tale coppia di punti. Possiamo allora applicare il teorema B.6 per:

$$f_{\tau^k}(x) := \frac{h_{t_0+\tau^k}(x) - h_{t_0}(x)}{\tau^k} \,,$$

ottendo che esiste il limite

$$\lim_{\tau^k \to 0} \int_X \frac{h_{t_0+\tau^k}(x) - h_{t_0}(x)}{\tau^k} d\mu(x) =: \left.\frac{\partial}{\partial t^k}\right|_{t_0} \int_X h_t(x) d\mu(x) \,,$$

e vale:

$$\int_X \lim_{\tau^k \to 0} \frac{h_{t_0+\tau^k}(x) - h_{t_0}(x)}{\tau^k} d\mu(x) =: \int_X \left. \frac{\partial h_t(x)}{\partial t^k} \right|_{t_0} d\mu(x) \,.$$

La tesi è stata provata per quanto riguarda (a) e (b). La dimostrazione di (c) è immediata: dal teorema B.6 tenendo conto dell'ipotesi (ii) si ha che ogni funzione, per $k = 1, \ldots, m$,

$$A \ni t \mapsto \frac{\partial}{\partial t^k} \int_X h_t(x) d\mu(x)$$

è continua, da cui la tesi. □

Osservazioni B.2. Nell'ipotesi di validità di (c) la funzione

$$A \ni t \mapsto \int_X h_t(x) d\mu(x)$$

risulta essere $C^1(A)$ e quindi differenziabile su A come funzione di più variabili.

Il teorema B.7 riproduce, come sottocaso il teorema classico B.3. È però fondamentale notare che il teorema B.7 ha validità molto più generale, in quanto lavora con l'integrale di Lebesgue o qualsiasi altra misura positiva, non richiede la continuità delle derivate nelle variabili congiuntamente: lo spazio X su cui si integra nel teorema B.7 potrebbe non essere uno spazio topologico e può anche avere misura infinita.

Proposizione B.3. *Si consideri una classe di funzioni a valori in $\mathbb{R}$, $\{f_t\}_{t \in A}$, definite sul compatto $K \subset \mathbb{R}^n$ e dove $A \subset \mathbb{R}$ è un intervallo aperto. Se valgono le condizioni seguenti:*

(i) $K \ni x \mapsto f_t(x)$ è continua sul compatto K per ogni $t \in I$,

(ii) esiste

$$A \times K \ni (t, x) \mapsto \frac{\partial f_t(x)}{\partial t}$$

ed è continua (congiuntamente nelle due variabili), allora la funzione (tutti gli integrali sono è indifferentemente intesi nel senso di Riemann o Lebesgue):

$$I \ni t \mapsto \int_K f_t(x) d^n x$$

è di classe $C^1(I)$ e vale l'identità :

$$\frac{d}{dt} \int_K f_t(x) d^n x = \int_K \frac{\partial f_t(x)}{\partial t} d^n x \,. \tag{B.4}$$

♢

Dimostrazione. Dato che si integra su un compatto, la richiesta di continuità in x di f e della sua derivata assicura che gli integrali di Riemann considerati esistano e coincidano con quelli

di Lebesgue. Per ogni $t_0 \in A$ e sia $A_0 \subset A$ un intervallo aperto a chiusura compatta, con $t_0 \in A_0$. La dimostrazione di (B.4) e del fatto che la funzione integrale appartenga a $C^1(I)$ è un'immediata conseguenza del fatto che la funzione

$$\overline{A_0} \times K \ni (t, x) \mapsto \frac{\partial f_t(x)}{\partial t}$$

essendo continua sarà limitata, in valore assoluto, da qualche costante $M > 0$. Pertanto possiamo applicare il teorema B.7 usando come funzione g quella che vale costantemente M su K. $\square$

Infine, per quanto riguarda le serie di funzioni, il teorema B.7 si specializza alla seguente proposizione lavorando sullo spazio con misura $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \delta)$.

Proposizione B.4. *Si consideri una successione di funzioni $\{f_n\}_{n\in\mathbb{N}}$ dove $f_n : A \to \mathbb{C}$ per ogni $n \in \mathbb{N}$, con $A \subset \mathbb{R}^m$ insieme aperto, $t = (t^1, \ldots, t^m)$ e si assuma che valga la convergenza assoluta della serie associata alle f_n:*

$$\sum_{n\in\mathbb{N}} |f_n(t)| < +\infty \,, \quad \textit{per ogni } t \in A \,,$$

per cui in particolare anche la serie delle f_n (senza valore assoluto) converge per ogni valore di t. Se sono verificate le seguenti due condizioni:

(i) per un certo valore k in $\{1, 2, \ldots, n\}$ esistono le derivate:

$$\frac{\partial f_n(t)}{\partial t^k} \,, \quad \textit{per ogni } n \in \mathbb{N} \textit{ e } t \in A$$

(ii) esiste una successione $\{g_n\}_{n\in\mathbb{N}}$ con $0 \le g_n$ costante, con $\sum_{n\in\mathbb{N}} g_n < +\infty$, e tale che:

$$\left| \frac{\partial f_n(t)}{\partial t^k} \right| \le g_n \,, \quad \textit{per ogni } t \in A \textit{ e } n \in \mathbb{N} \,,$$

allora valgono i seguenti fatti.

(a) $\sum_{n\in\mathbb{N}} |\frac{\partial f_n(t)}{\partial t^k}| < +\infty$, *per cui in particolare anche la serie delle derivate f_n rispetto a t^k (senza valore assoluto) converge per ogni valore di t,*

(b) *si possono scambiare i simboli di integrale con quello di somma per ogni $t \in A$:*

$$\frac{\partial}{\partial t^k} \sum_{n\in\mathbb{N}} f_n(t) = \sum_{n\in\mathbb{N}} \frac{\partial f_n(t)}{\partial t^k} \,. \tag{B.5}$$

Se infine:

(iii) per una fissata successione di costanti $\{g_n\}_{n\in\mathbb{N}}$ la condizione in (ii) vale contemporaneamente per tutti i valori di $k = 1, 2, \ldots, m$, e tutte le funzioni (per ogni $t \in A$ fissato):

$$A \ni t \mapsto \frac{\partial f_n(t)}{\partial t^k}$$

sono continue, allora
(c) *la funzione:*

$$A \ni t \mapsto \sum_{n\in\mathbb{N}} f_n(t)$$

è in $C^1(A)$. $\diamondsuit$

Diversamente dal teorema classico B.2 di derivazione sotto il segno di serie, questo teorema non richiede (eccetto che per la validità dell'ultimo punto) che le derivate delle funzioni nella serie siano funzioni continue. Non è nemmeno richiesta la convergenza uniforme della serie delle funzioni non derivate. La condizione (ii) in ogni caso assicura tra l'altro la convergenza uniforme della serie delle derivate (per il noto teorema di Weierstrass della convergenza totale [1].)

Bibliografia

[1] E. Giusti, *Analisi Matematica*, vol 1 e 2, Bollati-Boringhieri, Torino (2003).

[2] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, vol. II, Academic Press, New York, (1975).

[3] W. Rudin, *Analisi reale e complessa*, Bollati-Boringhieri, Torino (1982)

[4] W. Rudin, *Functional Aanlysis*, Mc Graw Hill, Boston (1991).

[5] M.E. Taylor, *Partial Differential Equations*, vol. I, II and III, New York, Springer-Verlag (1996).

[6] A.G. Svesnokov e A.N: Tichonov, *Teoria delle funzioni di una variabile complessa*, Editori Riuniti, Roma (1984).

[7] V. S. Vladimirov, *Equations of Mathematical Physics*, Mir, Moscow (1984).

La Struttura Matematica della Meccanica Quantistica e teoria spettrale in spazi di Hilbert.

Dispense per il Corso di Dottorato di Ricerca in Matematica

Anno accademico 2003-2004

Valter Moretti[1]
Dipartimento di Matematica
Facoltà di Scienze M.F.N
Università di Trento

[1]E-mail: moretti@science.unitn.it

Indice

1 Introduzione **4**
1.1 La MQ come teoria matematica. 4
1.2 La MQ nel panorama della Fisica attuale. 6
1.3 Convenzioni generali e prerequisiti fisico matematici. 7

I Elementi di teoria degli operatori limitati su spazi di Hilbert **9**

2 Alcune nozioni e teoremi generali nella teoria degli spazi normati e di Banach. **10**
2.1 Spazi normati, di Banach e algebre. 10
2.2 Operatori, spazi di operatori, norme di operatori. 16
2.3 I tre teoremi fondamentali negli spazi di Banach e le topologie deboli. 22
2.4 Proiettori. 31
2.5 *Norme equivalenti. 33

3 Spazi di Hilbert ed operatori limitati. **35**
3.1 Nozioni elementari, teorema di Riesz e riflessività. 35
3.2 Basi hilbertiane. 43
3.3 Aggiunto di un operatore limitato, operatori autoaggiunti, normali, unitari, positivi e proiettori ortogonali, C^*-algebre di operatori limitati. 54
3.4 Proiettori ortogonali. 60
3.5 Radici quadrate di operatori positivi e decomposizione polare di operatori limitati. 63
3.6 *La trasformata di Fourier-Plancherel. 69

4 Proprietà elementari degli operatori compatti, di Hilbert-Schmidt e di classe traccia nello spazio di Hilbert. **79**
4.1 Generalità sugli operatori compatti in spazi normati, di Banach e Hilbert. 79
4.2 Operatori compatti in spazi di Hilbert. 81
4.3 Operatori di Hilbert-Schmidt. 90
4.4 Operatori di classe traccia (o nucleari). 97

II Il formalismo della Meccanica Quantistica e la Teoria Spettrale 104

5 Brevi cenni di fenomenologia dei sistemi quantistici e di Meccanica Ondulatoria. 105
5.1 Generalità sui sistemi quantistici. 105
5.2 Alcune proprietà particellari delle onde elettromagnetiche. 107
5.2.1 Effetto Fotoelettrico. 107
5.2.2 Effetto Compton. 108
5.3 Cenni di Meccanica ondulatoria. 110
5.3.1 Onde di de Broglie. 110
5.3.2 Funzione d'onda di Schrödinger e interpretazione probabilistica di Born. . 111
5.4 Principio di indeterminazione di Heisenberg. 113
5.5 Le grandezze compatibili ed incompatibili. 115

6 La costruzione del formalismo matematico della MQ: il teorema di Gleason. 117
6.1 Le idee che stanno alla base dell'interpretazione standard della fenomenologia quantistica. 117
6.2 Stati classici come misure di probabilità sulla σ-algebra delle proposizioni elementari. 119
6.2.1 Stati e misure di Borel 119
6.2.2 Proposizioni e insiemi. 120
6.2.3 Interpretazione insiemistica dei connettivi logici. 121
6.2.4 Proposizioni "infinite" e grandezze fisiche. 121
6.2.5 Il reticolo distributivo, limitato, ortocomplementato e σ-completo delle proposizioni elementari. 124
6.3 Le proposizioni relative a sistemi quantistici come insiemi di proiettori ortogonali. 126
6.3.1 Reticoli di proiettori ortogonali su spazi di Hilbert. 126

7 Appendice A: Relazioni d'ordine e Insiemi Parzialmente Ordinati. 130

8 Appendice B: Dimostrazione di alcuni teoremi. 132

9 Appendice C: Integrale di funzioni a valori spazi di Banach. 133

Ringraziamenti.

Vorrei ringraziare il dr. Riccardo Aramini ed il dr. Gianmarco Bramanti per avere letto con cura queste dispense segnalandomi diversi errori di vario genere che ora sono stati corretti.

Capitolo 1

Introduzione

In queste dispense ci occuperemo di esporre i fondamenti matematici della Meccanica Quantistica (MQ) in modo matematicamente rigoroso. Escludendo alcune prime parti della dispensa, la fenomenologia fisica sarà lasciata sullo sfondo per concentrarci sugli aspetti logico-formali della teoria. In ogni caso daremo delle esemplificazioni fisiche del formalismo per non perdere il contatto con la realtà. Le dispense sono divise in due parti. Nella prima parte ci occuperemo di enunciare la teoria generale degli operatori limitati in spazi di Hilbert (dando anche alcune nozioni valide per contesti più generali: spazi di Banach). Nella seconda parte svilupperemo la teoria spettrale per operatori normali limitati ed autoaggiunti non limitati in spazi di Hilbert contestualmente al formalismo matematico della Meccanica Quantistica.

1.1 La MQ come teoria matematica.

Dal punto di vista matematico la Meccanica Quantistica rappresenta una rara sintesi di eleganza matematica e profondità decrittiva del contesto fisico. La teoria usa essenzialmente tecniche di analisi funzionale lineare, ma con diverse intersezioni con la teoria della misura, la teoria della probabilità e la logica matematica.
Esistono (almeno) due possibili formulazioni matematiche della Meccanica Quantistica. La più elementare e più antica in ordine storico, dovuta essenzialmente a von Neumann (1932), è formulata usando il linguaggio della teoria degli spazi di Hilbert e della teoria spettrale degli operatori non limitati su tali spazi. La formulazione più recente ed avanzata, sviluppata dalla scuola del matematico Gel'fand nel tentativo di risolvere alcuni problemi fisico-matematici della teoria quantistica dei campi, è presentata nel linguaggio delle algebre astratte ($*$-algebre e C^*-algebre) costruite sul modello delle algebre di operatori definite e studiate dallo stesso von Neumann (oggi note come W^**algebre* o *algebre di von Neumann*), ma emancipandosi dalla struttura di spazio di Hilbert (vedi per es. il testo classico sulle algebre di operatori [7]). Tale formulazione ha il suo centro nel famoso *teorema GNS* [9, 7]. La seconda formulazione, in un senso molto specifico che non possiamo chiarire qui, può considerarsi un'estensione della prima formulazione anche per i

nuovi contenuti fisici introdotti. In particolare essa permette di dare un senso matematicamente preciso alla richiesta di *località* delle teorie di campo quantistiche relativistiche [9] e permette l'estensione delle teorie quantistiche di campo in spaziotempo curvo.
In queste dispense ci occuperemo unicamente della prima formulazione che ha comunque una complessità matematica notevole accompagnata da una notevole eleganza formale.
Uno strumento matematico fondamentale per sviluppare la MQ è il cosiddetto *teorema spettrale per operatori autoaggiunti* (generalmente non limitati) definiti in varietà lineari dense in uno spazio di Hilbert. Tale teorema, che può essere esteso al caso di operatori normali, fu dimostrato per la prima volta proprio da von Neumann nel suo libro fondamentale sulla struttura matematica della MQ [10] che può considerari una pietra miliare della fisica matematica oltre che della matematica pura del XX secolo[1]. Il legame tra MQ e *teoria spettrale* è dovuto al seguente fatto. Si vede in modo naturale nell'interpretazione standard della MQ che le grandezze fisiche misurabili su sistemi quantistici possono essere associate ad operatori autoaggiunti non limitati in un opportuno spazio di Hilbert. Lo spettro di ciascuno di questi operatori coincide con l'insieme dei valori assumibili dalla grandezza associata. La procedura di costruzione delle grandezze fisiche a partire dalle proprietà o proposizioni elementari del tipo "il valore della grandezza cade nell'intervallo $(a, b]$", che nello schema matematico adottato corrispondono a proiettori ortogonali, non è altro che una procedura di integrazione su una appropriata misura spettrale a valori di proiezione. Tale procedura ricorda molto da vicino il metodo per definire l'integrale di Lebesgue di una funzione misurabile. Il teorema spettrale in sostanza altro non è cheun metodo che permette di costruire operatori più complessi partendo da proiettori o, viceversa, di decomporre operatori in termini di misure a valori di proiezione.
La formulazione che daremo della teoria spettrale è sicuramente differente da quella originale di von Neumann che però conteneva tutti gli elementi fondamentali. Ancora oggi il testo di von Neumann (che è stato scritto nel lontano 1932) rivela una profondità impressionante specie nei problemi più difficili dell'interpretazione fisica del formalismo della MQ di cui, leggendo il libro, si evince che von Neumann era chiaramente conscio a differenza di molti dei suoi colleghi. Sarebbe interessante fare un paragone tra il testo di von Neumann e il più famoso (a nostro parere immeritatamente) testo di Dirac [11] sui fondamenti della MQ, cosa che lasciamo al lettore interessato. In ogni caso la profondità dell'impostazione data da von Neumann alla MQ comincia anche ad essere riconosciuta da chi si occupa di Fisica sperimentale ed in particolare di misure quantistiche [8].
Le cosiddette *Logiche Quantistiche* nascono dal tentativo di formulare la MQ dal punto di vista più radicale possibile attibuendo alla stessa logica usata nel trattare i sistemi quantistici alcune proprietà differenti da quelle che si usano nella formulazione usuale del calcolo proposizionale e nella teoria dell'interpretazione. Per esempio, sono usati più di due valori di verità e il reticolo booleano delle proposizioni è rimpiazzato da una struttura non distributiva più complessa. Nella prima formulazione della logica quantistica, oggi denominata *Logica Quantistica Standard*, proposta da Von Neumann e Birkhoff nel 1936, la struttura dell'algebra booleana delle propo-

[1] La definizione del concetto di pazio di Hilbert infinitodimensionale e gran parte della teoria generale degli spazi di Hilbert così come la conosciamo oggi sono anch'essi dovuti a von Neumann e alla sua formulazione della MQ.

sizioni era rimpiazzata con quella di un *reticolo ortocomplementato, ortomodulare e completo* che di fatto ha come modello l'insieme dei proiettori ortogonali su uno spazio di Hilbert ovvero, l'insieme dei sottospazi chiusi su cui proiettano i proiettori [5], unitamente ad alcune regole di composizione. È noto che, a dispetto della sua eleganza, tale modellizzazione contiene alcuni difetti quando si cercs di tradurla in termini operativi (o più precisamente *operazionali*) fisici. Noi non ci spingeremo pertanto così avanti, non svilupperemo alcuna logica quantistica [5], non assumeremo la struttura completa di *reticolo ortocomplementato, ortomodulare e completo* anche se assumeremo come assioma fondamentale quello che afferma che le proprietà elementari dei sistemi fisici quantistici sono descritte da un reticolo di proiettori ortogonali su uno spazio di Hilbert complesso.

1.2 La MQ nel panorama della Fisica attuale.

La *Meccanica Quantistica* – genericamente parlando la teoria della fisica del mondo atomico e sub atomico – insieme alla *Teoria della Relatività Speciale e Generale* (RSG) – genericamente parlando la teoria fisica della gravità, del mondo macroscopico e della cosmologia – costituiscono i due paradigmi attraverso i quali si è sviluppata la fisica del XX secolo e quella dell'inizio del secolo attuale. I due paradigmi si sono fusi in vari contesti dando luogo a teorie quantistiche relativistiche, in particolare alla *Teoria Quantistica Relativistica dei Campi* [3, 4], che ha avuto uno sviluppo impressionante con straordinari successi esplicativi e predittivi nel contesto della teoria delle particelle elementari e delle forze fondamentali. A titolo d'esempio tale teoria ha previsto, all'interno del cosiddetto *modello standard delle particelle elementari*, l'unifcazione della *forza debole* ed *elettromagnetica* che è poi stata confermata sperimentalmente alla fine degli anni '80 con un esperimento spettacolare al C.E.R.N. di Ginevra in cui si sono osservate le particelle Z_0 e $W^{\pm}$ previste dalla teoria dell'unificazione elettrodebole.
La previsione del valore di una grandezza fisica che è stata poi confermata con una delle maggiori precisioni di tutta la storia della Fisica si è avuta nell'elettrodinamica quantistica. Si tratta del valore del cosiddetto rapporto giromagnetico dell'elettrone g. Tale grandezza fisica è un numero puro. Il valore previsto dall'elettrodinamica quantistica per $a := g/2 - 1$ è:

$$0.001159652359 \pm 0.000000000282\,,$$

quello ottenuto sperimentalmente è risultato essere

$$0.001159652209 \pm 0.000000000031\,.$$

Molti fisici ritiengono che la MQ sia la teoria fondamentale dell'Universo (più profonda delle teorie relativistiche) anche per il fatto che risulta essere valida per scale lineari di lunghezza che variano in uno spettro di ampiezza impressionante: da $1m$ (condensati di Bose-Einstein) *almeno fino* a $10^{-16}m$ (interno dei nucleoni: quarks). La MQ ha avuto un enorme successo sia teorico che sperimentale anche nella scienza che studia la struttura della materia solida, nell'ottica, nell'elettronica, con diverse importantissime ricadute tecnologiche: ogni oggetto tecnolgico di uso

comune che sia moderatamente sofisiticato (giocattoli per i bambini, telefonini, telecomandi...) da contenere qualche elemento *semiconduttore* sfrutta proprietà quantistiche della materia.
Tornando ai due paradigmi scientifici del XX secolo – MQ e RSG – rimangono diversi punti oscuri in cui i due paradigmi sembrano venire in conflitto, in particolare il problema della cosiddetta "quantizzazione della gravità" e della struttura dello spaziotempo alle *scale di Planck* – $10^{-33}cm$, $10^{43}s$ – le scale di lunghezza e di tempo che si ottengono combinando le costanti fondamentali delle due teorie: la velocità della luce, la costante di gravitazione universale e la costante di Planck. La necessità di una struttura discontinua dello spaziotempo a scale ultramicroscopiche è suggerita anche da alcune difficoltà matematiche (ma anche concettuali) non completamente risolte dalla cosiddetta teoria della *Rinormalizzazione* quantistica, dovute all'apparire di infiniti che si incontrano nei calcoli dei processi dovuti alle interazioni fondamentali tra le particelle elementari. Tutti questi problemi hanno dato luogo a recenti ed importanti sviluppi teorici, che hanno avuto influenze nello sviluppo della stessa matematica pura, come la teoria delle (super) *stringhe* (e *brane*) e le varie versioni di *Geometria non commutativa*, prima fra tutte quella di A. Connes. La difficoltà nel decidere quale di queste teorie abbia un senso fisico e descriva l'universo alle scale piccolissime è anche di natura tecnologica: la tecnologia attuale non è in grado di preparare esperimenti che permettano il discernimento tra le varie teorie proposte. Altri punti di contrasto tra MQ e RSG, su cui la discussione è oggi un po' più pacata rispetto al passato, riguardano il rapporto della MQ con concetti di *località di natura relativistica* (paradosso Einstein-Podolsky-Rosen[5]) in relazione ai fenomeni di *entanglement* della MQ. Anche ed in particolare in seguito a celebri esperimenti di Aspect che hanno dato torto alle aspettative di Einstein e ragione all'interpretazione di Copenhagen, sembra ormai condivisa dalla maggior parte dei fisici l'idea che la MQ coinvolga processi fisici non locali, ma che ciò non implichi alcuna reale violazioni dei fondamenti della Relatività (l'entanglement quantistico non coinvolge trasmissione superluminare di informazioni e violazione della causalità [5]).

Nell'interpretazione standard della MQ detta *di Copenhagen*, rimangono punti fisicamente e matematicamente poco chiari, ma di estremo interesse concettuale. In particolare non è per nulla chiaro come la meccanica classica si possa ottenere come sottocaso o caso limite della MQ e come si possa fissare un limite (anche provvisorio o impreciso) tra i due mondi. Ulteriormente rimane aperto il problema della descrizione fisica e matematica del cosiddetto *processo di misura quantistica* di cui parleremo più avanti e che è strettamente connesso a quello del limite classico della MQ. Anche prendendo spunto da questo problema sono nate altre interpretazioni del formalismo della MQ, profondamente differenti dall'interpretazione di Copenhagen. Tra queste nuove interpretazioni, una volta considerate eretiche, di grande interesse è in particolare quella a *variabili nascoste* di Bohm [5, 6].

1.3 Convenzioni generali e prerequisiti fisico matematici.

Nel seguito, se non sarà precisato altrimenti il *campo degli scalari di uno spazio* di Hilbert sarà sempre quello complesso $\mathbb{C}$. Il *complesso coniugato* di un numero c sarà indicato con $\overline{c}$.

Il *prodotto scalare hermitiano* tra due vettori ψ, ϕ di uno spazio di Hilbert sarà indicato con $(\psi|\phi)$ eccetto nel caso in cui si userà la notazione *bra-ket* di Dirac di cui diremo nel paragrafo relativo. Si supporrà sempre che l'entrata *di sinistra* del prodotto scalare sia quella antilineare: $(c\psi, \phi) = \overline{c}(\psi, \phi)$.

La locuzione *sottospazio* sarà riservata ai sottospazi *rispetto alla semplice struttura di spazio vettoriale* anche nel caso in cui esista un'ulteriore struttura (spazio di Hilbert, Banach o altro) nello spazio ambiente.

L'operazione di coniugazione hermitiana sarà sempre indicata con † e *operatore hermitiano*, *operatore simmetrico* ed *operatore autoaggiunto* **non** saranno considerati sinonimi; si vedano le definizioni corrispondenti nel seguito.

Prerequisiti matematici necessari per comprendere il contenuto di queste dispense sono essenzialmente, oltre ai contenuti di un normale corso completo di algebra lineare uno di analisi per funzioni di una e di più variabili, qualche nozione di topologia elementare degli spazi metrici, i fondamenti della teoria della misura su sigma algebre (in particolare il teorema di Riesz) [1], qualche nozione di teoria elementare delle funzioni di una variabile complessa.
Dal punto di vista fisico è necessaria la conoscenza di alcuni argomenti dei corsi universitari elementari di argomento Fisico. Più precisamente le nozioni di Meccanica elementare con alcuni elementi di Meccanica Analitica (formulazione di Hamilton della dinamica) unitamente ad alcune nozioni di Elettromagnetismo (proprietà elementari delle Onde Elettromagnetiche e fenomeni ondulatori principali quali interferenza, diffrazione, diffusione).
Le nozioni meno elementari sia di Matematica che di Fisica ed altre nozioni utili solo in alcuni punti verranno comunque riassunte brevemente nel testo.

Gli esempi presentati nei vari capitoli devono considerarsi come parte integrante del testo e non possono essere omessi quasi mai. Alcuni risultati ottenuti negli esercizi verranno successivamente usati nel testo come proposizioni note. Gli esercizi in questione sono corredati di soluzione.

I paragrafi il cui titolo è preceduto da un asterisco possono essere omessi in una prima lettura della parte I, ma possono essere richiesti per la comprensione della parte II.

Parte I

Elementi di teoria degli operatori limitati su spazi di Hilbert

Capitolo 2

Alcune nozioni e teoremi generali nella teoria degli spazi normati e di Banach.

In questo capitolo presenteremo alcune nozioni ed alcuni risultati fondamentali della teoria generale degli spazi normati e degli spazi di Banach. Introdurremo anche alcune strutture algebriche modellizzate su algebre naturali di operatori sugli spazi di Banach. Le algebre di Banch di operatori svolgono un ruolo di grande importanza nelle formulazioni moderne della Meccanica Quantistica.
Questo capitolo serve essenzialmente ad introdurre il linguaggio e gli strumenti elementari della teoria degli spazi di operatori lineari.
Le nozioni più importanti di questa sezione sono sicuramente la nozione di *operatore limitato* e le varie nozioni di *topologia (indotta da norme o da seminorme) negli spazi di operatori.* L'importanza di questi strumenti matematici deriva dal fatto che il linguaggio degli operatori lineari su spazi lineari è il linguaggio con cui è formulato la Meccanica Quantistica. In questo contesto la classe degli operatori limitati riveste un ruolo tecnico centrale anche se, per motivi di carattere fisico, in Meccanica Quantistica ci si trova costretti ad introdurre e lavorare anche con operatori non limitati. Vedremo tutto ciò nella seconda parte.

2.1 Spazi normati, di Banach e algebre.

Diamo di seguito le principali definizioni sugli spazi normati e sulle algebre.

Definizione 2.1. *Sia X uno spazio vettoriale sul campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$. Un'applicazione $N : X \to \mathbb{R}$ si dice* **norma** *su X e (X, N) si dice* **spazio normato**, *quando N soddisfa le seguenti proprietà:*
N0. $N(u) \geq 0$ *per* $u \in X$,
N1. $N(\lambda u) = |\lambda| N(u)$ *per* $\lambda \in \mathbb{K}$ *e* $u \in X$,
N2. $N(u + v) \leq N(u) + N(v)$, *per* $u, v \in X$,

N3. *$N(u) = 0 \Rightarrow u = 0$, per $u \in X$.*
Se **N0**, **N1** *e* **N2** *sono soddisfatte ma non lo è necessariamente* **N3**, *N si dice* **seminorma**.

Nota. È chiaro che da **N1** discende che $N(0) = 0$.

Notazione 2.1. Nel seguito $||\ ||$ e $p(\ )$, aggiungendo eventuali indici quando necessari, indicheranno sempre rispettivamente norme e seminorme.

È chiaro che ogni spazio normato ammette una topologia metrica naturale indotta dalla metrica $d(u, v) := ||u - v||$ per $u, v \in X$. ($d : X \times X \to [0, +\infty)$ è una metrica in quanto per costruzione è simmetrica, positiva, nulla solo se $u = v$ e vale la **disuguaglianza triangolare** $||x - y|| \leq ||x - z|| + ||z - y||$ per $x, y, z \in X$.)
Nel seguito ci riferemo quasi sempre a tale topologia usando concetti topologici in uno spazio normato.

Definizione 2.2. *Se $(X, ||\ ||_X)$ e $(Y, ||\ ||_Y)$ sono spazi normati sullo stesso campo $\mathbb{C}$ o $\mathbb{R}$, un'applicazione lineare $L : X \to Y$ è detta* **isometria** *se soddisfa $||L(x)||_Y = ||x||_X$ per ogni $x \in X$.*
Se l'isometria $L : X \to Y$ è anche suriettiva, è detta **isomorfismo di spazi normati**.
Se esiste un isomorfismo di spazi normati (L) dallo spazio normato X allo spazio normato Y, tali spazi si dicono **isomorfi (secondo L)**.

Note.
(1) Frequentemente si trova in letteratura una definizione alternativa e *non equivalente* di isomorfismo di spazi normati. Tale definizione richiede che l'isomorfismo sia un'applicazione lineare bicontinua (omeomorfismo). È chiaro che un isomorfismo nel senso della definizione 2.2 lo è anche rispetto a questa seconda definizione, tuttavia non vale il viceversa.
(2) È chiaro che ogni isometria $L : X \to Y$ è iniettiva per **N3**, ma può *non* essere suriettiva (se vale $X = Y$ la non suriettività può esserci solo se la dimensione dello spazio X non è finita). Ogni isometria è ovviamente continua rispetto alle due topologie metriche dei due spazi e, se è suriettiva (cioè se è un isomorfismo), la sua inversa è ancora un'isometria e quindi un isomorfismo.

Definizione 2.3. *Uno spazio normato si dice* **spazio di Banach** *quando è completo nella sua topologia metrica (cioè ogni successione di Cauchy di elementi dello spazio converge a qualche elemento dello spazio.)*

È chiaro che la proprietà di completezza è invariante per isomorfismi di spazi normati ma non lo è sotto omeomorfismi. Un controesempio è dato dalla coppia di spazi $\mathbb{R}$ e $(0, 1)$ entrambi dotati della norma valore assoluto. I due spazi sono omeomorfi ma il primo è completo, mentre il secondo non lo è.
È immediato provare che ogni sottospazio chiuso di uno spazio di Banach è a sua volta uno spazio di Banach rispetto alla restrizione della norma. È un fatto noto che ogni spazio normato può essere *completato* producendo uno spazio di Banach in cui lo spazio di partenza è rappresentato

da un sottospazio denso. Vale a tal proposito il:

Teorema del completamento per spazi di Banach. *Sia X uno spazio vettoriale sul campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$ dotato di una norma N.*
(a) *Esiste uno spazio di Banach (Y, M) su $\mathbb{K}$, detto* **completamento** *di X tale che X si identifica isometricamente con un sottospazio denso di Y tramite un applicazione lineare iniettiva $J : X \to Y$.*
In altre parole, esiste un applicazione lineare iniettiva $J : X \to Y$ con

$$\overline{J(X)} = Y \quad e \quad M(J(x)) = N(x) \quad per\ ogni\ x \in X.$$

(b) *Se la terna (J_1, Y_1, M_1) con $J_1 : X \to Y_1$ lineare isometrica e (Y_1, M_1) spazio di Banach su $\mathbb{K}$ è tale che X si identifica isometricamente con un sottospazio denso di Y_1 tramite J_1 allora esiste ed è unico un isomorfismo di spazi normati $\phi : Y \to Y_1$ tale che $J_1 = \phi \circ J$.*

Schema della dimostrazione. Diamo solo l'idea generale della dimostrazione. (a) Conviene considerare lo spazio C delle successioni di Cauchy di elementi di X e definire la relazione di equivalenza in C:

$$x_n \sim x'_n \quad \Leftrightarrow \quad \text{esiste} \quad \lim_{n\to\infty} ||x_n - x'_n|| = 0\,.$$

È chiaro che $X \subset C/\sim$ identificando ogni x di X con la classe di equivalenza della successione costante $x_n = x$. L'applicazione che definisce tale identificazione la indicheremo con J. Si prova facilmente che $C/\sim$ è uno spazio vettoriale su $\mathbb{K}$ normato rispetto alla struttura indotta naturalmente da quella di X. Si prova infine che $C/\sim$ è completo, che J è lineare, isometrica (e quindi iniettiva) e che $J(X)$ è denso in $Y := C/\sim$. (b) $J_1 \circ J^{-1} : J(X) \to Y_1$ è una trasformazione lineare isometrica continua definita su un insieme denso $J(X) \subset Y$ a valori in uno spazio di Banach Y_1, e pertanto si estende unicamente ad una trasformazione ϕ lineare continua ed isometrica su Y (vedi esercizi (3) e (4) in esercizi 2.1). Essendo ϕ isometrica, è anche iniettiva. La stessa cosa si può dire per l'estensione ϕ' di $J \circ J_1^{-1} : J_1(X) \to Y$ e per costruzione $(J \circ J_1^{-1}) \circ (J_1 \circ J^{-1}) = id_{J(X)}$. Estendendo per continuità su $\overline{J(X)} = Y$ troviamo che $\phi' \circ \phi = id_Y$ e con un analogo ragionamento troviamo anche che $\phi \circ \phi' = id_{Y_1}$. Concludiamo che ϕ e ϕ' sono anche surgettive ed in particolare ϕ è un isomorfismo di spazi normati e che per costruzione vale $J_1 = \phi \circ J$. L'unicità di un isomorfismo $\phi : Y \to Y'$ che soddisfa $J_1 = \phi \circ J$ si ha facilmente notando che ogni altro siffatto isomorfismo di spazi normati $\psi : Y \to Y_1$, per linearità deve soddisfare $J - J = (\phi - \psi) \circ J$ e quindi $(\phi - \psi) \restriction_{J(X)} = 0$. L'unicità dell'estensione dell'applicazione $(\phi - \psi) \restriction_{J(X)}$, continua con dominio $J(X)$ denso, a $\overline{J(X)} = Y$ prova che $\phi = \psi$. $\square$

Come vedremo esiste uno stretto legame tra *algebre* e spazi normati, tale legame passa per la nozione di operatore lineare su uno spazio normato. Le definizioni principali riguardante la nozione di algebra sono riassunte di seguito.

Definizione 2.4. *Un'* **algebra** A *sul campo* $\mathbb{K} = \mathbb{C}$ *o* $\mathbb{R}$ *è uno spazio vettoriale su* $\mathbb{K}$ *dotato di un'ulteriore applicazione detta* **prodotto** *dell'algebra:* $\circ : A \times A \to A$ *tale che:*
A1. $a \circ (b + c) = a \circ b + a \circ b$ *per* $a, b, c \in A$,
A2. $(b + c) \circ a = b \circ a + c \circ a$ *per* $a, b, c \in A$,
A3. $\alpha(a \circ b) = (\alpha a) \circ b = a \circ (\alpha b)$ *per* $\alpha \in \mathbb{K}$ *e* $a, b \in A$.
L'algebra A *è detta:*
commutativa *o* **abeliana** *se*
A4. $a \circ b = b \circ a$ *per ogni coppia* $a, b \in A$;
algebra con unità *se contiene un elemento* $\mathbb{I}$, *detto* **unità** *dell'algebra, tale che:*
A5. $\mathbb{I} \circ u = u \circ \mathbb{I} = u$ *per ogni* $a \in A$;
algebra normata *ovvero* **algebra normata con unità** *se è uno spazio vettoriale normato con norma* $|| \, ||$ *che soddisfi la relazione*
A6. $||a \circ b|| \leq ||a||||b||$ *per* $a, b \in A$;
e, nel caso di algebra normata con unità, valga anche:
A7. $||\mathbb{I}|| = 1$;
algebra di Banach *ovvero* **algebra di Banach con unità** *se* A *è spazio di Banach e rispettivamente algebra normata, o algebra normata con unità rispetto alla stessa norma.*
Un **omorfismo di algebre (con unità, normate, di Banach)**, $\phi : A \to A'$ *è un omomorfismo rispetto alla struttura di spazio vettoriale (funzione lineare) che preserva il prodotto delle algebre e l'unità se presente.* ϕ *è detto* **isomorfismo di algebre (con unità, normate, di Banach)** *se è anche biettivo. Se esiste un isomorfismo* $\phi : A \to A'$, *le algebre* A *e* A' *(con unità, normate, di Banach) sono dette* **isomorfe**.

Note.
(1) Si dimostra immediatamente che l'unità, se esiste è unica.
(2) La nozione di norma non viene coinvolta nella definizione di omomorfismo e isomorfismo di algebre con unità, normate, di Banach.

Notazione 2.2. Nel seguito, se ciò non darà luogo a fraintendimenti, indicheremo il prodotto di due elementi di un'algebra semplicemente con ab invece che con $a \circ b$.

Esempi 2.1.
(1) I campi $\mathbb{C}$ e $\mathbb{R}$ sono banalmente algebre commutative di Banach. In entrambi i casi la norma è l'operazione di estrazione del valore assoluto o modulo.

(2) Se X è un insieme qualsiasi e $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$, indichiamo con $L(X)$ l'insieme di tutte le funzioni $f : X \to \mathbb{K}$ *limitate* (cioè $\sup_{x \in X} |f(x)| < \infty$). $L(X)$ ha una struttura naturale di spazio vettoriale su $\mathbb{K}$ rispetto alla solita operazione di composizione lineare di funzioni. Possiamo aggiungere un prodotto che rende $L(X)$ un'algebra: se $f, g \in L(X)$, fg è definita come, punto per punto, $(fg)(x) := f(x) \cdot g(x)$. Si osservi che l'algebra è commutativa e con unità (data dalla funzione che vale sempre 1). Una norma che rende $L(X)$ algebra di Banach commutativa è quella del-

l'estremo superiore: $||f|| := \sup_{x \in X} |f(x)|$ [1].

(3) Se sull'insieme X di sopra definiamo una σ-algebra, Σ, la sottoalgebra delle funzioni Σ-misurabili, $M(X, \Sigma) \subset L(X)$, è un insieme chiuso in $L(X)$ rispetto alla topologia della norma dell'estremo superiore. Quindi $M(X, \Sigma)$ è a sua volta un algebra di Banach commutativa [1].

(4) Se X è uno spazio topologico, lo spazio vettoriale delle funzioni continue a valori nel campo $\mathbb{C}$ si indica con $C(X)$. $C_b(X) \subset C(X)$ denota il sottospazio delle funzioni continue limitate, $C_c(X) \subset C_b(X)$ denota infine il sottospazio delle funzioni continue a supporto compatto. Nel caso X sia compatto i tre spazi coincidono evidentemente. I tre spazi sono sicuramente algebre commutative rispetto alle operazioni dette nell'esempio (2). $C(X)$ e $C_b(X)$ sono algebre con unità data dalla funzione costante di valore 1, mentre non lo è $C_c(X)$ quando X non è compatto. Rispetto alla norma dell'estremo superiore definita nell'esempio (2) $C_b(X)$ è algebra di Banach. È un risultato importante delle teorie delle algebre di Banach [2] che ogni algebra di Banach con unità e commutativa sul campo $\mathbb{C}$ è isomorfa ad un'algebra $C(K)$ con K compatto. Se X è uno spazio topologico di Hausdorff (cioè per ogni coppia di punti $x, y \in X$ con $x \neq y$ esistono due aperti A, B tali che $A \ni x$, $B \ni y$ e $A \cap B = \emptyset$) che sia anche localmente compatto (cioè per ogni punto $p \in X$ c'è un aperto A ed un compatto K tali che $p \in A \subset K$), completando lo spazio normato $C_c(X)$ si ottiene un'algebra di Banach (senza unità) commutativa, tale algebra di Banach si chiama l'algebra delle funzioni continue *che tendono a zero all'infinito in* X [1]. Si tratta delle funzioni continue $f : X \to \mathbb{C}$, tali che, per ogni $\epsilon > 0$ esiste un compatto $K_\epsilon \subset X$ (dipendente da f in generale) con $|f(x)| < \epsilon$ se $x \in X \setminus K_\epsilon$.

(5) Se X è spazio di Hausdorff compatto, consideriamo in $C(X)$ una sottoalgebra A che soddisfi i seguenti requisiti: contenga l'unità (data dalla funzione che vale ovunque 1) e sia chiusa rispetto alla coniugazione complessa – se $f \in A$ allora $f^* \in A$, dove $f^*(x) := \overline{f(x)}$ per ogni $x \in X$ in cui la barra indica la coniugazione complessa –. Diremo che A *separa i punti* di X, se per ogni coppia $x, y \in X$ con $x \neq y$ esiste $f \in A$ con $f(x) \neq f(y)$. Il teorema *di Stone-Weierstrass* [2], prova che:
ogni sottoalgebra $A \subset C(X)$, con X di Hausdorff compatto, che contenga l'unità, sia chiusa rispetto alla coniugazione complessa e separi i punti, è tale che la sua chiusura rispetto alla norma dell'estremo superiore coincide con $C(X)$ stessa.
Un esempio tipico è quello in cui X è un compatto di $\mathbb{R}^n$ e A è l'algebra dei polinomi complessi ad n variabili (le coordinate di $\mathbb{R}^n$) ristretti ad A. Dal teorema di Stone-Weierstrass si ricava che con i polinomi possiamo approssimare uniformemente ogni funzione complessa continua definita su X. Tale risultato è utile nella teoria degli spazi di Hilbert per costruire basi hilbertiane.

(6) Se (X, Σ, μ) è uno spazio con misura positiva e $1 \leq p < \infty$ valgono le *disuguaglianze di*

Hölder e *di Minkowki*, rispettivamente:

$$\int_X |f(x)g(x)|d\mu(x) \leq \left(\int_X |f(x)|^p d\mu(x)\right)^{1/p} \left(\int_X |g(x)|^q d\mu(x)\right)^{1/q}, \quad (2.1)$$

$$\left(\int_X |f(x)+g(x)|^p d\mu(x)\right)^{1/p} \leq \left(\int_X |f(x)|^p d\mu(x)\right)^{1/p} + \left(\int_X |g(x)|^p d\mu(x)\right)^{1/p}, \quad (2.2)$$

dove $f, g : X \to \mathbb{C}$ sono funzioni misurabili e $1/p + 1/q = 1$ [1]. Se indichiamo con $\mathcal{L}^p(X,\mu)$ l'insieme contenente tutte le funzioni $f : X \to \mathbb{C}$ che sono μ-misurabili e tali che $\int_X |f(x)|^p d\mu(x) < \infty$, si prova facilmente che, in base alle disuguaglianze di sopra, $\mathcal{L}^p(X,\mu)$ è uno spazio vettoriale rispetto alle solite composizioni lineari di funzioni e che P_p definita sotto è effettivamente una *seminorma*:

$$P_p(f) := \left(\int_X |f(x)|^p d\mu(x)\right)^{1/p} \quad (2.3)$$

Per ottenere una vera norma, cioè per avere che sia soddisfatta **N3**, bisogna fare in modo da identificare con la funzione nulla ogni funzione che differisce da essa per un insieme di misura nulla. Tali funzioni pur non essendo nulle annullano l'integrale di sopra violando **N3** a P_p. A tal fine si può definire la relazione di equivalenza su $\mathcal{L}^p(X,\mu)$ data da $f \sim g$ se e solo se $f - g$ è nulla quasi ovunque rispetto a μ. Si indica con $L^p(X,\mu)$ lo spazio quoziente $\mathcal{L}^p(X,\mu)/\sim$. Questo spazio eredita naturalmente una struttura di spazio vettoriale su $\mathbb{C}$ da quella di $\mathcal{L}^p(X,\mu)$ semplicemente definendo $[f] + [g] := [f + g]$ e $\alpha[f] := [\alpha f]$ (si verifica che le definizioni non dipendono nei secondi membri dai rappresentanti scelti nelle classi di equivalenza usate a primo membro).
Si riesce a provare ([1], teorema 3.11) che $L^p(X,\mu)$ è spazio di Banach rispetto alla norma

$$||[f]||_p := \left(\int_X |f(x)|^p d\mu(x)\right)^{1/p} \quad (2.4)$$

dove f è un rappresentante arbitrario di $[f] \in L^p(X,\mu)$. Lo spazio di Banach ottenuto non è algebra rispetto al prodotto puntuale dato che il prodotto puntuale di funzioni di $\mathcal{L}^p(X,\mu)$ non è, in generale elemento dello stesso spazio.
N.B. *Nella letteratura corrente prevale l'uso del simbolo f per denotare la classe di equivalenza $[f] \in L^2(X,\mu)$. Noi seguiremo tale uso nelle situazioni in cui ciò non produrrà confusione.*

(7) In riferimento all'esempio (6), consideriamo il caso particolare in cui X è un insieme non necessariamente numerabile, Σ è l'insieme delle parti di X e μ è la misura *che conta i punti* : se $S \subset X$, $\mu(S) =$ *numero di elementi di* S e $\mu(S) = \infty$ se S contiene infiniti punti. In questo caso lo spazio $L^p(X,\mu)$ si indica semplicemente con $l^p(X)$. I suoi elementi sono le "successioni" di complessi etichettati su X $\{z_x\}_{x\in X}$ tali che

$$\sum_{x\in X} |z_x|^p < \infty ,$$

dove la somma è definita come

$$\sup\left\{\sum_{x\in X_0}|z_x|^p \;\middle|\; X_0\subset X,\quad X_0 \text{ finito}\right\}.$$

Si osservi che, nel caso X sia numerabile, $X=\mathbb{N}$ o $\mathbb{Z}$ in particolare, la definizione data sopra di somma di un insieme di numeri *positivi* etichettati su X si riduce a quella solita di somma di una serie (vedi definizione 3.4 e proposizione 3.2 nel capitolo 3).

(8) Se (X,Σ,μ) è uno spazio con misura positiva e $f:X\to\mathbb{C}$ una funzione misurabile definiamo $S_f:=\{r\in\mathbb{R}\mid \mu(|f|^{-1}(r,\infty))=0\}$. Le funzioni misurabili $f:X\to\mathbb{C}$ per cui $S_f\neq\emptyset$ costituiscono un'algebra commutativa con unità con le solite operazioni definite negli esempi di sopra per gli spazi di funzioni. È possibile definire una norma su tale algebra rispetto alla quale si abbia un'algebra di Banach (vedi [1], teorema 3.11 per la dimostrazione della completezza): $||f||:=\text{ess sup}f$ dove l'*estremo superiore essenziale di* f, ess sup$f$, è definito come $\inf S_f$ ossia

$$\text{ess sup}f:=\inf\left\{r\in\mathbb{R}\;\middle|\;\mu(|f|^{-1}(r,\infty))=0\right\}. \tag{2.5}$$

L'algebra di Banach che si ottiene si indica con $L^\infty(X,\mu)$ e si chiama l'algebra delle funzioni *essenzialmente limitate.*

(9) In riferimento all'esempio (8), nel caso particolare in cui Σ è l'insieme delle parti di X e μ è la misura che conta i punti, lo spazio $L^\infty(X,\mu)$ si indica semplicemente con $l^\infty(X)$. I suoi elementi sono le "successioni" di complessi etichettati su X $\{z_x\}_{x\in X}$ tali che $\sup_{x\in X}|z_x|<\infty$. (Per cui, in riferimento alla notazione usata nell'esempio (2), $l^\infty(X)=L(X)$.)

2.2 Operatori, spazi di operatori, norme di operatori.

Definizione 2.5. *Siano X e Y spazi vettoriali sullo stesso campo $\mathbb{K}:=\mathbb{R}$ o $\mathbb{C}$.*
(a) *$T:X\to Y$ è detto* **operatore lineare** *(o semplicemente* **operatore***)* **da** *X* **in** *Y se è lineare. $\mathfrak{L}(X,Y)$ denota l'insieme degli operatori lineari da X in Y. Quando X e Y sono normati, $\mathfrak{B}(X,Y)\subset\mathfrak{L}(X,Y)$ denota il sottoinsieme degli operatori lineari continui. In particolare $\mathfrak{L}(X):=\mathfrak{L}(X,X)$ e $\mathfrak{B}(X):=\mathfrak{B}(X,X)$.*
(b) *$T:X\to\mathbb{K}$ è detto* **funzionale lineare** *(o semplicemente* **funzionale***)* **su** *X se è lineare.*
(c) *Lo spazio $X^*:=\mathfrak{L}(X,\mathbb{K})$ è detto* **duale algebrico** *di X mentre, se $\mathbb{K}$ è inteso come uno spazio normato rispetto alla norma indotta dal valore assoluto, $X':=\mathfrak{B}(X,\mathbb{K})$ è detto* **duale topologico** *(o semplicemente* **duale***)* **di** *X.*

Al solito se $T,T'\in\mathfrak{L}(X,Y)$ e $\alpha,\beta\in\mathbb{K}$, la combinazione lineare $\alpha T+\beta T'$ è definita come l'applicazione $(\alpha T+\beta T'):u\mapsto\alpha(Tu)+\beta(T'u)$ per ogni $u\in X$ ed è quindi ancora un elemento di $\mathfrak{L}(X,Y)$. Dato che come è facile provare vale anche che ogni combinazione lineare di operatori continui è un operatore continuo concludiamo che $\mathfrak{L}(X,Y)$, $\mathfrak{L}(X)$ e X^* e $\mathfrak{B}(X,Y)$, $\mathfrak{B}(X)$ e X',

sono tutti *spazi vettoriali* su $\mathbb{K}$.

Introduciamo il concetto di operatore e funzionale *limitato*.

Teorema 2.1. *Siano* $(X, || \, ||_X), (Y, || \, ||_Y)$ *spazi normati sullo stesso campo* $\mathbb{K} = \mathbb{C}$ *o* $\mathbb{R}$. *Si consideri* $T \in \mathfrak{L}(X, Y)$.
(a) *Le seguenti due condizioni sono equivalenti*
(i) esiste $K \in \mathbb{R}$ *tale che* $||Tu||_Y \leq K||u||_X$ *per ogni* $u \in X$*;*
(ii) esiste $\sup_{u \in X \setminus \{0\}} \frac{||Tu||_Y}{||u||_X} < +\infty$.
(b) *Se vale la condizione (i) allora*

$$\sup \left\{ \frac{||Tu||_Y}{||u||_X} \;\middle|\; u \in X \setminus \{0\} \right\} = \inf \{K \in \mathbb{R} \mid ||Tu||_Y \leq K||u||_X \quad \textit{per ogni } u \in X\} \; .$$

Prova. (a) Se vale (i), per costruzione $\sup_{u \in X \setminus \{0\}} \frac{||Tu||_Y}{||u||_X} \leq K < \infty$. Se vale (ii) posto $A := \sup_{u \in X \setminus \{0\}} \frac{||Tu||_Y}{||u||_X}$, $K := A$ soddisfa (i).
(b) Detto I l'estremo inferiore dell'insieme dei K che soddisfano (i), valendo $\sup_{u \in X \setminus \{0\}} \frac{||Tu||_Y}{||u||_X} \leq K$ deve essere $\sup_{u \in X \setminus \{0\}} \frac{||Tu||_Y}{||u||_X} \leq I$. Supponiamo per assurdo che valga la disuguaglianza stretta. Sia allora J un punto con $\sup_{u \in X \setminus \{0\}} \frac{||Tu||_Y}{||u||_X} < J$ e $J < I$. Dalla prima condizione segue che $||Tu||_Y < J||u||_X$ per ogni $u \neq 0$ e quindi $||Tu||_Y \leq J||u||_X$ per ogni $u \in X$, ma allora J soddisfa la condizione (i) e ciò contraddice $J < I$. $\square$

Nel seguito ometteremo gli indici nelle norme per denotare gli spazi su cui sono definite se ciò sarà ovvio dal contesto.

Definizione 2.6. *Siano* X, Y *spazi normati sul medesimo campo* $\mathbb{C}$ *o* $\mathbb{R}$. $T \in \mathfrak{L}(X, Y)$ *è detto* **limitato** *se vale una delle due condizioni equivalenti dette in (a) nell'enunciato del teorema 2.1. In tal caso il numero*

$$||T|| := \sup_{||u|| \neq 0} \frac{||Tu||}{||u||} \; . \tag{2.6}$$

è detto **norma (operatoriale)** *di* T.

Nota.
Dalla definizione di $||T||$ si ha subito che se $T : X \to Y$ è limitato vale l'utile proprietà:

$$||Tu|| \leq ||T|| \, ||u|| \; , \quad \text{per ogni } u \in X \; . \tag{2.7}$$

La norma operatoriale può essere calcolata anche in altri modi talvolta utili nelle dimostrazioni. Vale la seguente proposizione in proposito.

Proposizione 2.1. *Siano X, Y spazi normati sul medesimo campo $\mathbb{C}$ o $\mathbb{R}$. $T \in \mathfrak{L}(X, Y)$ è limitato se e solo se esiste ed è finito uno dei secondi membri delle tre identità di sotto ed in tal caso tali identità sono verificate.*

$$||T|| = \sup_{||u||=1} ||Tu|| \,, \tag{2.8}$$

ovvero

$$||T|| = \sup_{||u||\leq 1} ||Tu|| \,, \tag{2.9}$$

ovvero

$$||T|| = \inf \{K \in \mathbb{R} \mid ||Tu|| \leq K||u|| \quad \textit{per ogni } u \in X\} \,. \tag{2.10}$$

Prova. Il fatto che T sia limitato se e solo se il secondo membro di (2.8) esista e sia finito e la stessa identità (2.8) seguono facilmente dalla linearità di T e dalla proprietà **N1** delle norme.
Il fatto che T sia limitato se e solo se il secondo membro di (2.9) esista e sia finito e la stessa identità (2.9) seguono facilmente dal seguente ragionamento. Dato che l'nsieme degli u con $||u|| \leq 1$ include l'insieme degli u con $||u|| = 1$, vale $\sup_{||u||\leq 1} ||Tu|| \geq \sup_{||u||=1} ||Tu||$. D'altra parte, se $||u|| \leq 1$, vale $||Tu|| \leq ||Tv||$ per qualche v con $||v|| = 1$ (qualunque v suddetto se $u = 0$ e $v = u/||u||$ altrimenti). Quindi vale anche $\sup_{||u||\leq 1} ||Tu|| \leq \sup_{||u||=1} ||Tu||$, da cui si ottiene $\sup_{||u||\leq 1} ||Tu|| = \sup_{||u||=1} ||Tu||$ che dimostra quanto volevamo.
Il fatto che T sia limitato se e solo se il secondo membro di (2.10) esista e sia finito e la stessa identità (2.10) seguono facilmente da (b) del teorema 2.1 □

Il legame tra *continuità* e *limitatezza* è dato dal seguente teorema che prova, tra l'altro, che gli operatori limitati sono tutti e soli quelli continui.

Teorema 2.2. *Sia $T \in \mathfrak{L}(X, Y)$ con X, Y spazi normati sul medesimo campo. I seguenti fatti sono equivalenti:*
(i) T è continuo in 0;
(ii) T è continuo;
(iii) T è limitato.

Prova. (i) $\Leftrightarrow$ (ii). La continuità implica banalmente la continuità in 0. Mostriamo che la continuità in 0 implica la continuità. Valendo $(Tu) - (Tv) = T(u - v)$ si ha che $(\lim_{u\to v} Tu) - Tv = \lim_{u\to v}(Tu - Tv) = \lim_{(u-v)\to 0} T(u - v) = 0$ per la continuità in 0.
(i) $\Rightarrow$ (iii). Dalla continuità in 0, esiste $\delta > 0$ tale che, se $||u|| < \delta$ allora $||Tu|| < 1$. Scelto $\delta' > 0$ con $\delta' < \delta$, se $v \in X \setminus \{0\}$, $u = \delta' v/||v||$ ha norma inferiore a δ per cui $||Tu|| < 1$, che in termini di v si scrive $||Tv|| < (1/\delta')||v||$. Vale allora la condizione (a) del teorema 2.1 con $K = 1/\delta'$ e pertanto, per la definizione 4.2 T è limitato.

(iii) $\Rightarrow$ (i). Si tratta di un fatto ovvio. Se T è limitato allora $||Tu|| \leq ||T||||u||$ da cui la continuità in 0. $\square$.

Il nome "norma" per $||T||$ non è casuale; in effetti la norma di operatori rende a tutti gli effetti $\mathfrak{B}(X,Y)$, *e quindi in particolare* $\mathfrak{B}(X)$ *e* X', uno spazio normato come proveremo tra poco. Più precisamente, vedremo che $\mathfrak{B}(X,Y)$ è uno spazio di Banach se Y è di Banch e quindi, in particolare X' è sempre spazio di Banach.
Il teorema che segue riguarda anche un altro importante fatto in relazione alla struttura di algebra. Cominciamo con l'osservare che la composizione di operatori di $\mathfrak{L}(X)$, rispettivamente $\mathfrak{B}(X)$, produce operatori nello stesso spazio (in particolare perché la composizione di funzioni continue produce funzioni continue). Ulteriormente, è immediato provare che lo spazio vettoriale $\mathfrak{L}(X)$, rispettivamente $\mathfrak{B}(X)$, soddisfa gli assiomi **A1**, **A2** e **A3** della definizione di algebra quando il prodotto dell'algebra è definito come la *composizione* di operatori. In questo modo risulta chiaro che $\mathfrak{L}(X)$ e $\mathfrak{B}(X)$ possiedono una struttura naturale di *algebra con unità*, quest'ultima data dalla funzione identità $I : X \to X$, inoltre $\mathfrak{B}(X)$ risulta essere una sottoalgebra di $\mathfrak{L}(X)$.
Nell'ultima parte del teorema seguente si rafforza ulteriormente il risultato, provando che $\mathfrak{B}(X)$ è sempre un'*algebra normata con unità* rispetto alla norma operatoriale ed è ulteriormente *algebra di Banach* se X è uno spazio di Banach.

Teorema 2.3. *Siano X, Y spazi normati sullo stesso campo.*
(a) *L'applicazione $|| \; || : T \mapsto ||T||$, dove $||T||$ è definita da (2.6) per $T \in \mathfrak{B}(X,Y)$, è una norma su $\mathfrak{B}(X,Y)$ e rende $\mathfrak{B}(X,Y)$ spazio normato.*
(b) *Sull'algebra con unità $\mathfrak{B}(X)$ valgono le ulteriori relazioni che la rendono algebra normata con unità:*
(i) $||TS|| \leq ||T||||S||$ e $T, S \in \mathfrak{B}(X)$,
(ii) $||I|| = 1$.
(c) *Se Y è completo $\mathfrak{B}(X,Y)$ è uno spazio di Banach.*
In particolare:
(i) se X è uno spazio di Banach, $\mathfrak{B}(X)$ è un'algebra di Banach;
(ii) X' è sempre uno spazio di Banach rispetto alla norma dei funzionali, anche se X non è completo.

Prova. (a) è diretta conseguenze della definizione di norma di un operatore: le proprietà definitorie della norma **N1**, **N2**, **N3** per la norma operatoriale possono essere immediatamente verificate usando le stesse proprietà **N1**, **N2**, **N3** per la norma dello spazio Y, la formula (2.8) per la norma operatoriale e la definizione di estremo superiore.
(b) Il punto (i) è immediata conseguenza della (2.7) e della (2.8). (ii) è di immediata verifica usando l'espressione per la norma operatoriale (2.8).
Passiamo a provare la parte (c). Proviamo che se Y è completo allora $\mathfrak{B}(X,Y)$ è uno spazio di Banach. Sia $\{T_n\} \subset \mathfrak{B}(X,Y)$ una successione di Cauchy rispetto alla norma operatoriale. Da

(2.7) segue che

$$||T_n u - T_m u|| \leq ||T_n - T_m||||u|| \,,$$

il fatto che $\{T_n\}$ sia di Cauchy implica che sia di Cauchy la successione dei vettori $T_n u$. Essendo Y completo, per ogni fissato $u \in X$ esisterà un vettore di Y:

$$Tu := \lim_{n\to\infty} T_n u \,.$$

$X \ni u \mapsto Tu$ è una funzione lineare essendo tali tutte le funzioni T_n. Mostriamo, per concludere, che $T \in \mathfrak{B}(X,Y)$ e che $||T - T_n|| \to 0$ per $n \to \infty$.
Essendo $\{T_n\}$ una successione di Cauchy, se $\epsilon > 0$, varrà $||T_n - T_m|| \leq \epsilon$ per n, m sufficientemente grandi e quindi anche $||T_n x - T_m x|| \leq ||T_n - T_m||||u|| \leq \epsilon||u||$. Allora:

$$||Tu - T_m u|| = ||\lim_{n\to+\infty} T_n u - T_m u|| = \lim_{n\to+\infty} ||T_n u - T_m u|| \ \leq \epsilon||u||$$

se m è grande a sufficienza. Dalla stima ottenuta, essendo $||Tu|| \leq ||Tu - T_m u|| + ||T_m u||$ ed usando (2.7), segue ancora che

$$||Tu|| \leq (\epsilon + ||T_m||)||u|| \,.$$

Ciò dimostra che T è limitato e quindi $T \in \mathfrak{B}(X,Y)$ per il teorema 2.2. Valendo, come provato sopra, $||Tu - T_m u|| \leq \epsilon||u||$ si ha anche che $||T - T_m|| \leq \epsilon$ dove ϵ può essere scelto arbitrariamente piccolo pur di scegliere m grande a sufficienza. In altre parole $||T - T_n|| \to 0$ se $n \to \infty$.
La prova dei sottocasi (i) e (ii) è immediata. (i) segue dal fatto che $\mathfrak{B}(X) = \mathfrak{B}(X,X)$ e (ii) segue dal fatto che $X' := \mathfrak{B}(X,\mathbb{K})$ ed il campo di X, $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$ è completo come spazio normato. □

Esempi 2.2
(1) Consideriamo X spazio topologico di Hausdorff localmente compatto. Sappiamo che l'algebra normata $C_c(X)$ ha come completamento nella norma dell'estremo superiore l'algebra di Banach $C_0(X)$ delle funzioni che si annullano all'infinito ((4) in esempi 2.1). Se μ è una *misura di Borel complessa* [1], e $|\mu|$ è la misura positiva di Borel detta *variazione totale di* μ (se E è insieme di Borel in X, $|\mu|(E) := \sup\{\sum_i |\mu(E_i)| \ | \ \{E_i\}$ partizione di E di insiemi di Borel$\}$).
La variazione totale di una misura complessa soddisfa sempre $|\mu|(X) < \infty$ per definizione di misura complessa. (N.B. Quando μ è una misura di Borel *positiva*, l'applicazione diretta della definizione di sopra produce $\mu = |\mu|$, ma in questo caso la finitezza di $|\mu|(X)$ non è assicurata e vale se e solo se μ è finita; le misure di Borel *con segno* si intendono qui un sottocaso di quelle complesse).
Nelle ipotesi fatte definiamo $||\mu|| := |\mu|(X)$. È chiaro che, se $f \in C_0(X)$,

$$\left|\int_X f d\mu\right| \leq ||\mu||||f||_\infty \,,$$

dove $||f||_\infty = \sup_{x\in X} |f(x)|$. Da tale fatto segue subito che ogni misura complessa definisce un elemento del duale (topologico) di $C_0(X)$. Ricordiamo che una misura di Borel complessa è

detta *regolare* se è tale la sua variazione totale[1]. Il *teorema di Riesz* nel caso di misure complesse [1] prova che questo è in realtà il caso generale precisando anche qualcosa in più:
se X è uno spazio topologico di Hausdorff localmente compatto, se $\Lambda : C_0(X) \to \mathbb{C}$ è un funzionale lineare continuo, allora esiste un'unica misura di Borel regolare μ_Λ tale che, per ogni $f \in C_0(X)$:

$$\Lambda(f) = \int_X f d\mu_\Lambda \, .$$

Vale inoltre $||\Lambda|| = ||\mu_\Lambda||$.
Si osservi che dato che $C_c(X)$ è denso in $C_0(X)$, un funzionale continuo sul primo spazio ne individua univocamente uno sul secondo, per cui il teorema caratterizza anche i funzionali, continui rispetto alla norma dell'estremo superiore, su $C_c(X)$.
(2) In riferimento all'esempio (6) in esempi 2.1, se $1 \leq p < +\infty$ e $1/p + 1/q = 1$, il duale di $L^p(X, \mu)$ risulta essere $L^q(X, \mu)$ nel senso che l'applicazione $L^q(X, \mu) \ni [g] \mapsto \Lambda_g$ dove $\Lambda_g(f) := \int_X fg \, d\mu$, è un isomorfismo di spazi normati da $L^q(X, \mu)$ a $(L^p(X, \mu))'$. Analogamente il duale di $L^1(X, \mu)$ si identifica con $L^\infty(X, \mu)$.

Esercizi 2.1.
(1) Provare che ogni seminorma p soddisfa $p(x) = p(-x)$.
(2) Si dimostri che in uno spazio normato (X, N) ogni successione convergente è successione di Cauchy.

Soluzione. Sia $\{x_n\}_{n\in\mathbb{N}} \subset X$ convergente a x nella topologia indotta dalla norma N. Valendo $N(x_n - x_m) = N(x_n - x + x - x_m) \leq N(x_n - x) + N(x - x_m)$ per la disuguaglianza trinagolare, si ha che per ogni $\epsilon > 0$ esiste $M_\epsilon > 0$ per cui se $n, m > M_\epsilon$, $N(x_n - x) < \epsilon/2$ e $N(x_m - x) < \epsilon/2$. Quindi per ogni $\epsilon > 0$ esiste $M_\epsilon > 0$ per cui se $n, m > N_\epsilon$, $N(x_n - x_m) < \epsilon$ e la successione è di Cauchy.
(3) Si consideri una coppia di spazi normati X, Y, con Y spazio di Banach, e un sottospazio $S \subset X$ la cui chiusura rispetto alla norma di X coincida con tutto X (in altre parole S è denso in X). Sia poi $T : S \to Y$ un operatore lineare limitato su S. Si dimostri che esiste un'unica estensione limitata di T, $T' : X \to Y$.

Soluzione. Se $x \in X$ ci sarà una successione $\{x_n\}$ di elementi di S che converge a x. Vale $||Tx_n - Tx_m|| \leq K||x_n - x_m||$ con $K \leq +\infty$ per ipotesi. Dato che $x_n \to x$, la successione degli x_n è di Cauchy e quindi lo è anche quella dei vettori Tx_n. Dato che Y è completo, esiste dunque $T'x := \lim_{n\to\infty} Tx_n \in Y$. Tale limite dipende solo da x e non dalla successione in S usata per approssimarlo: se $S \ni z_n \to x$ allora per la continuità delle norme

$$|| \lim_{n\to+\infty} Tx_n - \lim_{n\to+\infty} Tz_n|| = \lim_{n\to+\infty} ||Tx_n - Tz_n|| \leq \lim_{n\to+\infty} K||x_n - z_n|| = K||x - x|| = 0 \, .$$

Ovviamente $T' \restriction_S = T$, cioè T' è un'estensione di T e questo si prova scegliendo per ogni $x \in S$ la successione costante di termini $x_n := x$ che tende a x banalmente. Le proprietà di linearità di

[1] Se μ è una misura di Borel positiva sullo spazio di Hausdorff localmente compatto X, μ è detta regolare se per ogni boreliano E, $\mu(E)$ coincide con l'estremo superiore dell'insieme dei numeri $\mu(K)$ dove $K \subset E$ è compatto, e con l'estremo inferiore dell'insieme dei numeri $\mu(V)$ dove $V \supset E$ è aperto.

T' sono di verifica immediata dalla stessa definizione. Infine prendendo il limite a $n \to +\infty$ ad ambo membri di $||Tx_n|| \leq K||x_n||$ si ricava che $||T'x|| \leq K||x||$ per cui T' è limitato. Veniamo alla prova dell'unicità. Se U è un'altra estensione limitata di T su X allora, per ogni $x \in X$, per continuità $T'x - Ux = \lim_{n\to+\infty}(T'x_n - Ux_n)$ dove gli x_n appartengono a S (denso in X). Essendo $T' \restriction_S = T = U \restriction_S$ il limite è banale e fornisce $T'x = Ux$ per ogni $x \in X$ ossia $T' = U$.
(4) In riferimento all'esercizio precedente, si provi che l'unica estensione limitata T' di T soddisfa anche $||T'|| = ||T||$.

Soluzione. Se $x \in X$ e $\{x_n\} \subset S$ converge a x:

$$||T'x|| = \lim_{n\to+\infty} ||Tx_n|| \leq \lim_{n\to+\infty} ||T||||x_n|| = ||T||||x|| \,,$$

per cui $||T'|| \leq ||T||$. Ma essendo anche $S \subset X$ e $T' \restriction_S = T$,

$$||T'|| = \sup\left\{ \frac{||T'x||}{||x||} \,\middle|\, 0 \neq x \in X \right\} \geq \sup\left\{ \frac{||T'x||}{||x||} \,\middle|\, 0 \neq x \in S \right\} = \sup\left\{ \frac{||Tx||}{||x||} \,\middle|\, 0 \neq x \in S \right\} .$$

L'ultimo termine nella catena di identità scritta sopra è $||T||$. Quindi $||T'|| \geq ||T||$. Allora $||T'|| = ||T||$.
(5)* Se definiamo un isomorfismo di spazi normati come un'applicazione lineare continua con inversa continua, la proprietà di completezza di uno spazio normato è invariante sotto isomorfismi?

Suggerimento. Se non si arriva ad una conclusione si legga il paragrafo 2.5 generalizzando la proposizione 2.6 al caso di due spazi normati connessi da un'applicazione lineare continua con inversa continua.

2.3 I tre teoremi fondamentali negli spazi di Banach e le topologie deboli.

Riportiamo gli enunciati de i tre teoremi fondamentali della teoria degli spazi normati e di Banach nella versione più elementare possibile. Le dimostrazioni sono date in Appendice B. Dimostrazioni anche con formulazioni più complesse di questi teoremi, si possono trovare per esempio in [1] e [2].
Le applicazioni del secondo teorema, quello di Banach-Steinhaus, forniscono l'occasione per introdurre diverse topologie negli spazi di operatori. Tali topologie rivestono un ruolo importantissimo in Meccanica Quantistica quando lo spazio di Banch di partenza è lo spazio di Hilbert della teoria, l'algebra di operatori limitati di interesse è costituita da (alcune) osservabili della teoria, mentre le proprietà elementari del sistema quantistico associuate ai processi di misura sono una sottoclasse della classe degli operatori di proiezione ortogonale. Per passare con continuità dall'algebra delle osservabili a quella dei proiettori sono necessarie topologie più deboli rispetto a quella standard. Questo genere di problematiche che discuteremo più oltre hanno portato alla nozione di *algebra (di operatori) di Von Neumann.*

Il primo teorema è il famosissimo teorema di Hahn-Banach che si occupa del problema dell'estensione di un funzionale lineare continuo da un sottospazio dello spazio ambinete ad un funzionale continuo definito tutto lo spazio ma che conserva la norma iniziale. Esistono versioni del teorema valide in ambienti meno rigidi degli spazi normati (spazi vettoriali topologici con seminorme). Tuttavia noi ci accontenteremo della versione del teorema per spazi normati.
Per enunciare il teorema, notiamo che se X è uno spazio normato e $M \subset X$ un suo sottospazio (rispetto alla sola struttura di spazio vettoriale di X), la restrizione della norma di X a M definisce su M una struttura di spazio normato. In questo senso si può parlare di operatori o funzionali definiti su M e continui (ossia limitati rispetto alla struttura di spazio normato indotta da quella di X).

Teorema di Hahn-Banach (per spazi normati). *Sia M un sottospazio (non necessariamente chiuso) di uno spazio normato X con campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$ e sia $g : M \to \mathbb{K}$ un funzionale lineare continuo. Esiste un funzionale lineare $f : X \to \mathbb{K}$ continuo tale che $f \restriction_M = g$ e $||f||_X = ||g||_M$.*

Una delle più utili conseguenze del teorema di Hahn-Banach è il seguente corollario.

Corollario 1. *Sia X spazio normato e $x_0 \in X$ con $x_0 \neq 0$. Esiste $f \in X'$ con $||f|| = 1$ tale che $f(x_0) = ||x_0||$.*

Prova. Si scelga $M := \{\lambda x_0 \mid \lambda \in \mathbb{K}\}$ e $g : \lambda x_0 \to \lambda ||x_0||$. Sia $f \in X'$ il funzionale limitato che estende g secondo il teorema di Hahn-Banach. Per costruzione $f(x_0) = g(x_0) = ||x_0||$ e $||f||_X = ||g||_M = 1$.□

Un altro risultato, che ha importanti conseguenze nella teoria delle algebre di Banach, è il seguente.

Corollario 2. *Sia X spazio normato non banale. Gli elementi di X' separano i punti di X. Cioè, se $x_1 \neq x_2$ sono punti di X (ed esistono tali punti se X non è lo spazio vettoriale banale $\{0\}$) allora c'è almeno un funzionale $f \in X'$ per cui $f(x_1) \neq f(x_2)$.*

Prova. Basta scegliere nel corollario 1 $x_0 := x_1 - x_2$ ottenendo $f(x_1) - f(x_2) = f(x_1 - x_2) = ||x_1 - x_2|| \neq 0$. □

Se $x \in X$ e $f \in X'$ con $||f|| = 1$ allora $|f(x)| \leq 1||x||$ per cui

$$\sup\{|f(x)| \mid f \in X' \ , \ ||f|| = 1\} \leq ||x|| \, .$$

Il corollario 1 consente di rafforzare il risultato provando immediatamente che

$$\sup\{|f(x)| \mid f \in X' \ , \ ||f|| = 1\} = \max\{|f(x)| \mid f \in X' \ , \ ||f|| = 1\} = ||x|| \, .$$

Questo fatto apparentemente non molto profondo ha invece una certa rilevanza in una questione importante che nasce nella teoria degli spazi normati infinitodimensionali quando la si confronta

con quella nel caso finitodimensionale.
È noto dalla teoria elementare degli spazi vettoriali che lo spazio $(X^*)^*$, duale algebrico del duale algebrico di uno spazio vettoriale *di dimensione finita*, ha la notevole proprietà di essere *naturalmente isomorfo* a X stesso. L'isomorfismo è dato dalla applicazione lineare che associa ad $x \in X$ il funzionale lineare su X^*, $\mathfrak{I}(x)$, definito da $(\mathfrak{I}(x))(f) := f(x)$ per ogni $f \in X^*$.
Nel caso infinitodimensionale $\mathfrak{I}$ identifica X con un sottospazio di $(X^*)^*$, ma non, in generale con tutto $(X^*)^*$. C'è qualche proposizione generale a riguardo che vale lavorando con spazi normati infinitodimensionali considerando però i *duali in senso topologico*?
Si noti che $(X')'$ è il duale topologico di uno spazio normato (X' la cui norma è quella operatoriale). Di conseguenza $(X')'$ è uno spazio normato a sua volta, la norma essendo ancora una volta quella operatoriale.
Consideriamo ancora la trasformazione lineare naturale $\mathfrak{I} : X \to (X')^*$ che associa a $x \in X$ l'elemento $\mathfrak{I}(x) \in (X')^*$, cioè la funzione lineare $\mathfrak{I}(x) : X' \to \mathbb{K}$, definita come

$$(\mathfrak{I}(x))(f) := f(x) \quad \text{per ogni } f \in X' \text{ e } x \in X\,. \tag{2.11}$$

(È chiaro per costruzione che $\mathfrak{I}(x)$ è un funzionale lineare su X' per cui effettivamente $\mathfrak{I}(x) \in (X')^*$). Il fatto che

$$\sup\{|f(x)| \mid f \in X' \ , \ ||f|| = 1\} = ||x||$$

ha due implicazioni immediate: (1) $\mathfrak{I}(x)$ è un funzionale *limitato*, per cui appartiene a $(X')'$ e (2) $||\mathfrak{I}(x)|| = ||x||$. Per cui la trasformazione lineare $\mathfrak{I} : X \to (X')'$ è un *isometria* e quindi in particolare è *iniettiva*. In definitiva si ha l'inclusione isometrica $X \subset (X')'$ data dall'isometria $\mathfrak{I} : X \to (X')'$. Abbiamo provato il seguente corollario.

Corollario 3. La trasformazione lineare $\mathfrak{I} : X \to (X')'$ definita da $(\mathfrak{I}(x))(f) := f(x)$ per ogni $x \in X$ e $f \in X'$ è iun isometria. In tal modo X si identifica isometricamente con un sottospazio di $(X')'$.

Si possono trovare esempi, nel caso infinitodimensionali in cui X non ricopre tutto $(X')'$. Ciò porta a dare la seguente definizione.

Definizione 2.7. *Uno spazio normato X è detto* **riflessivo** *se l'isometria $\mathfrak{I} : X \to (X')'$, che associa $x \in X$ con $\mathfrak{I}(x)$ secondo la (2.11), è suriettiva (cioè è un isomorfismo di spazi normati).*

In altre parole X è riflessivo quando X e $(X')'$ sono isometricamente isomorfi tramite la trasformazione naturale $\mathfrak{I}$. Vedremo nella prossima sezione se X è uno spazio di Hilbert la riflessività è assicurata. Gli spazi di Banach $L^p(X, \mu)$ introdotti negli esempi 2.1 sono riflessivi se $1 < p < \infty$.

Passiamo alla formulazione del teorema di Banach-Steinhaus nella formulazione più elementare ed a quello dell'applicazione aperta.

Teorema di Banach-Steinhaus. *Sia $\{T_\alpha\}_{\alpha\in A}$ una famiglia di operatori in $\mathfrak{B}(X,Y)$ dove X è uno spazio di Banach e Y uno spazio normato. Se, per ogni $x \in X$*

$$\sup_{\alpha\in A} ||T_\alpha x|| < \infty \,,$$

allora esiste $K \in \mathbb{R}$ tale che

$$||T_\alpha|| < K \quad \textit{per ogni } \alpha \in A \,.$$

Un immediato ed utile corollario del teorema di Banach-Steinhaus è il seguente.

Corollario 1. *Nelle stesse ipotesi del teorema di Banach-Steinhaus la famiglia di operatori $\{T_\alpha\}_{\alpha\in A}$ è* **equicontinua** *ossia, per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che, per ogni coppia $x, x' \in X$, se $||x - x'||_X < \delta$ allora $||T_\alpha x - T_\alpha x'|| < \epsilon$ per ogni $\alpha \in A$.*

Prova. Nel seguito $C_\gamma := \{x \in X_1 \mid ||x||_X \leq \gamma\}$ per ogni $\gamma > 0$. Fissiamo $\epsilon > 0$, dobbiamo trovare il $\delta > 0$ che soddifa la proprietà di scritta nella tesi. Per il teorema di B-S ed usando la proposizione 2.1, $||T_\alpha x||_Y < c < \infty$ per ogni $\alpha \in A$ e $x \in B_1$. Se $c = 0$ non c'è nulla da provare, assumiamo pertanto $c > 0$. Scegliamo $\delta > 0$ per cui $C_\delta \subset C_{\epsilon/c}$. Con le scelte fatte, se $||x - x'||_X < \delta$, vale $c(x - x')/\epsilon \in B_{c\delta/\epsilon} \subset C_1$ e quindi

$$||T_\alpha x - T_\alpha x'||_Y = ||T_\alpha(x - x')||_Y = \frac{\epsilon}{c}||T_\alpha \frac{c(x - x')}{\epsilon}|| < \frac{\epsilon}{c} c = \epsilon \,,$$

indipendentemente da $\alpha \in A$. $\square$

Per introdurre altri due corollari del teorema di Banach-Steinhaus abbiamo bisogno di alcune nozioni topologiche.

Notazione 2.3. Se $\delta > 0$ e p è una seminorma sullo spazio vettoriale X, sul campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$, e $x \in X$, indicheremo con $B_{p,\delta}(x)$ la palla aperta associata a p, centrata sul vettore x e di raggio δ:

$$B_{p,\delta}(x) = \{z \in X \mid p(z - x) < \delta\} \,.$$

Nel caso in cui $x = 0$ scriveremo semplicemente $B_{p,\delta}$ in luogo di $B_{p,\delta}(0)$.
Se $A \subset X$, $x \in X$ e $\beta \in \mathbb{K}$ indicheremo l'nsieme $\{x + \beta u \mid u \in A\}$ con la notazione

$$x + \beta A = \{x + \beta u \mid u \in A\} \,.$$

Se $\mathcal{P} := \{p_i\}_{i\in I}$ è una famiglia di seminorme sullo spazio vettoriale X sul campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$, la **topologia su X generata da** $\mathcal{P}$, $\mathcal{T}(\mathcal{P})$, è l'unica che ammette come base la classe degli insiemi

$$x + \left(B_{p_{i_1},\delta_1} \cap \cdots \cap B_{p_{i_n},\delta_n}\right)$$

per ogni scelta di $x \in X$, di $n = 1, 2, \ldots$, degli indici $i_1, \ldots, i_n \in I$ e dei numeri $\delta_1 > 0, \ldots \delta_n > 0$.

Ricordiamo che in uno spazio topologico X, una successione di punti $\{x_n\}_{n\in\mathbb{N}} \subset X$ è detta *convergere* a $x \in X$ quando: per ogni aperto $A \ni x$ esiste $N_A \in \mathbb{N}$ tale che $x_n \in A$ se $n > N_A$. Si prova facilmente che:

Proposizione 2.2. *Una successione $\{x_n\}_{n\in\mathbb{N}} \subset X$ converge a $x_0 \in X$ nella topologia $\mathcal{T}(\mathcal{P})$ se e solo se, per ogni $p_i \in \mathcal{P}$, $p_i(x_n) \to p_i(x_0)$ quando $n \to \infty$.*

È chiaro che se la classe $\mathcal{P}$ si riduce ad un unico elemento dato da una *norma*, la topologia indotta da $\mathcal{P}$ è la solita topologia metrica indotta da una norma.

Con la nozione introdotta sopra di topologia indotta da una classe di seminorme si possono definire alcune tolologia "standard" su $\mathfrak{L}(X, Y)$, $\mathfrak{B}(X, Y)$ e sul duale X' una di queste topologie (e la corrispondente sul duale) la conosciamo già perché è quella indotta dalla norma operatoriale.

Definizione 2.8. *Siano X, Y spazi normati sullo stesso campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$.*
(a) *Si definiscono su $\mathfrak{L}(X, Y)$ e $\mathfrak{B}(X, Y)$) le seguenti topologie.*
(i) La topologia indotta su $\mathfrak{L}(X, Y)$ (ovvero $\mathfrak{B}(X, Y)$) dalla classe di tutte le seminorme definite come, se $T \in \mathfrak{L}(X, Y)$) (rispettivamente $\mathfrak{B}(X, Y)$),

$$p_{x,f}(T) := |f(T(x))|$$

per ogni $x \in X$ e $f \in Y'$, che è detta **topologia debole** *su $\mathfrak{L}(X, Y)$) (rispettivamente $\mathfrak{B}(X, Y)$);*
(ii) La topologia indotta su $\mathfrak{L}(X, Y)$ (ovvero $\mathfrak{B}(X, Y)$) dalla classe di tutte le seminorme definite come, se $T \in \mathfrak{L}(X, Y)$) (rispettivamente $\mathfrak{B}(X, Y)$),

$$p_x(T) := ||T(x))||_Y$$

per ogni $x \in X$, che è detta **topologia forte** *su $\mathfrak{L}(X, Y)$) (rispettivamente $\mathfrak{B}(X, Y)$)*
(iii) La topologia indotta su $\mathfrak{B}(X, Y)$ dalla norma operatoriale (2.6) è detta **topologia uniforme** *su $\mathfrak{B}(X, Y)$.*
(b) *Se $Y = \mathbb{K}$, la topologia uniforme definita in (iii) prende il nome di* **topologia (duale) forte** *e le topologie debole e forte definite in (i) e (ii), che risultano coincidere, prendono il nome di* **topologia *-debole**.

Note.
(1) Si prova facilmente che la topologia forte è più forte di quella debole (cioè gli aperti della topologia debole sono aperti anche nella topologia forte). Nello stesso modo si prova facilmente che la topologia uniforme è più forte della topologia forte. Per gli spazi duali vale ovviamente l'analoga proprietà: la topologia forte è più forte della topologia *-debole.
(2) Data una successione di operatori $\{T_n\} \subset \mathfrak{L}(X, Y)$ (o $\mathfrak{B}(X, Y)$) e un operatore $T \in \mathfrak{L}(X, Y)$

(o rispettivamente $\mathfrak{B}(X,Y)$), si prova facilemente che:
$T_n \to T$ nella topologia debole, solo se

$$f(T_n(x)) \to f(T(x))$$

per ogni scelta di $x \in X$ e $f \in Y'$.
Nello stesso modo, si prova facilemente che:
$T_n \to T$ nella topologia forte se e solo se

$$||T_n(x) - T(x)||_Y \to 0$$

per ogni scelta di $x \in X$.
È chiaro che la convergenza di una successione di operatori di $\mathfrak{B}(X,Y)$ in senso uniforme (cioè rispetto alla topologia uniforme) implica la convergenza della stessa successione in senso forte (cioè rispetto alla topologia forte).
Similmente la convergenza di una successione di operatori di $\mathfrak{L}(X,,Y)$ o $\mathfrak{B}(X,Y)$ in senso forte implica la convergenza della stessa successione in senso debole (cioè rispetto alla topologia debole).
(3) Data una successione di funzionalii $\{f_n\} \subset X'$ e un funzionale $f \in X'$, si prova facilemente che:
$f_n \to f$ nella topologia *-debole se e solo se

$$f_n(x) \to f(x)$$

per ogni scelta di $x \in X$.
È chiaro che la convergenza di una successione di funzionali di X' in senso forte (cioè rispetto alla topologia duale forte) implica la convergenza della stessa successione in senso *-debole (cioè rispetto alla topologia *-debole).

Notazione 2.4. Per denotare i limiti rispetto alla topologie forti e deboli solitamente si usano le seguenti convenzioni notazionali che adotteremo anche in questo testo.

$$T = s\text{-}\lim T_n$$

significa che T è il limite nella topologia forte della successione di operatori $\{T_n\}$. La stessa notazione si usa nel caso in cui gli operatori coinvolti siano funzionali e la topologia è quella duale forte.

$$T = w\text{-}\lim T_n$$

significa che T è il limite nella topologia debole della successione di operatori $\{T_n\}$.

$$f = w^*\text{-}\lim f_n$$

significa che f è il limite nella topologia *-debole della successione $\{f_n\}$.

Le nozioni topologiche acquisite permettono di provare altri due rilevanti corollari del teorema di Banch-Steinhaus.
Sappiamo già che se X è uno spazio metrico allora X' è completo rispetto alla topologia duale fortecome provato in (ii) di (c) del teorema 2.3 Si può provare che la completezza sussiste anche rispetto alal topologia *-debole purché X sia spazio di Banach.

Corollario 2. *Se X è uno spazio di Banach sul campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$ allora X' è completo anche rispetto alla topologia $*$-debole. In altre parole Se $\{f_n\} \subset X'$ è tale che $\{f_n(x)\}$ è una successione di Cauchy per ogni $x \in X$, allora esiste $f \in X'$ tale che $f = w^*\text{-}lim f_n$.*

Prova. Il campo di X è completo per ipotesi, di conseguenza, per ogni $x \in X$ esisterà $f(x) \in \mathbb{K}$ con $f_n(x) \to f(x)$. È immediato verificare che $f : X \ni x \mapsto f(x)$ definisce un funzionale lineare. Per concludere la dimostrazioneproviamo che f è continuo. Per ogni $x \in X$, $f_n(x)$ è limitata per cui, dal teorema di B-S, $|f_n(x)| < c < \infty$ per ogni $x \in X$ con $||x|| \leq 1$. Facendo il limite a $n \to \infty$ segue che $|f(x)| < c$ se $||x|| \leq 1$ da cui $||f|| < c < \infty$ e pertanto, dal teorema 2.2, f è continuo. □

Corollario 3. *Sia X è uno spazio normato sul campo $\mathbb{C}$ o $\mathbb{R}$. Se $S \subset X$ è debolmente limitato, cioè per ogni $f \in X'$ esiste $c(f) \geq 0$ tale che $|f(x)| \leq c(f)$ per ogni $x \in X$, allora S è limitato rispetto alla norma di X.*

Prova. Consideriamo gli elementi $x \in S \subset X$ come funzionali del duale di X', $(X')'$ facendo uso della trasformazione isometrica $\mathfrak{I} : X \to (X')'$ definita nel corollario 3 al teorema di Hahn-Banach. La famiglia di funzionali su X', $S \subset (X')'$ è limitata su ogni $f \in X'$ per ipotesi essendo (scriviamo x in luogo di $\mathfrak{I}(x)$) $|x(f)| = |f(x)| \leq ||f||||x|| \leq c(f)||x||$. Dato che X' è completo possiamo applicare il teorema di B-S concludendo che $\sup\{|x(f)| \mid ||f|| = 1\} < K < \infty$ per ogni $x \in S$ ossia (essendo $\mathfrak{I}$ un'isometria) $||x|| < K < \infty$ per ogni $x \in S$. □
Concludiamo la sezione con l'ultimo importante teorema generale.

Teorema dell'applicazione aperta. *Siano X e Y due spazi di Banach sullo stesso campo $\mathbb{C}$ o $\mathbb{R}$ e sia $B_1 \subset X$ la palla aperta di raggio 1 centrata in 0. Per ogni $T \in \mathfrak{B}(X, Y)$ suriettivo esiste una palla aperta, $B_\delta \subset Y$, centrata nell'origine di raggio $\delta > 0$ tale che*

$$T(B_1) \supset B_\delta \,. \tag{2.12}$$

Nota. Il nome del teorema è dovuto al seguente fatto evidente. (2.12) e la linearità di T implicano che l'immagine secondo T di qualunque palla aperta $B_\epsilon(x) = x + \epsilon B_1$ centrata in qualunque punto $x \in X$, contiene la palla aperta di Y centrata in $T(x)$: $B_{\delta_\epsilon} := T(x) + \epsilon B_\delta$ (per $\delta > 0$ sufficientemente piccolo). Quindi l'immagine secondo T di un qualsiasi aperto di X, $A = \cup_{x \in A} B_{\epsilon_x}(x)$ risulta essere un aperto di Y: $T(A) = \cup_{x \in A} B_{\delta_{\epsilon_x}}(T(x))$. In altre parole, la funzione T è *aperta*.

La conseguenza elementare più importante del teorema della funzione aperta è sicuramente il **Teorema di Banach dell'operatore inverso** nel caso di spazi di Banch (esiste una versione che usa solo l'potesi di spazi vettoriali dotati metrica e completi).

Corollario 1 (Teorema di Banach dell'operatore inverso). *Siano X e Y due spazi di Banach sullo stesso campo $\mathbb{C}$ o $\mathbb{R}$. Se $T \in \mathfrak{B}(X, Y)$ è iniettivo e suriettivo, allora valgono i seguenti due fatti.*
(a) *$T^{-1} : Y \to X$ è un operatore limitato cioè $T^{-1} \in \mathfrak{B}(X', X)$.*
(b) *Esiste $c > 0$ tale che:*

$$||x|| \leq c||Tx|| \,, \quad \textit{per ogni } x \in X. \tag{2.13}$$

Prova. (a) Il fatto che T^{-1} sia lineare è di immediata verifica. Bisogna solo provare che T^{-1} è continuo. Essendo T aperta la controimmagine di un aperto di X secondo T^{-1} è un aperto, quindi T^{-1} è continua. (b) Osserviamo che la (2.13) è banalmente vera per $x = 0$. Dobbiamo quindi provarla per $x \neq 0$. Se T è *iniettiva*, dalla (2.12) si evince che $||Tu|| \leq \delta \Rightarrow ||u|| < 1$. Per ogni $x \in X \setminus \{0\}$ vale sicuramente,

$$\left|\left| T\left(\frac{\delta x}{||Tx||}\right)\right|\right| \leq \delta$$

per cui, chiamando u l'argomento ti T, dovrà essere $||u|| < 1$ ossia.

$$\left|\left| \frac{\delta x}{||Tx||}\right|\right| \leq 1 \,.$$

Questa non è altro che la (2.13) per $c := 1/\delta$ (notare che $\delta > 0$): $||x|| \leq c||Tx||$. $\square$

Per concludere passiamo a dimostrare un utilissimo teorema in teoria degli operatori chiamato il teorema del grafico chiuso.

Notazione 2.5. Se $\emptyset \neq K_1, \ldots, K_n \subset X$, con X spazio vettoriale, $< K_1, \cdots, K_n >$ denota il sottospazio vettoriale di X *generato* dai sottospazi K_i cioè il sottospazio di X contenente tutte le combinazioni lineari (finite) di elementi degli spazi K_i. Se $K \subset X$ e $\emptyset \neq K_1, \ldots, K_n \subset X$ sono sottospazi, $K = K_1 \oplus \cdots \oplus K_n$ significa che K è *somma diretta* degli spazi K_i. In altre parole vale che $K =< K_1, \cdots, K_n >$ e $K_i \cap K_j = \{0\}$ per ogni coppia di indici i, j con $i \neq j$, ovvero equivalentemente, ogni vettore di K può essere decomposto in maniera unica come una somma di vettori degli spazi K_i

Per dimostrare il teorema abbiamo bisogno di alcuni risultati preliminari. Prima di tutto notiamo che se (X, N_X) e (Y, N_Y) sono spazi normati sullo stesso campo $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$, possiamo considerare lo spazio vettoriale S costituito sul prodotto cartesiano $X \times Y$ definendo le operazioni di prodotto per scalare e somma di vettori come: $\alpha(x, y) = (\alpha x, \alpha y)$ e $(x, y) + (x', y') = (x + x'6, y + y')$. È

chiaro che X e Y si identificano naturalmente con due sottospazi di S, che indicheremo ancora con X e Y rispettivamente, e che che vale $S = X \oplus Y$. S è dotato naturalmente della *topologia prodotto* indotta da quella dei fattori X e Y.
(Ricordiamo che la topologia prodotto su uno spazio vettoriale prodotto $T = \times_{\alpha \in A} T_\alpha$ è la meno forte topologia per cui tutte le proiezioni canoniche $\Pi_\alpha : X \to X_\alpha$ sono continue.
Nel caso di un numero finito di fattori (come nel caso considerato), tale topologia ammette come base di aperti i prodotti prodotti cartesiani di tutte le palle aperte negli spazi X_α.)
Le operazioni della struttura di spazio vettoriale di S sono continue rispetto alla topologia prodotto di X e Y.
La topologia prodotto di S è **normabile**: cioè esiste un norma su S la cui topologia metrica associata è proprio la topologia prodotto di S. Tale norma è definita come

$$||(x, y)|| := \max\{N_x(x), N_Y(y)\} \quad \text{per ogni } (x, y) \in X \oplus Y \,. \tag{2.14}$$

Il fatto che questa norma generi la stessa topologia generata dei prodotti delle palle metriche di X e Y si prova come segue. Considerando il prodotto di due palle aperte $B_\delta(x_0) \times B_\mu(y_0)$ rispettivamente in X e Y, la palla aperta di $X \oplus Y$:

$$\{(x, y) \in X \times Y \mid ||(x, y) - (x_0, y_0)|| < \min\{\delta, \mu\}\}$$

è contenuta in $B_\delta(x_0) \times B_\mu(y_0)$. Viceversa, Il prodotto $B_\delta(x_0) \times B_\delta(y_0)$ e contenuto nella palla metrica

$$\{(x, y) \in X \times Y \mid ||(x, y) - (x_0, y_0)|| < \epsilon\} \,,$$

se $\epsilon > \delta$. Questo conclude la prova perchè ne segue che ogni palla metrica di $X \oplus Y$ è unione di prodotti di palle metriche di X e Y e ogni prodotto di palle metriche in X e Y è unione di palle metriche di $X \oplus Y$ e quindi le due topologie coincidono.
Si prova immediatamente che $(X \oplus Y, ||\,||)$ *è uno spazio di Banach se sono tali* (X, N_X) *e* (Y, N_Y).
(Questo fatto implica, in base alla proposizione 2.3 della prossima sezione, che qualunque norma che generi la topologia prodotto rende $X \oplus Y$ spazio di Banach.) Sia infatti $\{(x_n, y_n)\}$ una successione di Cauchy in $X \oplus Y$, allora $\{x_n\}$ e $\{y_n\}$ risultano essere di Cauchy in X e Y rispettivamente dalla stessa definizione di norma su $X \oplus Y$ data sopra. Siano $x \in X$ e $y \in Y$ i limiti di queste due successioni che esistono essendo X e Y spazi di Banach. Quindi, se $\epsilon > 0$ esistono N_x e N_y interi positivi tali che, se $n > \max\{N_x, N_y\}$ allora

$$||(x, y) - (x_n, y_n)|| < \epsilon \,,$$

e quindi $(x_n, y_n) \to (x, y)$ per $n \to +\infty$ nella topologia indotta dalla norma di $X \oplus Y$. Tale spazio è quindi spazio di Banach.

Definizione 2.9. *Siano X, Y spazi normati sullo stesso campo $\mathbb{C}$ o $\mathbb{R}$. $T \in \mathfrak{L}(X, Y)$ è detto* **chiuso** *se il* **grafico** *di T, cioè il sottospazio $G(T) := \{(x, Tx) \mid x \in X\} \subset X \oplus Y$, è chiuso nella topologia prodotto. In altre parole T è chiuso quando per ogni successione $\{x_n\} \subset X$ tale*

che $\{x_n\}$ *e* $\{Tx_n\}$ *convergono in* X *e* Y *rispettivamente,* $\lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n)$.

Possiamo enunciare e dimostrare il *teorema dell'opertaore chiuso o grafico chiuso.*

Corollario 2 (Teorema dell'operatore chiuso o del grafico chiuso). *Siano* X, Y *spazi di Banach sullo stesso campo* $\mathbb{K} = \mathbb{C}$ *o* $\mathbb{R}$ *e sia* $T : X \to Y$ *un operatore lineare.* T *è limitato se e solo se è chiuso.*

Prova. Se T è limitato allora è banalmente chiuso. Si supponga quindi che T sia chiuso e mostriamo che è limitato. Consideriamo l'applicazione lineare biettiva $M : G(T) \ni (x, Tx) \mapsto x \in X$. $G(T)$ è per ipotesi un sottospazio chiuso nello spazio di Banach $X \oplus Y$ per cui è a sua volta uno spazio di Banach rispetto alla restrizione della norma $|| \, ||$ (2.14). Dalla definizione di tale norma troviamo che vale banalmente $||M(x, Tx)||_X = ||x||_X \leq ||(x, Tx)||_{G(T)}$ per cui M è limitato. Per il teorema dell'inverso limitato di Banach $M^{-1} : X \to G(T) \subset X \oplus Y$ è limitato. Dato che la proiezione canonica di $\Pi_Y : X \oplus Y \to Y$ è continua nella topologia prodotto, concludiamo che l'applicazione lineare $\Pi_Y \circ M^{-1} : x \mapsto Tx$ per ogni $x \in X$ è continua ossia limitata. $\square$

2.4 Proiettori.

Introduciamo, usando il teorema del grafico chiuso, la nozione e le proprietà elementari di alcuni operatori continui detti *proiettori.* La nozione di proiettore giocherà il ruolo central nella formulazione della Meccanica Quantistica quando lo spazio normato sarà uno spazio di Hilbert.

Definizione 2.10 *Sia* $(X, || \, ||)$ *spazio normato sul campo* $\mathbb{C}$ *o* $\mathbb{R}$. $P \in \mathfrak{B}(X)$ *è detto* **proiettore** *se è idempotente, cioè se*

$$PP = P \, . \tag{2.15}$$

$M := P(X)$ *è detto* **spazio di proiezione** *di* P *e si dice anche che* P **proietta** *su* M.

I proiettori sono naturalmente associati ad una decomposizione diretta dello spazio ambiente X in una coppia di sottospazi chiusi.

Proposizione 2.4. *Sia* $P \in \mathfrak{B}(X)$ *un proiettore sullo spazio normato* $(X, || \, ||)$. *Vale quanto segue.*
(a) *Se* $Q : X \to X$ *è l'applicazione lineare che soddisfa*

$$Q + P = I \, , \tag{2.16}$$

Q *è proiettore e*

$$PQ = QP = 0 \, , \tag{2.17}$$

dove 0 *è l'operatore nullo (quello che trasforma nel vettore* 0 *ogni vettore di* X*).*
(b) *Gli spazi di proiezione* $M := P(X)$ *ed* $N := Q(X)$ *sono sottospazi chiusi di* X *e vale*

$$X = M \oplus N \,. \tag{2.18}$$

Prova. (a) Q è continuo perè somma di operatori continui, $QQ = (I-P)(I-P) = I-2P+PP = I-2P+P = I-P = Q$. $PQ = P(I-P) = P-PP = P-P = O$, $(I-P)P = P-PP = P-P = O$.
(b) Se $P(x_n) \to y$ per $n \to \infty$ allora, dalla continuità di P: $PP(x_n) \to P(y)$. Per (2.15) questo si scrive $P(x_n) \to P(y)$ da cui $y = P(y)$ per l'unicità del limite (essendo X di Hausdorff). In definitiva $y \in \overline{M}$ implica $y \in M(\subset \overline{M})$ per cui $M = \overline{M}$ ed M è dunque un insieme chiuso. Con lo stesso ragionamento si prova che N è chiuso. Il fatto che M ed N siano sottospazi è immediata conseguenza della linearità di P e Q rispettivamente. Sia ora $x \in X$, dalla definizione di Q vale

$$x = P(x) + Q(x) \,,$$

di conseguenza $X =< M, N >$. Per concludere basta provare che $M \cap N = \{0\}$. Se $x \in M, N$ allora $x = P(x)$ e $x = Q(x)$ per (2.15): per es. se $x \in M$ allora $x = Pz$ per qualche $z \in X$, ma allora $Px = PPz = Pz = x$. Applicando Q alla identità $x = Px$ e tenendo conto di $x = Qx$ si ha: $x = Q(x) = QP(x)$. Ma $QP(x) = 0$ per (2.17). Abbiamo ottenuto che $x \in M \cap N$ implica $x = 0$ che è quanto volevamo provare. □

Il teorema del grafico chiuso prova che la proposizione 2.4 può essere invertita come segue assumendo l'ulteriore ipotesi che lo spazio ambiente sia di Banach.

Proposizione 2.5. *Sia* $(X, || \, ||)$ *spazio di Banach e* $M, N \subset X$ *due sottospazi chiusi tali che* $X = M \oplus N$*. Si considerino le applicazioni* $P : X \to M$ *e* $Q : X \to N$ *che associano a* $x \in M$ *i rispettivi elementi di* M *ed* N *secondo la decomposizione diretta* $X = M \oplus N$*. Allora vale quanto segue.*
(a) P *e* Q *sono proiettori che proiettano rispettivamente su* M *ed* N*.*
(b) *Valgono le idnetità (2.16) e (2.17).*

Prova. Per ipotesi, se $x \in X$ allora vale la decomposizione $x = u_M + u_N$ con $u_M \in M$ e $u_n \in N$ e tale decomposizione è l'unica possibile rispetto a tale coppia di sottospazi. L'unicità ed il fatto che M ed N sono rispettivamente chiusi per combinazioni lineari implicano che le funzioni $P : x \mapsto u_M$ e $Q : x \mapsto u_N$ siano lineari e che valga $PP = P$ e $QQ = Q$. Si noti anche che $P(X) = M$ e $Q(X) = N$ per costruzione e che (2.16) vale in quanto $X =< M, N >$, mentre (2.17) vale perchè $M \cap N = \{0\}$. Per concludere la dimostrazione bisogna solo provare che P e Q sono continui. Mostriamo che P è chiuso, il teorema dell'operatore chiuso implica allora che P è continuo. La dimostrazione per Q è analoga. Sia $\{x_n\} \subset X$ una successione che converge a

$x \in X$ e tale che $\{Px_n\}$ converge anch'essa ad un elemento di X. Dobbiamo provare che

$$Px = \lim_{n\to+\infty} Px_n \, .$$

Dato che N è chiuso:

$$N \ni Qx_n = x_n - Px_n \to x - \lim_{n\to\infty} Px_n = z \in N \, .$$

Abbiamo ottenuto che

$$x = \lim_{n\to+\infty} Px_n + z \, ,$$

con $z \in N$ ma anche $\lim_n Px_n \in M$ perchè M è chiuso e $Px_n \in M$ per ogni n. D'altra parte sappiamo che vale anche

$$x = Px + Qx$$

Per l'unicità della decomposizione di x come somma di vettori in N ed M deve essere

$$Px = \lim_{n\to+\infty} Px_n$$

e $z = Qx$ quindi l'operatore P è chiuso ed è quindi continuo. □

2.5 *Norme equivalenti.

Un'interessante applicazione del teorema dell'operatore inverso di Banach è un criterio per stabilire quando due norme sullo stesso spazio vettoriale, che risulta essere completo per entrambe, sono associate alla stessa topologia. Introduciamo quanche nozione prima di enunciare il criterio detto nella proposizione 2.4

Definizione 2.11. *Due norme N_1 e N_2 sullo stesso spazio vettoriale X (su $\mathbb{C}$ o $\mathbb{R}$) si dicono* **equivalenti** *se esistono due costanti $c, c' > 0$ tali che:*

$$cN_2(x) \leq N_1(x) \leq c'N_2(x) \, , \quad \textit{per ogni } x \in X. \tag{2.19}$$

Osservazioni.
(1) Si osservi che (2.19) è equivalente all'analoga disuguaglianza che si ottiene scambiando N_1 con N_2 e sostituendo c, c' con $1/c', 1/c$ rispettivamente.
(2) Si prova immediatamente che *la proprietà di completezza di uno spazio normato è invariante sotto lo scambio della norma iniziale con un'altra equivalente.*
(3) Se due norme sullo stesso spazio sono equivalenti allora generano la stessa topologia. La proposizione di sotto prova che vale anche il viceversa.

Proposizione 2.6. *Sia X uno spazio vettoriale su $\mathbb{C}$ o $\mathbb{R}$. Le norme N_1 e N_2 su X sono equivalenti se e solo se l'applicazione identica $I : (X, N_2) \ni x \mapsto x \in (X, N_1)$ è un omeomorfismo (che equivale a dire che le topologie metriche generate dalle due norme coincidono).*

Prova. Bisogna solo dimostrare che se I è omeomorfismo allora le norme sono equivalenti. L'altra implicazione è ovvia dalla definizione di norme equivalenti. Se I è un omeomorfismo allora è continuo nell'origine e quindi, in particolare, la palla metrica rispetto alla norma N_1 di raggio 1 e centrata in 0 deve contenere completamente una palla aperta rispetto alla norma N_2 centrata in 0 di raggio $\delta > 0$ sufficientemente piccolo. Ossia $N_2(x) \leq \delta \Rightarrow N_1(x) < 1$. In particolare, per $x \neq 0$, $N_2(\delta x/N_2(x)) < \delta$ per cui: $N_1(\delta x/N_2(x)) < 1$ ossia $\delta N_1(x) < N_2(x)$. Per $x = 0$ tale identità vale banalmente. Abbiamo provato che esiste $c' = 1/\delta > 0$ per cui $N_1(x) \leq c' N_2(x)$ per ogni $x \in X$. L'altra parte della (2.19) si prova analogamente scambiando il ruolo dei due spazi. □

Proposizione 2.7. *Sia X uno spazio vettoriale su $\mathbb{C}$ o $\mathbb{R}$. Si supponga che X sia spazio di Banach rispetto ad entrambe le norme N_1, N_2. Se esiste $k > 0$ tale che, per ogni $x \in X$:*

$$N_1(x) \leq k N_2(x)$$

allora le due norme sono equivalenti.

Prova. Si consideri la funzione identità $I : x \mapsto x$. Tale funzione è lineare e continua quando la pensiamo come $I : (X, N_1) \to (X, N_2)$ perché $N_2(x) \leq (1/k) N_1(x)$ per ogni $x \in X$. Per il punto (b) nella tesi del teorema di Banach della funzione inversa esisterà $c > 0$ tale che $N_1(x) \leq c N_2(x)$ per ogni $x \in X$. N_1 e N_2 soddisfano allora la (2.19) con $c' := k$. □

Capitolo 3

Spazi di Hilbert ed operatori limitati.

In questa sezione introdurremo alcune prime nozioni matematiche relative agli spazi di Hilbert che ci serviranno per dare un primo fondamento matematico alla MQ.

3.1 Nozioni elementari, teorema di Riesz e riflessività.

Definizione 3.1. *Se X è uno spazio vettoriale sul campo complesso, un 'applicazione $S : X \times X \to \mathbb{C}$ si dice* **prodotto scalare (hermitiano)** *e* (X, S) spazio con prodotto scalare *quando S soddisfa:*

H0. $S(u, u) \geq 0$ *per ogni* $u \in X$*;*

H1. $S(u, \alpha v + \beta w) = \alpha S(u, v) + \beta S(u, w)$ *per ogni* $\alpha, \beta \in \mathbb{C}$ *e* $u, v, w \in X$*;*

H2. $S(u, v) = \overline{S(v, u)}$ *per ogni* $u, v \in X$*;*

H3. $S(u, u) = 0 \Rightarrow u = 0$*, pero ogni* $u \in X$*.*

Se valgono **H0**, **H1**, **H2** *ma non necessariamente* **H3**, *S si dice* **semi prodotto scalare**.
Nello spazio vettoriale X con (semi) prodotto scalare hermitiano S, vettori $u, v \in X$ si dicono **ortogonali** *se:*

$$S(u, v) = 0 \, .$$

In tal caso si dice anche che u è **ortogonale**, *o equivalentemente* **normale**, *a v e v è* **ortogonale**, *o equivalentemente,* **normale** *a u.*
Se $\emptyset \neq K \subset X$, lo **spazio ortogonale a** *K, $K^{\perp}$, è l'insieme dei vettori di X ortogonali a ciascun elemento di K.*

Note.
(1) H1 e **H2** implicano insieme che S è *antilineare* nel primo argomento:

$$S(\alpha v + \beta w, u) = \overline{\alpha} S(v, u) + \overline{\beta} S(w, u)$$

per ogni $\alpha, \beta \in \mathbb{C}$ e $u, v, w \in X$.
(2) Dalla **H2** segue che u è ortogonale a v se e solo se v è ortogonale a u, per cui la definizione di ortogonalità è ben posta.
(3) Si verifica immediatamente che $K^\perp$ è un sottospazio vettoriale di X per la proprietà **H1**, per cui il termine *spazio* ortogonale è appropriato.

Proposizione 3.1. *Sia X uno spazio vettoriale su $\mathbb{C}$ dotato di un semi prodotto scalare S.*
(a) *Vale la* **disuguaglianza di Cauchy-Schwartz** *per ogni coppia $x, y \in X$*

$$|S(x,y)|^2 \leq S(x,x)S(y,y)\,, \tag{3.1}$$

in (3.1) vale il segno di uguaglianza se x e y sono linearmente dipendenti e solo in tal caso se S è prodotto scalare.
(b) *Al variare di $x \in X$,*

$$p(x) := \sqrt{S(x,x)} \tag{3.2}$$

definisce una seminorma, ovvero un norma se S è un prodotto scalare, che si dice essere **indotta da** *S.*
(c) *p soddisfa* **l'identità del parallelogramma***:*

$$p(x+y)^2 + p(x-y)^2 = 2(p(x)^2 + p(y)^2)\,, \tag{3.3}$$

per ogni $x, y \in X$.

Prova. (a) Se $\alpha \in \mathbb{C}$ vale, usando le proprietà del semi prodotto scalare,

$$0 \leq S(x-\alpha y, x-\alpha y) = S(x,x) - \alpha S(x,y) - \overline{\alpha} S(x,y) + |\alpha|^2 S(y,y)\,. \tag{3.4}$$

Supponiamo $S(y,y) \neq 0$. In tal caso ponendo $\alpha := \overline{S(x,y)}/S(y,y)$ si ottiene da (3.4)

$$0 \leq S(x,x) - |S(x,y)|^2/S(y,y)\,,$$

che è la tesi. Se $S(y,y) = 0$ troviamo invece da (3.4) che deve comunque valere per ogni $\alpha \in \mathbb{C}$:

$$0 \leq S(x,x) - \alpha S(x,y) - \overline{\alpha} S(y,x)\,.$$

Scegliendo $\alpha \in \mathbb{R}$ sufficientemente grande in valore assoluto, si vede che la disuguaglianza (3.4) è impossibile a meno che $S(x,y) + S(y,x) = 0$. Scegliendo poi $\alpha = i\lambda$ con $\lambda \in \mathbb{R}$ sufficientemente grande in valore assoluto, si vede che la disuguaglianza (3.4) può valere solo se oltre all'identità $S(x,y) = -S(y,x)$ provata sopra, vale anche $S(x,y) - S(y,x) = 0$ e quindi $S(x,y) = 0$. In definitiva, se $S(y,y) = 0$ vale ancora (3.1) perchè $S(x,y) = 0$. Se x e y sono linearmente dipendenti allora vale $x = \alpha y$ oppure $y = \alpha x$ per qualche $\alpha \in \mathbb{C}$. Si verifica immediatamente che in tal caso i due membri di (3.1) sono uguali. Supponiamo ora che S sia un prodotto scalare e che valga $|S(x|y)|^2 = S(x|x)S(y|y)$. Se almeno uno dei due vettori x, y è nullo (ed in tal caso l'uguaglianza vale) allora è banalmente vero che x e y sono linearmente dipendenti.

Mettiamoci nel caso in cui entrambi i vettori siano non nulli e valga l'uguaglianza detta. In tal caso, ridefinendo $u = x/\sqrt{S(x|x)}$ e $v = y/\sqrt{S(y|y)}$, vale $|(u|v)| = 1$ e quindi $(u|v) = e^{i\eta}$ per qualche $\eta \in \mathbb{R}$. Per **H3** l'equazione in α e β, $\alpha u + \beta v = 0$ equivale a $S(\alpha u + \beta v|\alpha u + \beta v) = 0$. sviluppandola si ottiene che essa equivale a

$$|\alpha|^2 + |\beta|^2 + \overline{\alpha}\beta e^{i\eta} + \overline{\beta}\alpha e^{-i\eta} = 0\,.$$

Scegliendo $\alpha = e^{i\mu}$ e $\beta = e^{i\nu}$ in modo che $-\mu + \nu + \eta = \pi$, l'equazione di sopra risulta essere soddisfatta. Abbiamo trovato che, per opportuni $\mu, \nu \in \mathbb{R}$ vale

$$e^{i\mu}\sqrt{S(y|y)}x + e^{i\nu}\sqrt{S(x|x)}y = 0$$

ossia che x e y sono linearmente dipendenti.
(b) Tutte le proprietà delle seminorme si verificano immediatamente dalla definizione di p e dalle proprietà del prodotto scalare, eccetto la disuguaglianza trinagolare **N2** che proviamo ora. Dalle proprietà del prodotto scalare risulta subito che

$$p(x+y)^2 = p(x)^2 + 2ReS(x,y) + p(y)^2\,,$$

doeve $Rez := (z + \overline{z})/2$ indica la parte reale del numero complesso z. $2ReS(x,y) \leq 2|S(x,y)|$ e quindi per (3.1), $2ReS(x,y) \leq 2p(x)p(y)$, da cui si ottiene:

$$p(x+y)^2 \leq p(x)^2 + 2p(x)p(x) + p(y)^2\,,$$

ossia

$$p(x,y)^2 \leq (p(x) + p(y))^2\,,$$

che implica **N2**. La proprietà **N3** segue immediatamente da **H3**, nel caso S sia un prodotto scalare.
(c) si verifica immediatamente usando la definizione di p e le proprietà del prodotto scalare. □

Nota. La disuguaglianza di Cauchy-Schwartz ha come immediata conseguenza il fatto che un prodotto scalare $S : X \times X \to \mathbb{C}$ è una funzione *continua* sullo spazio $X \times X$ dotato della topologia prodotto, quando su X si usi la *topologia metrica indotta dal prodotto scalare* tramite la norma (3.2). In particolare il prodotto scalare è anche continuo nei suoi singoli argomenti separatamente.

Definizione 3.2. *Se (X, S_X) e (Y, S_Y) sono spazi con prodotto scalare (hermitiano), un'applicazione lineare $L : X \to Y$ è detta* **isometria** *se soddisfa $S_Y(L(x), L(y)) = S_X(x,y)$ per ogni $x, y \in X$.*
Se l'isometria $L : X \to Y$ è anche surgettiva è detta **isomorfismo di spazi con prodotto scalare**.
Se esiste un isomorfismo di spazi con prodotto scalare (L) dallo spazio X allo spazio Y, tali spazi si dicono **isomorfi (secondo L)**.

È chiaro che ogni isometria $L : X \to Y$ è iniettiva per **H3**, ma può *non* essere surgettiva (anche se $X = Y$ se la dimensione dello spazio X non è finita). Inoltre ogni isometria è ovviamente continua rispetto alle due topologie metriche dei due spazi indotte dalle norme indotte dai prodotti scalari e, se è surgettiva (cioè se è un isomorfismo), la sua inversa è ancora un'isometria (e quindi un isomorfismo).

Esercizi 3.1.
(1) Mostrare che se, su uno spazio vettoriale sul campo $\mathbb{C}$, una seminorma (norma) p soddisfa l'identità del parallelogramma allora esiste, ed è unico, un semi prodotto scalare (rispettivamente un prodotto scalare) S che indice p tramite (3.2).

Suggerimento. Se S è un semiprodotto scalare vale

$$4S(x, y) = S(x + y, x + y) - S(x - y, x - y) - iS(x + iy, x + iy) + iS(x - iy, x - iy).$$

(2) Si può dare la definizione di (semi) prodotto scalare anche per spazi vettoriali sul campo reale, semplicemente cambiando **H2** in $S(u, v) = S(v, u)$ ed usando ovviamente combinazioni lineari a coefficienti in $\mathbb{R}$ in **H1**. Mostrare che vale ancora la proposizione 3.1
(3) Si consideri uno spazio vettoriale con campo reale. Provare che se una seminorma (norma) p su tale spazio soddisfa l'identità del parallelogramma allora esiste, ed è unico, un semi prodotto scalare (rispettivamente un prodotto scalare) S che indice p tramite (3.2).

Suggerimento. Se S è un semiprodotto scalare vale

$$4S(x, y) = S(x + y, x + y) - S(x - y, x - y).$$

(4) Dimostrare quanto asserito nella nota sotto la proposizione 3.1 per uno spazio con prodotto scalare (X, S).

Suggerimento. Se $X \times X \ni (x_n, y_n) \to (x, y) \in X \times X$ per $n \to \infty$, usando la disuguaglianza di C.-S. provare che se S è il prodotto scalare e p la norma associata, allora vale

$$|S(x, y) - S(x_n, y_n)| \leq p(x_n)p(y_n - y) + p(x_n - x)p(y) \ ,$$

e tenere conto che $p(x_n) \to p(x)$ e che la proiezione canonica è continua nella topologia prodotto.
(5) Siano (X, S) e (X_1, S_1) spazi con prodotto scalare hermitiano. Sia $L : X \to X_1$ lineare e tale che $S_1(L(x), L(x)) = S(x, x)$ per ogni $x \in X$. Provare che L è un isometria (di spazi con prodotto scalare). Si può generalizzare il risultato al caso di spazi sul campo reale?

Suggerimento. Usare i suggerimenti agli esercizi (1) e (3).

Notazione 3.1. Nel seguito, se non sarà specificato altrimenti, $(\ |\ )$ denoterà un prodotto scalare e $||\ \ ||$ la norma indotta da esso come in proposizione 3.1

Definizione 3.3. *Uno spazio con prodotto scalare (hermitiano) si dice* **spazio di Hilbert** *se è spazio di Banach rispetto alla norma indotta dal prodotto scalare.*

Vale un teorema del completamento analogo a quello valido per spazi di Banach.

Teorema del completamento per spazi di Hilbert. *Sia X uno spazio vettoriale sul campo $\mathbb{C}$ dotato di un prodotto scalare S.*
(a) *Esiste uno spazio di di Hilbert $(H, (\ |\ ))$ detto* **completamento** *di X tale che X si identifica con un sottospazio denso (rispetto alla topologia metrica di H indotta dal prodotto scalare $(\ |\ )$) di H tramite un applicazione lineare iniettiva $J : X \to H$ che estende il prodotto scalare S in $(\ |\ )$. In altre parole, esiste un applicazione lineare iniettiva $J : X \to H$ con*

$$\overline{J(X)} = H \quad e \quad (J(x)|J(y)) = S(x, y) \quad per\ ogni\ x, y \in X.$$

(b) *Se la terna $(J_1, H_1, (\ |\ )_1)$ con $J_1 : X \to H_1$ lineare e $(H_1, (\ |\ )_1)$ spazio di Hilbert è tale che X si identifica con un sottospazio denso di H_1 tramite J_1 estendendo il prodotto scalare S in $(\ |\ )_1$, allora esiste ed è unico un isomorfismo di spazi con prodotto scalare hermitiano $\phi : H \to H_1$ tale che $J_1 = \phi \circ J$.*

Schema della dimostrazione. **(a)** Conviene usare il teorema del completamento per gli spazi di Banach e costruire il completamento di Banach dello spazio normato (X, N) dove $N(x) := \sqrt{S(x, x)}$. Dato che S è continuo e X è denso nel completamento secondo la funzione lineare J, S induce un semiprodotto scalare $(|)$ sul completamento di Banach H. $(|)$ è un prodotto scalare su H in quanto, sempre per continutà risulta indurre la stessa norma dello spazio di Banach. Di conseguenza H è spazio di Hilbert e la funzione lineare J che identifica X con un sottospazio denso di H soddisfa le proprietà prescritte nella tesi. (b) La dimostrazione è essenzialmente uguale a quella data nel caso di spazi di Banach. □

Esempi 3.1.
(1) $\mathbb{C}^n$ dotato del prodotto scalare $(u|v) := \sum_{i=1}^n \overline{u_i} v_i$ dove $u = (u_1, \ldots, u_n)$ e $v = (v_1, \ldots, v_n)$ è uno spazio di Hilbert.
(2) Un esempio di spazio di Hilbert estremamente importante in fisica si ricava dall'esempio (6) in esempi 2.1: lo spazio $L^2(X, \mu)$, dove X è uno spazio con misura positiva μ su una σ-algebra Σ. Sappiamo che $L^2(X, \mu)$ è uno spazio di Banach rispetto alla norma

$$||f||_2 := \int_X \overline{f(x)} f(x) d\mu(x)$$

essendo f un qualsiasi rappresentante della classe di equivalenza di funzioni uguale quasi ovunque rispetto a μ, $f \in L^2(X, \mu)$. Usando essenzialmente la stessa dimostrazione per provare la disuguaglianza di Cauchy-Schwartz e tenendo conto della finitezza delle norme degli elementi dello spazio di Banach $L^2(X, \mu)$, è facile provare che

$$(f|g) := \int_X \overline{f(x)} g(x) d\mu(x) \tag{3.5}$$

è finito se $f, g \in L^2(X, \mu)$ e definisce un prodotto scalare hermitiano su tale spazio di Banach. È ovvio che tale prodotto scalare induce la norma $||\ ||_2$ per cui $L^2(X, \mu)$

e uno spazio di Hilbert rispetto al prodotto scalare (3.5).

Passiamo ora a provare che gli spazi di Hilbert sono *riflessivi*. Per fare ciò dobbiamo introdurre alcuni utili strumenti legati al concetto di spazio ortogonale di centrale importanza nello sviluppo del formalismo della MQ.

Notazione 3.2. Se $K \subset X$, con X spazio topologico, $\overline{X}$ indica la *chiusura* di X, cioè il più piccolo chiuso (nel senso dell'intersezione) che include K, che è equivalente a dire che $\overline{X}$ è l'unione di K e di tutti i suoi punti di accumulazione.

Ricordiamo che se $\emptyset \neq K \subset X$, con X spazio vettoriale, K è detto **convesso** se $\alpha u + \beta v \in K$ quando $\alpha, \beta \geq 0$, $\alpha + \beta = 1$ e $u, v \in K$. Chiaramente ogni sottospazio di X è convesso, ma non tutti i sottoinsiemi convessi di X sono sottospazi di X. Per esempio una palla aperta di raggio finito rispetto alla norma su uno spazio vettoriale è un insieme convesso che non è sottospazio.

Teorema 3.1. *Sia $(H, (\ |\ ))$ spazio di Hilbert e $K \subset H$ un sottoinsieme non vuoto.*
(a) *$K^\perp$ è un sottospazio chiuso di H.*
(b) $K^\perp = < K >^\perp = \overline{< K >^\perp} = \overline{< K >}^\perp$.
(c) *Se K è chiuso e convesso allora, per ogni $x \in H$ esiste ed è unico $P_K(x) \in K$ tale che $||x - P_K(x)|| = \min\{||x - y|| \mid y \in K\}$ dove $||\ ||$ denota la norma indotta da $(\ |\ )$.*
(d) *Se K è un sottospazio chiuso allora ogni $x \in H$ si decompone in modo unico come $z_x + y_x$ con $z_x \in K$ e $y_x \in K^\perp$ e quindi vale:*

$$H = K \oplus K^\perp \,. \tag{3.6}$$

Infine risulta $z_x := P_K(x)$ definito in (c).
(e) $(K^\perp)^\perp = \overline{< K >}$.

Nota. In realtà (a) e (b) valgono anche se lo spazio non è di Hilbert ma solo spazio vettoriale con prodotto scalare e la topologia è quella indotta dal prodotto scalare.

Prova del teorema 3.1. (a) $K^\perp$ è sottospazio per le proprietà di (anti) linearità del prodotto scalare. Dalla continuità del prodotto scalare segue subito che se $\{x_n\} \subset K^\perp$ converge a $x \in H$ allora $(x|y) = 0$ per ogni $y \in K$, per cui $x \in K^\perp$. Quindi $K^\perp$ è chiuso contenendo tutti i suoi punti di accumulazione.
(b) La dimostrazione della prima identità è ovvia dalla definizione di ortogonale e dalla linearità (antilinerarità) del prodotto scalare. La seconda identità segue quindi immediatamente da (a). Riguardo all'ultima identità notiamo che sicuramente, essendo $< K > \subset \overline{< K >}$ vale $< K >^\perp \supset \overline{< K >}^\perp$. D'altra parte $< K >^\perp \subset \overline{< K >}^\perp$ per la continuità del prodotto scalare. Di conseguenza $< K >^\perp = \overline{< K >}^\perp$ che conclude la catena di uguaglianze avendo già provato che $< K >^\perp = \overline{< K >^\perp}$.
(c) Sia $0 \leq d = \inf_{y \in K} ||x - y||$ che esiste perchè l'insieme dei numeri $||x - y||$ con $y \in K$ è limi-

tato inferiormente e non vuoto. Sia $\{y_n\} \subset K$ una successione di punti per cui $||x - y_n|| \to d$. Mostriamo che si tratta di una successione di Cauchy. Dall'dentità del parallelogramma (3.3), con x e y in tale formula sostituiti con $x - y_n$ e $x - y_m$ rispettivamente, segue che

$$||y_n - y_m||^2 = 2||x - y_n||^2 + 2||x - y_m||^2 - ||2x - y_n - y_n||^2 \,.$$

Osserviamo ora che $||2x - y_n - y_n||^2 = 4||x - (y_n + y_m)/2||^2 \geq 4d$ dato che $y_n/2 + y_m/2 \in K$ nell'ipotesi di convessità di K e che d è l'estremo inferiore dei numeri $||x - y||$ se $y \in K$. Fissato $\epsilon > 0$, prendendo n e m grandi a sufficienza avremo: $||x - y_n||^2 \leq d^2 + \epsilon$, $||x - y_m||^2 \leq d^2 + \epsilon$, da cui

$$||y_n - y_m||^2 \leq 4(d^2 + \epsilon) - 4d^2 = 4\epsilon \,.$$

Quindi la successione è di Cauchy. Essendo H completo y_n converge a qualche $y \in K$ dato che K è chiuso. Per la continuità della norma $d = ||x - y||$. Mostrimo che $y \in K$ è l'unico punto che soddisfa $d = ||x - y||$. Ogni altro $y' \in K$ con la stessa proprietà soddisfa, dall'dentità del parallelogramma:

$$||y - y'||^2 = 2||x - y||^2 + 2||x - y'||^2 - ||2x - y - y'||^2 \leq 2d^2 + 2d^2 - 4d^2 = 0 \,,$$

dove abbiamo ancora usato $||2x - y - y||^2 = 4||x - (y + y)/2||^2 \geq 4d$ (dato che $y/2 + y'/2 \in K$ nell'ipotesi di convessità di K e che d è l'estremo inferiore dei numeri $||x - z||$ se $z \in K$). Valendo $||y - y'|| = 0$ deve essere $y = y'$. Quindi $P_K(x) := y$ soddisfa tutte le richieste.
(d) Sia $x \in H$ e $x_1 \in K$ a distanza minima da x. Poniamo $x_2 := x - x_1$ e mostriamo che $x_2 \in K^\perp$. Sia $y \in K$, la funzione $\mathbb{R} \ni t \mapsto f(t) := ||x - x_1 + ty||^2$ ha un minimo per $t = 0$. Si noti che tale fatto vale nell'ipotesi che K sia sottospazio per cui $-x_1 + ty \in K$ per ogni $t \in \mathbb{R}$ se $x_1, y \in K$. Quindi la sua derivata si annulla in $t = 0$:

$$f'(0) = \lim_{t \to 0} \frac{||x_2 + ty||^2 - ||x_2||^2}{t} = (x_2|y) + (y|x_2) = 2Re(x_2|y) \,.$$

Di conseguenza $Re(x_2|y) = 0$. Sostituendo y con iy si ricava che anche la parte immaginaria di $(x_2|y)$ vale zero per cui $(x_2|y) = 0$ e $x_2 \in K^\perp$. Con questo abbiamo provato che $< K, K^\perp >= H$. Non resta che mostrare che $K \cap K^\perp = \{0\}$. Ma questo è ovvio dal fatto che se $x \in K \cap K^\perp$ allora x deve essere ortogonale a x allora $||x||^2 = (x|x) = 0$ e quindi $x = 0$.
(e) Da (b) si ha che, se $y \in \overline{< K >}$, allora y è ortogonale ad ogni elemento di $K^\perp$. In altre parole

$$\overline{< K >} \subset (K^\perp)^\perp \,. \tag{3.7}$$

Usando (d) e sostituendo al posto di K il sottospazio chiuso $\overline{< K >}$ otteniamo che $H = \overline{< K >} \oplus \overline{< K >}^\perp$. Ossia, tenendo conto di (b)

$$H = \overline{< K >} \oplus K^\perp \,. \tag{3.8}$$

Se $u \in (K^\perp)^\perp$, in base alla (3.8) c'è una decomposizione fatta di vettori ortogonali ($(u_0|v) = 0$) $u = u_0 + v$ con $u_0 \in \overline{< K >}$ e $v \in K^\perp$ e quindi $(u|v) = (v|v)$. Valendo $(u|v) = 0$ perchè

$u \in (K^\perp)^\perp$ e $v \in K^\perp$ deve essere $(v|v) = 0$ e dunque $(K^\perp)^\perp \ni u = u_0 \in \overline{< K >}$. Concludiamo che $\overline{< K >} \supset (K^\perp)^\perp$ da cui la tesi valendo (3.7). $\square$

Il teorema provato ha il seguente corollario che segue immediatamente da (b) e (d).

Corollario. *Se S è un sottoinsieme di uno spazio di Hilbert H, $< S >$ è denso in H se e solo se $S^\perp = \{0\}$.*

Possiamo ora enunciare e provare il teorema dovuto a F. Riesz, che è sicuramente il teorema più imporatante della teoria degli spazi di Hilbert.

Teorema 3.2 (Riesz). *Sia $(H, (\ |\ ))$ spazio di Hilbert. Ogni funzionale lineare continuo su H, f, è univocamente rappresentabile come*

$$f(x) = (y_f|x)\,, \quad \textit{per ogni } x \in H\,,$$

dove $y_f \in H$. Al variare di $f \in H'$, y_f assume ogni valore in H.

Prova. È chiaro che ogni applicazione $H \ni x \mapsto (y|x)$ definisce un elemento di H' per la continuità del prodotto scalare. D'altra parte si ottengono due diverse applicazioni usando due y differenti: $(y|x) = (y'|x)$ per ogni $x \in K$ implica $||y - y'||^2 = (y - y'|y - y') = 0$ da cui $y = y$. Quindi l'applicazione $H' \ni f \to y_f \in H$ se esiste è iniettiva.
Per concludere il teorema è quindi sufficiente provare che per ogni $f \in H'$ esiste $y_f \in H$ tale che $f(x) = (y_f|x)$ per ogni $x \in H$. Il nucleo di f, $Kerf$, è un sottospazio chiuso dato che f è continuo. Se $Kerf = H$ allora $y_f = 0$, altrimenti $H = Kerf \oplus (Kerf)^\perp$ per il teorema 3.1. Proviamo che $(Kerf)^\perp$ ha dimensione 1. Sia $0 \neq y \in K^\perp$. In tal caso $f(y) \neq 0$ ($y \notin Kerf$!). Per ogni $z \in (Kerf)^\perp$, il vettore $z - \frac{f(z)}{f(y)}y$ cade in $(Kerf)^\perp$ perché combinazione lineare di elementi di $(Kerf)^\perp$, ma anche $z - \frac{f(z)}{f(y)}y \in Kerf$ come si verifica immediatamente. Quindi $z - \frac{f(z)}{f(y)}y \in Kerf \cap (Kerf)^\perp$ e dunque $z - \frac{f(z)}{f(y)}y = 0$. Quindi y è una base per $(Kerf)^\perp$ dato che ogni altro vettore $z \in (Kerf)^\perp$ è una combinazione lineare di y: $z = \frac{f(z)}{f(y)}y$. Definiamo

$$y_f := \frac{f(y)}{(y|y)}y\,.$$

Mostriamo che y_f rappresenta f nel senso voluto.
Se $x \in H$, possiamo decomporre x rispetto alla decomposizione $Kerf \oplus (Kerf)^\perp$ ottenendo: $x = n + x^\perp$, dove

$$x^\perp = \frac{f(x^\perp)}{f(y)}y = \frac{f(x)}{f(y)}y$$

per quanto provato sopra ed avendo tenuto conto del fatto che $f(x^\perp) = f(x)$ per linearità essendo $f(n) = 0$. Il calcolo diretto di $(y_f|x)$ fornisce immediatamente: $(y_f|x) = f(x)$ concludendo

la dimostrazione. $\square$.

Corollario. *Ogni spazio di Hilbert è riflessivo.*

Prova. Prima di tutto notiamo che possiamo mettere su H' un prodotto scalare definito come $(f|g)' := (y_g|y_f)$, dove $f(x) = (y_f|x)$ e $g(x) = (y_g|x)$ per ogni $x \in H$. Con tale prodotto scalare H' è spazio di Hilbert per il teorema 3.2 e l'applicazione *anti* lineare biettiva $H' \ni f \mapsto y_f \in H$ preserva il prodotto scalare per costruzione. In questo senso H e H' sono *anti isomorfi*. Si osservi che la norma indotta da $(\ |\ )'$ in H' coincide con la norma naturale di H'

$$||f|| := \sup_{||x||=1} |f(x)| \, .$$

Infatti questa definizione si può scrivere, per il teorema 3.2:

$$||f|| = \sup_{||x||=1} |(y_f|x)| \, ;$$

la disuguaglianza di Cauchy-Schwartz implica subito che $||f|| \leq ||y_f||$, inoltre vale $|(y_f|x)| = ||y_f||$ se $x = y_f/||y_f||$. Quindi $||f|| = ||y_f||$, ma quest'ultima è proprio, come volevamo, la norma di f indotta dal prodotto scalare $(\ |\ )'$.
Quindi $(H', (\ |\ )')$ è uno spazio di Hilbert e $(H')'$ il suo duale. Il teorema 3.2 assicura che gli elementi di $(H')'$, F, sono tutti e soli quelli per cui esiste qualche $g_F \in H'$ tale che $F(f) = (g_F|f)'$ per ogni $f \in H'$. Ma $(g_F|f)' = (y_f|y_{g_F}) = f(y_{g_F})$. Abbiamo cioè che, per ogni $F \in (H')'$ esiste (ed è unico per il corollario 3 del teorema di Hahn-Banach) un vettore $y_{g_F} \in H$ tale che, per ogni $f \in H'$:

$$F(f) = f(y_{g_F}) \, .$$

Questa proprietà altro non è che la riflessività di H'. $\square$

3.2 Basi hilbertiane.

Passiamo ora ad introdurre tutto l'armamentario matematico relativo al concetto di *base hilbertiana*. Abbiamo bisogno di qualche risultato preliminare.
Chiameremo **insieme indiciato**, e lo indicheremo con $\{\alpha_i\}_{i\in I}$, ogni applicazione $I \ni i \mapsto \alpha_i$. L'insieme I è detto **insieme degli indici** la coppia (i, α_i), è detto **elemento i-esimo** dell'insieme indiciato. Si osservi che può accadere che $\alpha_i = \alpha_j$ con $i \neq j$.

Definizione 3.4. *Se $A = \{\alpha_i\}_{i\in I}$ è un insieme non vuoto di numeri reali non negativi indiciati sull'insieme I di cardinalità arbitraria, la* **somma dell'insieme indiciato** *A è il numero in $[0, +\infty) \cup \{+\infty\}$ definito come:*

$$\sum_{i\in I} \alpha_i := \sup \left\{ \sum_{j\in F} \alpha_j \;\middle|\; F \subset I \, , \quad F \textit{ finito} \right\} \, . \tag{3.9}$$

Proposizione 3.2. *In riferimento alla definizione 3.4 vale quanto segue.*
(a) *Se I è finito o numerabile, la somma dell'insieme indiciato A coincide con la somma, ovvero rispettivamente con la somma della della serie $\{\alpha_{i_n}\}_{n\in\mathbb{N}}$ (che converge sempre al più a $+\infty$ essendo serie di termini non negativi) indipendentemente dall'ordinamento, cioè dalla funzione bigettiva $\mathbb{N} \ni n \mapsto i_n \in I$.*
(b) *Se la somma dell'insieme A è finita allora il sottoinsieme di I per cui $\alpha_i \neq 0$ è finito o numerabile. In tal caso, restringendosi al sottoinsieme detto, la somma di A coincide con la somma su un insieme finito di indici o, rispettivamente, la somma di una serie come precisato in (a).*
(c) *Se μ è la misura che conta i punti di I definita sulla σ-algebra dell'insieme della parti di I (se $J \subset I$ allora $\mu(J) \leq +\infty$ è il numero di elementi di J per definizione), allora*

$$\sum_{i\in I} \alpha_i = \int_A \alpha_i \, d\mu(i) \, . \tag{3.10}$$

Prova. (a) Il caso di I finito è ovvio, consideriamo il caso di I infinito numerabile e supponiamo di avere scelto un particolare ordinamento di I, per cui possiamo riscrivere A come $A = \{\alpha_{i_n}\}_{n\in\mathbb{N}}$. Mostriamo che la somma della serie di elementi $\{\alpha_{i_n}\}_{n\in\mathbb{N}}$, $\sum_{n=0}^{\infty} \alpha_{i_n}$, coincide con la somma nel senso (3.9) che, per definizione, non dipende dall'ordinamento scelto. Vale, per (3.9):

$$\sum_{n=0}^{N} \alpha_{i_n} \leq \sum_{i\in I} \alpha_i \, .$$

E quindi prendendo il limite per $N \to \infty$, che esiste e coincide con l'estremo superiore dell'insieme delle ridotte della serie, essendo non decrescente la successione delle ridotte:

$$\sum_{n=0}^{\infty} \alpha_{i_n} \leq \sum_{i\in I} \alpha_i \, . \tag{3.11}$$

D'altra parte se $F \subset I$ è finito possiamo sempre fissare N_F sufficientemente grande in modo tale che tale che $\{\alpha_i\}_{i\in F} \subset \{\alpha_{i_0}, \alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_{N_F}}\}$. In tal modo

$$\sum_{i\in F} \alpha_i \leq \sum_{n=0}^{N_F} \alpha_{I_n} \, ,$$

Prendendo l'estremo superiore al variare degli insiemi $F \subset I$ finiti, e ricordando che l'estremo superiore dell'insieme delle ridotte coincide con la somma della serie abbiamo:

$$\sum_{i\in I} \alpha_i \leq \sum_{n=0}^{\infty} \alpha_{i_n} \, . \tag{3.12}$$

(3.11) e (3.12) implicano la tesi.
(b) sia $S < \infty$, $S \geq 0$, la somma dell'insieme indiciato A. Se $S = 0$ tutti gli elementi di A devono essere nulli e la tesi è verificata banalmente. Si supponga che $S > 0$. È chiaro che ogni α_i è contenuto in $[0, S]$ – altrimenti la somma sarebbe maggiore di S – in particolare se $\alpha_i \neq 0$ allora $\alpha_i \in (0, S]$. Poniamo $S_n := S/n$ per $n = 1, 2, \ldots$. Se N_k è il numero di elementi $i \in I$ per cui α_i cade nell'inervallo, $(S_k, S_{k+1}]$ sicuramente $\sum_{i \in I} \alpha_i \geq S_k N_k$ ovvero $\sum_{i \in I} \alpha_i = +\infty$ se N_k è infinito. Quindi N_k deve essere finito per ogni k. Dato che l'unione degli intervalli disgiunti $(S_k, S_{k+1}]$ è $(0, S]$, ogni $\alpha_i \neq 0$ cade esattamente in uno solo di tali intervalli, che contiene un numero finito di $\alpha_i \neq 0$ come provato sopra. La cardinalità dell'insieme degli i per cui $\alpha_i \neq 0$ è quindi al più numerabile perchè tale insieme è unione finita o numerabile di insiemi finiti.
(c) Notando che ogni funzione è misurabile rispetto alla misura detta (dato che la σ-algebra è l'insieme delle parti) l'identità (3.10) è immediata conseguenza della definizione di integrale di una funzione positiva [1]. □

Passiamo a definire in vari passi il concetto di *sistema ortonormale completo* ovvero *base hilbertiana.*

Definizione 3.5. *Sia $(H, (\ |\ ))$ spazio con prodotto scalare hermitano e $\emptyset \neq N \subset H$.*
(a) *N è detto* **insieme ortogonale** *se $N \not\ni \{0\}$ e se $x, y \in N$ sono ortogonali quando $x \neq y$.*
(b) *N è detto* **insieme ortonormale** *se è ortogonale ed ogni suo elemento ha norma (indotta da prodotto scalare) unitaria.*
(c) *Se $(H, (\ |\ ))$ è spazio di Hilbert, N è detto* **sistema ortonormale completo** *o equivalentemente* **base hilbertiana***, se è un insieme ortonormale e soddifa la proprietà:*

$$N^{\perp} = \{0\}\,. \tag{3.13}$$

Nota. È chiaro che un insieme di vettori ortogonali N è *linearmente indipendente*: se $F \subset N$ è finito e $0 = \sum_{x \in F} \alpha_x x$ allora $0 = (\sum_{x \in F} \alpha_x x | \sum_{y \in F} \alpha_y y) = \sum_{x \in F} \sum_{y \in F} \overline{\alpha_x} \alpha_y (x|y) = \sum_{x \in F} |\alpha_x|^2 ||x||^2$. Dato che $||x|| > 0$ e $|\alpha_x|^2 \geq 0$ deve necessariamente essere $|\alpha_x| = 0$ per ogni $x \in L$. Ossia $\alpha_x = 0$ per ogni $x \in L$.

Teorema 3.3 (Disuguaglianza di Bessel). *Sia $(H, (\,|\,))$ spazio con prodotto scalare hermitano e N un insieme ortonormale in H. Per ogni $x \in H$*

$$\sum_{z \in N} |(x|z)|^2 \leq ||x||^2\,. \tag{3.14}$$

In particolare solo una quantità numerabile di elementi $(x|z)$ sono non nulli.

Prova. È sufficiente provare (3.14) per i sottoinsiemi $F \subset N$ finiti. Se (3.14) vale in tali casi, la prova nel caso generale includendo la proposizione enunciata sotto (3.14) è immediata conseguenza della definizione 3.4 e di (b) della proposizione. Dimostriamo dunque la (3.14) per

$F \subset N$ finito contenente n elementi $z_1, \ldots, z_n$. Siano $x \in H$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. Sviluppando la norma $||x - \sum_{k=1}^n \alpha_k z_k||^2$ in termini del prodotto scalare di H e tenendo conto dell'ortonormalità degli elementi z_p e z_q e delle proprietà di (anti) linearità del prodotto scalare, si trova subito che:

$$\left|\left| x - \sum_{k=1}^n \alpha_k z_k \right|\right|^2 = ||x||^2 + \sum_{k=1}^n |\alpha_k|^2 - \sum_{k=1}^n \alpha_k (x|z_k) - \sum_{k=1}^n \overline{\alpha_k (x|z_k)} \, .$$

Il secondo membro può riscriversi

$$||x||^2 - \sum_{k=1}^n |(x|z_k)|^2 + \sum_{k=1}^n \left(|(x|z_k)|^2 - \alpha_k (x|z_k) - \overline{\alpha_k (x|z_k)} + |\alpha_k|^2 \right) \, ,$$

ossia

$$\left|\left| x - \sum_{k=1}^n \alpha_k z_k \right|\right|^2 = ||x||^2 - \sum_{k=1}^n |(x|z_k)|^2 + \sum_{k=1}^n |\, (z_k|x) - \alpha_k \,|^2 \, .$$

È chiaro che il secondo membro ha un unico minimo assoluto per $\alpha_k = (z_k|x)$ $k = 1, \ldots, n$. In tal caso si ottiene proprio:

$$0 \leq \left|\left| x - \sum_{k=1}^n (z_k|x) z_k \right|\right|^2 = ||x||^2 - \sum_{k=1}^n |(x|z_k)|^2 \, ,$$

da cui

$$\sum_{k=1}^n |(x|z_k)|^2 \leq ||x||^2 \, .$$

Ciò completa la dimostrazione $\square$.

Lemma 3.1. *Sia $\{x_n\}_{n \in \mathbb{N}}$ un insieme ortogonale numerabile indicizzato su $\mathbb{N}$ nello spazio di Hilbert $(H, (\ |\ ))$. Sia $||\ ||$ la norma indotta da $(\ |\ )$. Se*

$$\sum_{n=0}^\infty ||x_n||^2 < \infty \, , \tag{3.15}$$

allora:
(a) *esiste, ed è unico, $x \in H$ tale che*

$$\sum_{n=0}^\infty x_n = x \, , \tag{3.16}$$

dove la convergenza della serie è intesa come convergenza della successione delle ridotte nella topologia metrica di $||\ ||$;
(b) *la serie in (3.17) può essere riordinata, ossia, per ogni $f : \mathbb{N} \to \mathbb{N}$ biettiva,*

$$\sum_{n=0}^\infty x_{f(n)} = x \, . \tag{3.17}$$

Prova.
(a) Sia $A_n := \sum_{k=0}^{n} x_k$, dalle proprietà di ortonormalità dei vettori x_k e dalla definizione della norma in termini di prodotto scalare segue subito che, se $m > n$ allora

$$||A_n - A_m||^2 = \sum_{k=m+1}^{n} ||x_k||^2 .$$

Dato che la serie a secondo membro converge per ipotesi,

$$||A_n - A_m||^2 = \sum_{k=m+1}^{n} ||x_k||^2 \leq \sum_{k=m+1}^{\infty} ||x_k||^2 \to 0 \quad \text{per } m \to +\infty$$

e questo implica immediatamente che la successione delle ridotte $\{A_n\}$ sia di Cauchy. Dato che H è completo, esisterà $x \in H$ limite di tali msuccessioni e quindi, per definizione, limite della serie. Dato che H è spazio normato e quindi metrico, vale la proprietà di Hausdorff che implica l'unicità dei limiti e quindi in particolare quella di x.
(b) Si fissi $f : \mathbb{N} \to \mathbb{N}$ biettiva. Definiamo i vettori $B_n := \sum_{k=0}^{n} x_{f(k)}$, e i numeri complessi $A_n^* := \sum_{k=0}^{n} ||x_k||^2$ e $S^* := \sum_{k=0}^{\infty} ||x_k||^2$. Per dimostrare la tesi è sufficiente provare che $B_n \to x$ per $n \to +\infty$. Proviamolo. $A_n \to x$ e $A_n^* \to S^*$ (in modo monotono crescente) per $n \to +\infty$, quindi, se $\epsilon > 0$, esiste $N > 0$ tale che $||A_N - x||^2 < \sqrt{\epsilon}/2$ e $S^* - A_N^* < \epsilon/4$. Per tale N sia $M := \max\{f^{-1}(1), f^{-1}(2), \ldots, f^{-1}(n)\}$ allora $\{1, 2, \ldots, N\} \subset \{f(1), f(2), \ldots, f(M)\}$ per la surgettività di f. Per $n \geq M$ vale allora

$$||B_n - x||^2 = (||B_n - A_N|| + ||A_N - x||)^2 \leq (||B_n - A_N|| + \sqrt{\epsilon}/2)^2 .$$

Inoltre

$$||B_n - A_N||^2 = \left|\left| \sum_{k=0}^{n} x_{f(k)} - \sum_{k=0}^{N} x_k \right|\right|^2 .$$

I termini $x_1, \ldots, x_N$ appaiono in entrambe le somme e si cancellano reciprocamente rimanendo:

$$||B_n - A_N||^2 = \left|\left| \sum_{k=N+1}^{n} x_k \right|\right|^2 = \sum_{k=N+1}^{n} ||x_k||^2 \leq S^* - A_N^* \leq \epsilon/4 .$$

In definitiva $||B_n - A_N|| \leq \sqrt{\epsilon}/2$ e quindi $||B_n - x|| < \epsilon$ se $n > M$. Abbiamo ottenuto che $B_n \to x$ per $n \to +\infty$ concludendo la dimostrazione. □

Possiamo quindi enunciare e dimostrare il teorema fondamentale delle basi hilbertiane. Tale teorema prova che le basi hilbertiane soddifano proprietà che sono dirette generalizzazione delle basi ortonormali negli spazi vettoriali *di dimensione finita* dotati di prodotto scalare. La differenza è che ora sono ammesse, usando la topologia di H, anche *combinazioni lineari infinite*:

ogni elemento dello spazio di Hilbert può essere scritto in modo univoco come una combinazione lineare infinita di elementi di una base hilbertiana.
Si deve osservare che, indipendentemente dal fatto che negli spazi di Hilbert esistano tali basi hilbertiane, esistono comunque, come conseguenza del lemma di Zorn (o assioma della scelta) anche *basi algebriche* che non richiedono alcuna nozione topologica. La differenza tra le basi hilbertiane e quella algebriche è quindi che le seconde riguardano solo *combinazioni lineari finite*: ogni elemento dello spazio vettoriale (in questo caso di Hilbert) può essere scritto in modo univoco come una combinazione lineare finita degli elementi della base.

Teorema 3.4. *Sia $(H, (\ |\ ))$ spazio di Hilbert e $N \subset H$ un insieme ortonormale. I seguenti fatti sono equivalenti.*
(a) *N è una base hilbertiana (ossia N è insieme ortonormale e $N^{\perp} = \{0\}$).*
(b) *Fissato $x \in H$, una quantità al più numerabile di numeri $(z|x)$ è non nulla per $z \in N$ e vale:*

$$x = \sum_{z \in N} (z|x) z \ , \tag{3.18}$$

dove la convergenza della serie è intesa come convergenza della successione delle ridotte nella topologia indotta dal prodotto scalare.
(c) *Fissati $x, y \in H$ una quantità al più numerabile di numeri $(z|x), (y|z)$ è non nulla se $z \in N$ e vale:*

$$(x|y) = \sum_{z \in N} (x|z)(z|y) \ . \tag{3.19}$$

(d) *Se $x \in H$ vale, nel senso della definizione 3.4:*

$$||x||^2 = \sum_{z \in N} |(z|x)|^2 \ . \tag{3.20}$$

(e) *$\overline{< N >} = H$, cioè il sottospazio generato da N è denso in H.*
Se vale una delle proprietà equivalenti dette sopra, in (3.18), (3.19) non conta l'ordine con cui si etichettano i coefficienti $(x|z)$, $(z|x) = \overline{(x|z)}$ e $(z|y)$ non nulli.

Prova.
(a) $\Rightarrow$ (b). Per il teorema 3.3 solo una quantità al più numerabile di coefficienti $(z|x)$ è non nullo. Indichiamo con $(z_n|x)$ tali coefficienti, $n \in N$, e poniamo $S_N := \sum_{n=0}^{N} (z_n|x) z_n$. L'insieme $\{(z_n|x) z_n\}_{n \in \mathbb{N}}$ è per costruzione ortogonale, inoltre essendo $||z_n|| = 1$ la disuguaglianza di Bessel implica che $\sum_n ||(z_n|x) z_n||^2 < \infty$. Usando (a) del lemma 1.3 segue che la serie (3.18) converge ad un unico $x' \in H$ con $x' = \sum_{n=0}^{\infty} (z_n|x) z_n$. Inoltre la serie può essere riordinata pur convergendo allo stesso x' per (b) di lemma 1.3. Mostriamo che $x' = x$. Dalla linearità e continuità del prodotto scalare si ha che, se $z_0 \in N$:

$$(x - x'|z_0) = (x|z_0) - \sum_{z \in N} (z|x)(z|z_0) = (x|z_0) - (x|z_0) = 0$$

dove si è tenuto conto del fatto che l'insieme dei coefficienti z costituisce un insieme ortonormale. Dato che $z_0 \in N$ è arbitrario, $x - x' \in N^\perp$ e quindi $x - x' = 0$ in quanto $N^\perp = \{0\}$ per ipotesi. Questo prova che (3.18) vale indipendentemente dall'ordine con cui si etichettano i coefficienti $(z|x) \neq 0$.
(b) $\Rightarrow$ (c). Se (b) è valida, (c) è una sua ovvia conseguenza dovuta alla continuità e (anti) linearità del prodotto scalare e al fatto che N è ortonormale.
(c) $\Rightarrow$ (d). (d) è un caso particolare di (c) quando $y = x$.
(d) $\Rightarrow$ (a). Se (d) è vera e se $x \in H$ è tale che $(x|z) = 0$ per ogni $z \in N$, allora $||x|| = 0$ ossia $x = 0$. In altre parole $N^\perp = \{0\}$, cioè vale (a).
Abbiamo provato che (a),(b),(c) e (d) sono equivalenti. Per concludere notiamo che (b) implica immediatamente (e) mentra (e) implica (a): se $x \in N^\perp$ allora, per la linearità e continuità del prodotto scalare $x \in \overline{< N >^\perp}$. Ma per (b) del teorema 1.3 $\overline{< N >^\perp} = \overline{< N >}^\perp$. Essendo $\overline{< N >} = H$ per ipotesi, $x \in H^\perp = \{0\}$. In altre parole essendo $N^\perp = \{0\}$ e quindi vale (a).
Il fatto che la serie di numeri complessi in (3.19) può essere riordinata arbitrariamente conservandone la somma segue dal fatto che essa è assolutamente convergente, essendo per la disuguaglianza di Cauchy-Schwartz (fissato un arbitrario ordinamento)

$$\sum_{z \in N} |(x|z)||(z|y)| \leq (\sum_{z \in N} |(x|z)|^2)^{1/2} (\sum_{z \in N} |(x|z)|^2)^{1/2} < \infty$$

per (d). □

Il lemma di Zorn implica immediatamente che ogni spazio di Hilbert ammette una base hilbertiana.

Teorema 3.5. *Ogni spazio di Hilbert ammette una base hilbertiana.*

Prova. Sia H uno spazio di Hilbert. Consideriamo la classe $\mathcal{A}$ i cui elementi sono gli insiemi ortonormali in H. Definiamo in $\mathcal{A}$ la relazione d'ordine parziale data dall'inclusione insiemistica. Per costruzione ogni sottoinsieme ordinato $\mathcal{E}$ di $\mathcal{A}$ è superiormente limitato dall'elemento costituito dall'unione degli elementi di $\mathcal{E}$. Per il lemma di Zorn esite allora in $\mathcal{A}$ un elemento massimale N. Con le definizione date, la massimalità di N implica che non ci sono in H vettori normali a tutti gli elementi di N, non nulli e non appartenenti a N stesso. In altre parole N è un sistema ortonormale completo.□

Prima di passare al caso degli spazi di Hilbert separabili, diamo ancora un importante risultato della teoria generale.

Teorema 3.6. *Sia H uno spazio di Hilbert con base hilbertiana N. Valgono i due seguenti fatti.*
(a) *H è isomorfo (come spazio con prodotto scalare hermitiano) allo spazio di Hilbert $L^2(N, \mu)$ dove μ è la misura positiva che conta i punti di N (vedi (6) e (7) in esempi 2.1 e (2) in esempi*

3.1); l'isomorfismo è dato dall'applicazione

$$H \ni x \mapsto \{(z|x)\}_{z\in N} \in L^2(N,\mu)\,. \tag{3.21}$$

(b) *Tutte le basi hilbertiane di H hanno la stessa cardinalità (pari a quella di N) che si dice* **dimensione** *dello spazio di Hilbert.*
(c) *Se H_1 è spazio di Hilbert con la stessa dimensione di H allora i due spazi sono isomorfi (come spazi con prodotto scalare hermitiano).*

Prova. (a) L'applicazione $U : H \ni x \mapsto \{(z|x)\}_{z\in N} \in L^2(N,\mu)$ è ben definita, in quanto, se $x \in H$ e N è una base hilbertiana allora vale la propietà (d) del teorema 3.4 che afferma proprio che $\{(z|x)\}_{z\in N} \in L^2(N,\mu)$. Tale applicazione è sicuramente iniettiva: se $x, x' \in H$ producono gli stessi coefficienti $(z|x) = (z|x')$ per ogni $z \in N$, allora $x = x'$ per (b) del teorema 3.4. L'applicazione è anche suriettiva: se $\{\alpha_z\}_{z\in N} \in L^2(N,\mu)$ e quindi $\sum_{z\in N} |\alpha_z|^2 < \infty$, allora per il lemma 3.1 esiste $x := \sum_{z\in N} \alpha_z z$ e $(z|x) = \alpha_z$ per la continuità del prodotto scalare e l'ortonormalità dell'insieme N. Infine (c) del teorema 3.4 implica che U sia isometrica. Quindi $U : H \to L^2(N,\mu)$ è un isomorfismo di spazi con prodotto scalare e quindi H e $L^2(N,\mu)$ sono isomorfi come spazi con prodotto scalare hermitiano.
(b) Se una delle due basi hilbertiane ha cardinalità finita c allora, banalmente è anche base algebrica di H. Con la geometria elementare si prova che se esiste una base con cardinalità finita c allora ogni altro insieme di vettori linearmente indipendenti M ha cardinalità $\leq c$ e $= c$ se e solo se M genera finitamente tutto lo spazio. Tenendo conto che ogni base hilbertiana, in quanto insieme ortogonale, è costituita da vettori linearmente indipendenti, si conclude facilmente che ogni base hilbertiana di H deve avere cardinalità sicuramente $\leq c$ e quindi $= c$ perchè genera (finitamente) H. Quanto detto esclude anche il caso in cui, in H, una base hilbertiana abbia cardinalità finita ed un'altra abbia cardinalità infinita. Siano dunque N e M basi hilbertiane di H entrambe di cardinalità infinita. Se $x \in M$ definiamo $N_x := \{z \in N \mid (x|z) \neq 0\}$. Dato che $1 = (x|x) = \sum_{z\in N} |(x|z)|^2$ deve accadere che per ogni $z \in N$ esiste $x \in M$ tale che $z \in N_x$. Quindi $N \subset \cup_{x\in M} N_x$ e allora la cardinalità di N sarà inferiore o uguale a quella di $\cup_{x\in M} N_x$ che coincide con quella di M dato che ogni insieme N_x è al più numearbile per (b) del teorema 3.4. Quindi la cardinalità di N è inferiore o uguale a quella di M. Scambiando i ruoli di N ed M si ottiene anche che la cardinalità di M è inferiore o uguale a quella di N. Il teorema di Schröder-Bernstein implica infine che la cardinalità di M e quella di N coincidano.
(c) Siano N e N_1 due basi hilbertiane di H e H_1 rispettivamente, e si supponga che N e N_1 abbiano la stessa cardinalità. Allora esiste un'applicazione biettiva che identifica i punti di N con quelli di N_1 e tale applicazione genera naturalmente un isomorfismo di spazi con prodotto scalare V dallo spazio L^2 su N allo spazio L^2 su N_1 rispetto alla misura che conta i punti. Se $U_1 : H_1 \to L^2(N_1,\mu)$ è l'isomorfismo analogo a $U : H \to L^2(N,\mu)$ descritto sopra, $UVU_1^{-1} : H_1 \to H$ è un isomorfismo di spazi con prodotto scalare per costruzione e dunque H e H_1 sono isomorfi come spazi con prodotto scalare hermitiano.□

Di particolare interesse in fisica sono i cosiddetti spazi di Hilbert *separabili.*

Definizione 3.6. *Uno spazio di Hilbert si dice separabile se ammette un insieme denso e numerabile.*

Vale un ben noto e utile teorema di caratterizzazione.

Teorema 3.7. *Uno spazio di Hilbert è separabile se e solo se ammette una base hilbertiana numerabile. Ulteriormente, se uno spazio di Hilbert è separabile, ogni sua base hilbertiana è numerabile.*

Prova. Se lo spazio di Hilbert ammette una base hilbertiana numerabile, allora, in base a (b) del teorema 3.4, esiste un insieme denso numerabile: quello costituito dalle combinazioni lineari finite di elementi della base usando come coefficienti numeri complessi di parte reale ed immaginaria razionale. Lasciamo i dettagli della facile dimostrazione al lettore. Supponiamo viceversa che uno spazio di Hilbert sia separabile. Sappiamo che, per il teorema 3.5 esistono sempre basi hilbertiane in tale spazio, mostriamo che ciascuna di esse devone essere numerabile. Tale fatto completa la dimostrazione del teorema. (In realtà basterebbe provare che esiste una base hilbertiana numerabile, l'ultima proposizione del teorema sarebbe infatti a questo punto conseguenza di (b) del teorema 3.6. Tuttavia diamo un adimostrazione indipendente di ciò). Supponiamo per assurdo che N sia base hilbertiana non numerabile per lo spazio di Hilbert H separabile. Scelti $z, z' \in N$ con $z \neq z'$, ogni punto $x \in H$ è tale che, per la disuguaglianza triangolare riferita alla norma indotta dal prodotto scalare: $||z - z'|| \leq ||x - z'|| + ||z - x||$. D'altra parte essendo $\{z, z'\}$ un insieme ortonormale $||z - z'||^2 = (z - z'|z - z') = ||z||^2 + ||z'||^2 + 0 = 1 + 1 = 2$. Quindi $||x - z|| + ||x - z'|| \geq \sqrt{2}$. Questo risultato implica immediatamente che palle aperte di raggio $\epsilon < \sqrt{2}/2$ centrate rispettivamente in z e z' sono disgiunte comunque scegliamo $z, z' \in N$. Sia $B(z)$ con $z \in N$ una classe di tale sfere disgiunte a due a due e ciscuna centrata sul rispettivo $z \in N$. Se $D \subset H$ è l'insieme denso e numerabile che esiste per l'ipotesi di separabilità dovrà accadere che, per ogni $z \in N$ esiste $x \in D$ con $x \in B(z)$. Dato che le sfere sono disgiunte avremo un diverso x per ogni sfera. Ma la cardinalità delle sfere non è numerabile per cui nemmeno D può essere numerabile e questo è impossibile. $\square$

Un'altra utile proposizione per spazi di Hilbert separabili è la seguente.

Proposizione 3.4. *Sia $(H, (\ |\ ))$ spazio di Hilbert e $Y := \{y_n\}_{n\in\mathbb{N}}$ un insieme di vettori linearmente indipendenti e tali che $Y^\perp := \{0\}$ o equivalentemente $\overline{< Y >} = H$. Allora H è separabile ed esiste una base hilbertiana di H, $X := \{x_n\}_{n\in\mathbb{N}}$, tale che, per ogni $p \in \mathbb{N}$, il sottospazio spazio generato dai vettori $y_0, y_1, \ldots, y_p$ coincide con quello generato dai vettori $x_0, x_1, \ldots, x_p$.*

Schema di dimostrazione. Diamo solo una traccia della dimostrazione perchè si tratta essenzialmente della nota procedura di ortogonalizzazione di Gram-Schmidt nota dai corsi di geometria elementare.
$y_0 \neq 0$ per ipotesi. Poniamo quindi $x_0 := y_0/||y_0||$. Successivamente consideriamo il vettore non nullo (perchè y_0 e y_1 sono linearmente indipendenti) $z_1 := y_1 - (x_0|y_1)x_0$. È chiaro che x_0, z_1

sono non nulli, ortogonali (quindi linearmente indipendenti) e generano la stesso spazio di y_0 e y_1. Non resta che porre $x_1 := z_1/||z_1||$ per ottenere l'insieme ortonormale $\{x_0, x_1\}$ che genera lo stesso spazio di y_0, y_1. Possiamo ora ripetere la procedura definendo induttivamente l'insieme di vettori:

$$z_n := y_n - \sum_{k=0}^{n-1} (x_k|y_n)x_k \ ,$$

e quindi quello dei vettori $x_n := z_n/||z_n||$. Per induzione si verifica facilmente che i vettori $z_0, \ldots, z_k$ sono non nulli, ortogonali (quindi linearmente indipendenti) e generano lo stesso spazio generato dai vettori linearmente indipendenti $y_0, \ldots, y_k$, di conseguenza i vettori $x_0, \ldots, x_k$ formano un insieme ortonormale che genera lo stesso spazio generato dai vettori linearmente indipendenti $y_0, \ldots, y_k$. Se $u \perp y_n$ per ogni $n \in \mathbb{N}$, allora vale anche (scrivendo gli x_n come combinazione lineare di $y_0, \ldots, y_n$) $u \perp x_n$ per ogni $n \in \mathbb{N}$ e viceversa (scrivendo ogni y_n come combinazione lineare di $x_0, \ldots, x_n$). Quindi $X^\perp = Y^\perp = \{0\}$ e dunque X è una base hilbertiana per H. $\square$

Esempi 3.2.
(1) Consideriamo lo spazio di Hilbert $L^2([-L/2, L/2], dx)$ (vedi (2) in esempi 3.1) dove dx è la solita misura di Lebesgue sui boreliani di $\mathbb{R}$ e $l > 0$. Consideriamo le funzioni misurabili (perchè continue) definite da, per $n \in \mathbb{Z}$ e $x \in [-L/2, L/2]$

$$f_n(x) := \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \ . \tag{3.22}$$

Si verifica immediatamente che le funzioni f_n appartengono allo spazio considerato e costituiscono un insieme ortonormale rispetto al prodotto scalare di $L^2([-L/2, L/2], dx)$ (vedi (2) in esempi 3.1)

$$(f|g) := \int_{-L/2}^{L/2} \overline{f(x)} g(x) dx \ . \tag{3.23}$$

Consideriamo l'algebra di Banch di funzioni continue $C([-L/2, L/2])$ (che è sottospazio di $L^2([-L/2, L/2], dx)$) con norma data dalla norma dell'estremo superiore ((4) e (5) in esempi 2.1). La sottoalgebra S ottenuta prendendo combinazioni lineari finite di prodotti di elementi f_0, f_1, f_{-1} coincide con il sottospazio vettoriale S di $C([-L/2, L/2])$ generato da tutti i vettori f_n, $n \in \mathbb{Z}$. S contiene 1, è chiusa rispetto alla coniugazione complessa e come è facile provare, separa i punti di $[-L/2, L/2]$ (l'insieme delle funzioni f_n da solo separa i punti) per cui, per il teorema di Stone-Weierstrass ((5) in esempi 2.1) è densa in $C([-L/2, L/2])$. D'altra parte è noto che le funzioni continue di $[-L/2, L/2]$ costituiscono uno spazio vettoriale denso in $L^2([-L/2, L/2], dx)$ nella topologia di quest'ultimo ([1] p.85) infine la topologia di $C([-L/2, L/2])$ è più forte di quella di $L^2([-L/2, L/2], dx)$ valendo $(f|f) \leq L \sup |f|^2 = L(\sup |f|)^2$ se $f \in C([-L/2, L/2])$. Ne consegue che S è denso in $L^2([-L/2, L/2], dx)$. Per (e) del teorema 3.4, i vettori f_n costituiscono una base hilbertiana di $L^2([-L/2, L/2], dx)$, che di conseguenza risulta anche essere separabile.
(2) Consideriamo lo spazio di Hilbert $L^2([0, 1], dx)$ dove dx è come nell'esempio precedente e

l'algebra di Banach $C([0,1])$ che è ancora sottospazio denso in $L^2([0,1],dx)$ nella topologia di quest'ultimo. Definiamo ora diversamente dall'esempio precedente, per $n = 0,1,2,\ldots$, $x \in [0,1]$:

$$g_n(x) := x^n \,. \tag{3.24}$$

Si dimostra subito che i vettori scritti sopra sono linearmente indipendenti. La sottoalgebra S ottenuta prendendo combinazioni lineari finite di prodotti di elementi g_0, g_1 coincide con il sottospazio vettoriale S di $C([0,1])$ generato da tutti i vettori f_n, $n \in \mathbb{N}$. S contiene 1, è chiusa rispetto alla coniugazione complessa e separa i punti per cui, per il teorema di Stone-Weierstrass è densa in $C([0,1])$ e anche in $L^2([0,1],dx)$ ragionando come nell'esempio precedente. La differenza dal caso precedente è ora che lefunzioni g_n non costituiscono un insieme ortonormale. Tuttavia, usando la proposizione 3.4 si costruisce immediatamente un sistema ortonormale completo per $L^2([0,1],dx)$. Gli elementi di tale base hilbertiana si chiamano i *polinomi di Legendere.*
(3) I due esempi precedenti esibiscono due spazi L^2 separabili. Si può provare che $L^p(X,\mu)$ $(1 \leq p < \infty)$ è separabile se e solo se μ è separabile, cioè ammette un sottoinsieme denso e numerabile lo spazio metrico costruito sul sottoinsieme della σ-algebra Σ della misura μ contenente insiemi di misura finita, prendendo il quoziente rispetto alla relazione di equivalenza che identifica insiemi che differeiscono per insiemi di misura nulla, e infine definendo la distanza come:

$$d(A,B) := \mu((A \setminus B) \cup (B \setminus A)) \,.$$

A tal fine vale la seguente proposizione:
μ (e quindi $L^p(X,\mu)$) è separabile se μ è σ-finita (cioè X è unione numerabile di insiemi di misura finita) ed esiste una classe al più numerabile di insiemi misurabili la cui σ-algebra generata coincide con quella della misura μ. [13]
Si osservi che di conseguenza, ogni misura di Borel σ-finita riferita ad uno spazio topologico a base numerabile produce spazi L^p separabili. In particolare questo accade per la misura di Lebesgue su $\mathbb{R}^n$. Le misure di Borel positive definite su spazi di Hausdorff localmente compatti si chiamano *misure di Radon* se sono regolari e i compatti hanno misura finita. Una misura di Radon è σ-finita se lo spazio su cui è definita è σ-compatto, cioè unione numerabile di compatti.
(4) Consideriamo lo spazio $L^2((a,b),dx)$ dove $-\infty \leq a < b \leq +\infty$ e dx denota l'usuale misura di Lebesgue du $\mathbb{R}$. Sussiste la seguente utilissima proposizione la cui dimostrazione verrà data negli esempi 3.3 dopo avere introdotto la trasformata di Fourier e Fourier-Plancherel:
Sia $f : (a,b) \to \mathbb{C}$ misurabile tale che (1) l'insieme $\{x \in (a,b) \mid f(x) = 0\}$ ha misura nulla, (2) esistono $C, \delta > 0$ per cui $|f(x)| < Ce^{-\delta|x|}$ per ogni $x \in (a,b)$.
In questo caso lo spazio lineare finitamente generato da tutte le funzioni $x \mapsto x^n f(x)$ per $n = 0,1,2,\ldots$ è denso in $L^2((a,b),dx)$.
L'importanza di questo risultato è che consente di costruire facilmente delle basi hilbertiane in $L^2((a,b),dx)$ anche se a o b sono infiniti (per cui non è possibile usare Stone-Weierstrass). Infatti, la procedura di ortogonalizzazione di Gram-Schmidt applicata ai vettori f_n con $f_n(x) := x^n f(x)$ fornisce una base hilbertiana secondo la procedura illustrata in proposizione 3.4.
A titolo di esempio, se $f(x) := e^{-x^2/2}$, la procedura di Gram-Schmidt definisce la base hilbertiana di $L^2(\mathbb{R},dx)$ delle cosiddette *Funzioni di Hermite*, che hanno tutte la forma $g_n(x) := H_n(x)e^{-x^2/2}$

dove H_n è un polinomio di grado $n = 0, 1, 2, \ldots$ chiamato n-esimo *polinomio di Hermite.* Nella Meccanica Quantistica questa base hilbertiana è importante quando si studia il sistema fisico detto *oscillatore armonico unidimensionale.*
Con la stessa procedura ed usando $f(x) := e^{-x}$ si trova una base hilbertiana di $L^2((0, +\infty), dx)$ data dalle cosiddette *funzioni di Laguerre* che sono della forma $e^{-x}L_n(x)$ con $n = 0, 1, \ldots$. Ogni polinomio L_n è di grado n è si dice n-esimo *polinomio di Laguerre.* In Meccanica Quantistica questa base hilbertiana è importante quando si lavora con sistemi fisici con simmetria sferica come l'atomo di idrogeno.

3.3 Aggiunto di un operatore limitato, operatori autoaggiunti, normali, unitari, positivi e proiettori ortogonali, C^*-algebre di operatori limitati.

Esaminamo ora una delle più importanti nozioni della teoria degli operatori in uno spazio di Hilbert: la nozione di *operatore aggiunto (hermitiano).* In questa sezione considereremo solo il caso di operatori limitati. Il caso generale sarà trattato più avanti dopo avere dato la formalizzazione di base della Meccanica Quantistica.

Sia $(H, (\ |\ ))$ uno spazio di Hilbert e $T \in \mathfrak{B}(H)$. Consideriamo, per $u \in H$ fissato, l'applicazione

$$H \ni v \mapsto (u|Tv) \in \mathbb{C}\,. \tag{3.25}$$

L'applicazione di sopra è sicuramente lineare ed anche limitata: $|(u|Tv)| \leq ||u||||T||||v||$. Si tratta dunque di un elemento di H'. per il teorema di Riesz (teorema 3.2), esisterà $w_{T,u} \in H$ tale che

$$(u|Tv) = (w_{T,u}|v)\,, \quad \text{per ogni } v \in H.\,. \tag{3.26}$$

Possiamo ancora notare che l'applicazione $H \ni u \mapsto w_{T,u}$ è lineare. Infatti:

$$(w_{T,\alpha u+\beta u'}|v) = (\alpha u + \beta u'|Tv) = \overline{\alpha}(u|Tv) + \overline{\beta}(u'|Tv) = (\alpha w_{T,u} + \beta w_{T,u'}|v)\,,$$

per cui, per ogni $v \in H$,

$$(w_{T,\alpha u+\beta u'} - \alpha w_{T,u} + \beta w_{T,u'}|v)\,;$$

scegliendo $v := w_{T,\alpha u+\beta u'} - \alpha w_{T,u} + \beta w_{T,u'}$ si ha che deve essere $w_{T,\alpha u+\beta u'} - \alpha w_{T,u} + \beta w_{T,u'} = 0$ e quindi

$$w_{T,\alpha u+\beta u'} = \alpha w_{T,u} + \beta w_{T,u'}$$

per ogni $\alpha, \beta \in C$ e $u, u' \in H$. Quindi esiste un operatore lineare $T^\dagger : H \ni u \mapsto w_{T,u}$. Questo operatore soddisfa $(u|Tv) = (T^\dagger u|v)$ per ogni coppia $u, v \in H$ ed è l'unico operatore lineare a soddisfare questa identità. Se ci fosse un altro operatore $B \in \mathfrak{L}(H)$ che soddisfa l'identità detta, varrebbe anche $(T^\dagger u|v) = (Bu|v)$ per ogni $v \in H$. Di conseguenza avremmo

che $((T^\dagger - B)u|v) = 0$ per ogni $v \in H$. Scegliendo $v := (T^\dagger - B)u$ si ha che deve essere $||(T^\dagger - B)u||^2 = 0$ e quindi $T^\dagger u - Bu = 0$. Dato che $u \in H$ è arbitrario, deve infine valere $T^\dagger = B$. Abbiamo in definitiva provato che vale la proposizione seguente.

Proposizione 3.5. *Sia $(H, (\ |\ ))$ uno spazio di Hilbert e $T \in \mathfrak{B}(H)$.*
Esiste ed è unico un operatore lineare $T^\dagger : H \to H$ tale che:

$$(u|Tv) = (T^\dagger u|v)\,, \quad \textit{per ogni coppia } u, v \in H. \tag{3.27}$$

Possiamo dare la definizione di *operatore aggiunto hermitiano*. Nel seguito ometteremo l'aggettivo "hermitiano", dato che in queste dispense non lavoreremo mai con operatori aggiunti non hermitiani (si può dare un'analoga definizione di operatore aggiunto in spazi di Banach.)

Definizione 3.7. *Sia $(H, (\ |\ ))$ uno spazio di Hilbert e $T \in \mathfrak{B}(H)$. L'unico operatore lineare $T^\dagger \in \mathfrak{L}(H)$ che soddisfa (3.27) si dice* **aggiunto (hermitiano)** *o anche* **coniugato hermitiano** *dell'operatore T.*

L"operazione di coniugazione hermitiana gode delle seguenti proprietà elementari.

Proposizione 3.6. *Sia $(H, (\ |\ ))$ uno spazio di Hilbert e $T \in \mathfrak{B}(H)$. Valgono i seguenti fatti.*
(a) *$T^\dagger \in \mathfrak{B}(H)$ e più precisamente*

$$||T^\dagger|| = ||T||\,, \tag{3.28}$$
$$||T^\dagger T|| = ||T||^2 = ||TT^\dagger||\,. \tag{3.29}$$

(b) *L'operazione di coniugazione hermitiana è* **involutiva***:*

$$(T^\dagger)^\dagger = T\,.$$

(c) *Se $S \in \mathfrak{B}(H)$ e $\alpha, \beta \in \mathbb{C}$ valgono le identità*

$$(\alpha T + \beta S)^\dagger = \overline{\alpha} T^\dagger + \overline{\beta} S^\dagger\,, \tag{3.30}$$
$$(TS)^\dagger = S^\dagger T^\dagger\,. \tag{3.31}$$

(d) *Se $Ran(A)$ e $Ker(A)$ indicano rispettivamente il rango (o immagine) dell'operatore A e il nucleo dell'operatore A, allora*

$$Ker(T) = Ran(T^\dagger)^\perp\,, \quad Ker(T^\dagger) = Ran(T)^\perp\,. \tag{3.32}$$

Prova. (a) Per ogni coppia $u, x \in H$ vale $|(T^\dagger u|x)| = |(u|Tx)| \leq ||u||\,||T||\,||x||$. Scegliendo $x := T^\dagger u$ si ha in particolare $||T^\dagger u||^2 \leq ||T||\,||u||\,||T^\dagger u||$ e quindi $||T^\dagger u|| \leq ||T||\,||u||$. Quindi $T^\dagger$ è limitato e $||T^\dagger|| \leq ||T||$. Ha senso quindi definiere $(T^\dagger)^\dagger$ ed ottenere $||(T^\dagger)^\dagger|| \leq ||T^\dagger||$. Questa

disuguaglianza si può riscrivere $||T|| \leq ||T^\dagger||$ per (b), la cui dimostrazione usa solo il fatto che $T^\dagger$ sia limitato. Valendo $||T^\dagger|| \leq ||T||$ e $||T|| \leq ||T^\dagger||$ vale (3.28). Passiamo a dimostrare (3.29). È sufficiente dimostrare la prima delle due identità, la seconda segue dalla prima e da (3.28). Notiamo che, per (b)(i) del teorema 2.3, $||T^\dagger T|| \leq ||T^\dagger||\,||T|| = ||T||^2$. D'altra parte:

$$||T||^2 = (\sup_{||x||\leq 1} ||Tx||)^2 = \sup_{||x||\leq 1} ||Tx||^2 = \sup_{||x||\leq 1} (Tx|Tx)\,.$$

Usando la definizione di aggiunto e poi la disuguaglianza di Cauchy-Schwartz sull'ultimo termine della catena di uguaglianze scritte sopra, troviamo che :

$$||T||^2 = \sup_{||x||\leq 1} (Tx|Tx) = \sup_{||x||\leq 1} |(T^\dagger Tx|x)| \leq \sup_{||x||\leq 1} ||T^\dagger Tx|| = ||T^\dagger T||\,.$$

In definitiva abbiamo ottenuto che $||T^\dagger T|| \leq ||T||^2$ e $||T||^2 \leq ||T^\dagger T||$ per cui vale $||T^\dagger T|| = ||T||^2$.
(b) Segue immediatamente dall'unicità dell'operatore aggiunto valendo, per le note proprietà del prodotto scalare e usando la stessa definizione di operatore aggiunto dell'operatore T:

$$(v|T^\dagger u) = \overline{(T^\dagger u|v)} = \overline{(u|Tv)} = (Tv|u)\,.$$

(c) Se $u, v \in H$

$$(u|(\alpha T + \beta S)v) = \alpha(u|Tv) + \beta(u|Sv) = \alpha(T^\dagger u|v) + \beta(S^\dagger u|v) = ((\overline{\alpha}T^\dagger + \overline{\beta}S^\dagger)u|v)\,.$$

Per l'unicità dell'operatore aggiunto vale (3.30). Se $u, v \in H$

$$(u|(TS)v) = (T^\dagger u|Sv) = ((S^\dagger T^\dagger)u|v)\,.$$

Per l'unicità dell'operatore aggiunto vale (3.31).
(d) È sufficiente provare la prima delle due identità, la seconda segue subito dalla prima usando (b). Dato che $(T^\dagger u|v) = (u|Tv)$, se $v \in Ker(T)$ allora $(T^\dagger u|v) = 0$ per ogni $u \in H$ e quindi $v \in Ran(T^\dagger)^\perp$. Viceversa, sempre per $(T^\dagger u|v) = (u|Tv)$, se $v \in Ran(T^\dagger)^\perp$, allora $(u|Tv) = 0$ per ogni $u \in H$. Scegliendo $u := Tv$ segue che $Tv = 0$ e quindi $v \in Ker(T)$. □

L'operazione di coniugazione hermitiana consente di introdurre uno dei più importanti concetti matematici utili nelle formulazioni avanzate della Meccanica Quantistica, stiamo parlando della la struttura di C^*-*algebra* (anche detta B^*-*algebra*).

Definizione 3.8 *Sia A un'algebra (commutativa, con unità, normata con norma $||\,||$, di Banach) sul campo $\mathbb{C}$. Se esiste un'applicazione $^* : A \to A$ che gode delle seguenti propietà:*
I1. *(antilinearità) $(\alpha x + \beta y)^* = \overline{\alpha}x^* + \overline{\beta}y^*$ per ogni $x, y \in A$, $\alpha, \beta \in \mathbb{C}$,*
I2. *(involutività) $(x^*)^* = x$ per ogni $x \in A$,*
I3. *$(xy)^* = y^*x^*$ per ogni $x, y \in A$,*
tale applicazione è detta **involuzione** *e la struttura $(A,^*)$ si dice *-algebra (rispettivamente commutativa, con unità, normata, di Banach).*

*Una *-algebra di Banach (con unità) si dice C^*-**algebra** (rispettivamente con unità) se vale la ulteriore proprietà*

$$||x^*x|| = ||x||^2 \,. \tag{3.33}$$

Un elemento x di una *-algebra è detto **normale** *o* **hermitiano** *se rispettivamente $xx^* = x^*x$ ovvero $x^* = x$.*

Nota. Abbiamo immediatamente che, in virtù di (a), (b) ed (c) in proposizione 3.6:
l'algebra di Banach $\mathfrak{B}(H)$ è anche una C^*-algebra se l'involuzione è definita come la coniugazione hermitiana e gli operatori autoaggiunti coincidonocon gli elementi hermitiani di $\mathfrak{B}(H)$.

Prima di tornare alla C^*-algebra $\mathfrak{B}(H)$, diamo alcune immediate proprietà generali delle $*$-algebre che seguono dalla stessa definizione.

Proposizione 3.7. *Sia A una $*$-algebra. Indicando con * l'involuzione in A, vale quanto segue.*
(a) *Se A è C^*-algebra con norma $||\,||$ e $x \in A$ è normale allora*

$$||x^2|| = ||x||^2 \,.$$

(b) *Se A è C^*-algebra norma $||\,||$ e $x \in A$ è normale allora*

$$||x^*|| = ||x|| \,.$$

(c) *Se A ammette unità $\mathbb{I}$, vale $\mathbb{I}^* = \mathbb{I}$.*

Prova. (a) Usando ripetutamente (3.33), **I2** e **I3** e il fatto che $xx^* = x^*x$:

$$||x^2||^2 = ||(x^2)^*x^2|| = ||(x^*)^2x^2|| = ||(xx^*)^*(xx^*)|| = ||x^*x||^2 = (||x||^2)^2$$

da cui $||x^2|| = ||x||^2$ per la positività della norma.
(b) dato che $xx^* = x^*x$, (3.33) implica che $||xx^*|| = ||x||^2$. D'altra parte: $||xx^*|| = ||(x^*)^*x^*|| = ||x^*||^2$, per cui $||x|| = ||x^*||$.
(c) $\mathbb{I}\mathbb{I}^* = \mathbb{I}^*$ (1) per definizione di unità, d'altra parte $\mathbb{I}\mathbb{I}^* = (\mathbb{I}^*)^*\mathbb{I}^* = (\mathbb{I}^*\mathbb{I})^*$ (2). Da (1) e (2) segue che $\mathbb{I}^* = (\mathbb{I}^*\mathbb{I})^* = (\mathbb{I}^*)^* = \mathbb{I}$. $\square$
Tornando alla C^* algebra $\mathfrak{B}(H)$ possiamo dare alcune definizioni riguardanti i più importanti tipi di operatori.

Definizione 3.9. *Se $(H, (\,|\,)$ uno spazio di Hilbert e $T \in gB(H)$.*
(a) *T è detto* **normale** *se $TT^\dagger = T^\dagger T$.*
(b) *T è detto* **autoaggiunto** *se $T = T^\dagger$.*
(c) *T è detto* **isometrico** *se $T^\dagger T = I$; equivalentemente $T \in \mathfrak{L}(H)$ è isometrico se $(Tx|Ty) = (x|y)$ per ogni coppia $x, y \in H$.*
(d) *T è detto* **unitario** *se $T^\dagger T = TT^\dagger = I$; equivalentemente $T \in \mathfrak{L}(H)$ è unitario se è isometrico e suriettivo.*

(e) *T è detto* **positivo**, *e si scrive $T \geq 0$, se $(u|Tu) \geq 0$ per ogni $x \in H$.*
(f) *Se $U \in \mathfrak{B}(H)$, si dice che T* **maggiora** *U, e si scrive $T \geq U$ se $T - U \geq 0$.*

Note.
(1) Si osservi, a commento di (c), che se $T \in \mathfrak{B}(H)$ e $T^\dagger T = I$ allora vale $(Tx|Ty) = (x|y)$ per ogni coppia $x, y \in H$ in quanto $(x|T^\dagger Ty) = (Tx|Ty)$. Viceversa se $T \in \mathfrak{L}(H)$ e vale $(Tx|Ty) = (x|y)$ per ogni $x, y \in H$, allora T è limitato (bata porre $y = x$) e dunque esiste $T^\dagger$; infine vale $T^\dagger T = I$ perchè $(x|T^\dagger Ty) = (Tx|Ty) = (x|y)$ per ogni coppia $x, y \in H$ e quindi in particolare $(x|(T^\dagger T - I)y) = 0$ con $x = (T^\dagger T - I)y$.
A commento di (d) si osservi che ogni operatore isometrico T è ovviamente iniettivo, perchè $Tu = 0$ implica $||u|| = 0$ e quindi $u = 0$. Allora la suriettività è equivalente all'esistenza di un inversa destra che deve coincidere con quella sinistra che esiste già per l'iniettività ed è $T^\dagger$. Da ciò segue immediatamente che $T^\dagger T = TT^\dagger = I$ è equivalente a dire che $T \in \mathfrak{L}(H)$ è isometrico (e quindi limitato) ed è anche suriettivo.
(2) Esistono operatori isometrici che non sono operatori unitari (ovviamente ciò non accade se H ha dimensione finita). Un esempio, sullo spazio $l^2(\mathbb{N})$ è l'operatore

$$A : (z_0, z_1, z_2, \ldots) \mapsto (0, z_0, z_1, , \ldots) ,$$

per ogni $(z_0, z_1, z_2, \ldots) \in l^2(\mathbb{N})$.
(3) Gli operatori unitari ed autoaggiunti sono operatori normali, non vale il viceversa in generale.

Per concludere questa sezione diamo alcune proprietà elementari degli operatori normali, autoaggiunti, unitari e positivi nella seguente proposizione preceduta da una definizione.

Definizione 3.10. *Sia X spazio vettoriale sul campo $\mathbb{K}$ e $T \in \mathfrak{L}(X)$. $\lambda \in \mathbb{K}$ è detto* **autovalore** *dell'operatore T se*

$$Tu = \lambda u \tag{3.34}$$

per qualche $u \in X \setminus \{0\}$. In tal caso u è detto **autovettore** *di T con autovalore o associato all'autovalore λ. Il sottospazio di tutti gli autovettori con autovalore λ è detto* **autospazio** *di T con autovalore o associato all'autovalore λ.*

Proposizione 3.8. *Sia $(H, (\,|\,))$ spazio di Hilbert.*
(a) *Se $T \in \mathfrak{B}(H)$ è autoaggiunto allora:*

$$||T|| = \sup \{(x|Tx) \mid x \in H \,,\ ||x|| = 1\} \ . \tag{3.35}$$

(b) *Se $T \in \mathfrak{B}(H)$ è normale (e quindi in particolare autoaggiunto oppure unitario):*
(i) $\lambda \in \mathbb{C}$ è autovalore di T con autovettore u se e solo se $\overline{\lambda}$ è autovalore per $T^\dagger$ con lo stesso autovettore u;
(ii) autospazi di T con autovalori distinti sono ortogonali.
(c) *Se $T \in \mathfrak{L}(H)$ valgono i seguenti fatti:*

(i) se T è positivo i suoi possibili autovalori sono reali non negativi;
(ii) se $T \in \mathfrak{B}(H)$ è autoaggiunto i suoi possibili autovalori sono reali;
(iii) se T è isometrico (ed in particolare unitario) i suoi possibili autovalori sono complessi con modulo uguale a 1.
(d) *Se $T \in \mathfrak{L}(H)$ soddisfa $(y|Tx) = (Ty|x)$ per ogni coppia $x, y \in H$, allora T è limitato ed è autoaggiunto.*
(e) *Se $T \in \mathfrak{B}(H)$ è positivo allora è autoaggiunto.*
(f) *$\geq$ è una relazione di ordine parziale in $\mathfrak{B}(H)$.*

Prova. (a) Posto $Q := \sup\{(x|Tx) \mid x \in H\ ,\ ||x|| = 1\}$, dato che $||x|| = 1$,

$$|(x|Tx)| \leq ||Tx||||x|| \leq ||Tx|| \leq ||T||$$

Quindi: $Q \leq ||T|| = \sup\{||Tx|| \mid x \in H\ ,\ ||x|| = 1\}$. Per concludere, è sufficiente provare che $||T|| \leq Q$. Vale l'identità di immediata verifica

$$4(x|Ty) = (x+y|T(x+y)) - (x-y|T(x-y)) + i(x+iy|T(x+iy)) - i(x-iy|T(x-iy))\ .$$

da tale identittà e tenendo conto del fatto che $\overline{(z|Tz)} = (Tz|z) = (z|Tz)$, si ricava che

$$4Re(x|Ty) = (x+y|T(x+y)) - (x-y|T(x-y)) \leq Q||x+y||^2 + Q||x-y||^2 = 2Q||x||^2 + 2Q||y||^2\ .$$

Sia $y \in H$ con $||y|| = 1$. Se $Ty = 0$ allora è ovvio che $||Ty|| \leq Q$ altrimenti definiamo $x := Ty/||Ty||$ e otteniamo dalla disuguaglianza provata sopra:

$$4||Ty|| = 4Re(x|Ty) \leq 2Q(||x||^2 + ||y||^2) = 2Q(1+1) = 4Q$$

da cui, ancora $||Ty|| \leq Q$. In definitiva $||Ty|| \leq Q$ se $||y|| = 1$ e quindi $||T|| \leq Q$.
(b)(i) Se A è normale

$$||Au||^2 = (Au|Au)^2 = (A^\dagger Au|u) = (AA^\dagger u|u) = (A^\dagger u|A^\dagger u) = ||A^\dagger u||^2\ .$$

Se T è normale, $T - \lambda I$ è normale con aggiunto $T^\dagger - \overline{\lambda} I$, per cui applicando il risultato di sopra,

$$||Tu - \lambda u||^2 = ||T^\dagger u - \overline{\lambda} u||^2\ .$$

La tesi segue immediatamente. (ii) Sia u autovettore con autovalore λ e v autovettore con autovalore μ. Per (i), $\lambda(v|u) = (v|Tu) = (T^\dagger v|u) = \mu(v|u)$ per cui $(\lambda - \mu)(v|u) = 0$. Essendo $\lambda \neq \mu$, deve essere $(v|u) = 0$.
(c) Se $T \geq 0$ e $Tu = \lambda u$ con $u \neq 0$ allora $0 \leq (u|Tu) = \lambda(u|u)$ essendo $(u|u) > 0$ segue che deve essere $\lambda \geq 0$. Sia poi $T = T^\dagger$ e $Tu = \lambda u$ con $u \neq 0$. $\lambda(u|u) = (u|Tu) = (Tu|u) = \overline{\lambda}(u|u)$. Essendo $(u|u) \neq 0$ si ha $\lambda = \overline{\lambda}$ ossia $\lambda \in \mathbb{R}$. Se viceversa T è isometrico: $(u|u) = (Tu|Tu) = |\lambda|^2(u|u)$. Essendo $u \neq 0$ si ha che $|\lambda| = 1$.
(d) La tesi si prova dimostrando che T è limitato. L'unicità dell'aggiunto implica allora che $T = T^\dagger$ in quanto $(y|Tx) = (Ty|x)$ per ogni coppia $x, y \in H$. A causa del teorema del grafico

chiuso (Corollario 2 del teorema dell'applicazione aperta 2.3 in cap.2), per dimostrare che T è limitato è sufficiente provare che T è chiuso. Sia allora $\{x_n\} \subset H$ una successione convergente a x e supponiamo che i vettori Tx_n definiscano una successione convergente, dobbiamo mostrare che $Tx_n \to Tx$. Nelle nostre ipotesi, fissato $y \in H$

$$(y|Tx_n) = (Ty|x_n) \to (Ty|x) = (y|Tx)$$

Per la continuità del prodotto scalare e tenendo conto che esiste per ipotesi $\lim_{n\to+\infty} Tx_n$, possiamo allora scrivere:

$$\left(y \middle| Tx - \lim_{n\to+\infty} Tx_n\right) = 0\,.$$

Dato che y è arbitrario, scegliendo proprio $y := Tx - \lim_{n\to+\infty} Tx_n$ concludiamo che $Tx - \lim_{n\to+\infty} Tx_n = 0$.
(e) Se $T \in \mathfrak{B}(H)$ è positivo allora $(x|Tx)$ è reale e coincide quindi con il suo complesso coniugato che vale $(Tx|x)$ per le proprietà del prodotto scalare. Quindi $((T^\dagger - T)x|x) = 0$ per ogni $x \in H$. Per (2) in esercizi 3.2 deve essere $T^\dagger - T = 0$ oia $T = T^\dagger$.
(f) Bisogna provare tre fatti. (i) $T \geq T$: questo è ovvio perchè significa che $((T-T)x|x) \geq 0$ per ogni $x \in H$. (ii) se $T \geq U$ e $U \geq S$ allora $T \geq S$: questo si prova immediatamente notando che $T - S = (T-U) + (U-S)$ e quindi $((T-S)x|x) = ((T-U)x|x) + ((U-S)x|x) \geq 0$ per ogni $x \in H$ dato che $T \geq U$ e $U \geq S$. (iii) se $T \geq U$ e $U \geq T$ allora $T = U$. Per provare (iii) si osservi che nelle ipotesi fatte $(x|(T-U)x) = 0$ per ogni $x \in H$. Per (2) in esercizi 3.2 deve essere $T - U = 0$ ossia $T = U$.
Si osservi che per spazi di Hilbert sul campo reale $\geq$ non è una relazione d'ordine parziale. □

3.4 Proiettori ortogonali.

Come ultimo concetto elementare introduciamo la nozione di *proiettore ortogonale* che giocherà il ruolo centrale nel costruire il formalismo della meccanica quantistica.
In riferimento alla definizione 2.10 ed alle successive proposizioni 4.2 e 5.2 possiamo dare la seguente definizione.

Definizione 3.11. *Se $(H, (\,|\,))$ è spazio di Hilbert, un proiettore $P \in \mathfrak{B}(H)$ è detto* **proiettore ortogonale** *se $P^\dagger = P$.*

Nota. Quindi i proiettori ortogonali sono tutti e soli gli operatori limitati da H in H definiti dalle due condizioni $P = PP$ (idempotenza) e $P = P^\dagger$ (auto aggiuntezza). Si osservi che come immediata conseguenza della definizione data si ha la *positività* dei proiettori ortogonali: per ogni $x \in H$

$$(u|Pu) = (u|PPu) = (P^\dagger u|Pu) = (Pu|Pu) = ||Pu||^2 \geq 0\,.$$

Abbiamo la seguente coppia di proposizioni che caratterizzano i proiettori ortogonali.

Proposizione 3.9. *Sia $P \in \mathfrak{B}(H)$ un proiettore ortogonale nello spazio di Hilbert H che proietta su M, allora vale quanto segue.*
(a) *Il proiettore associato $Q := I - P$ è ancora un proiettore ortogonale.*
(b) *$Q(H) = M^\perp$ per cui la decomposizione diretta associata a P secondo (b) di proposizione 2.4 è quella dovuta a M ed al suo ortogonale $M^\perp$*

$$H = M \oplus M^\perp \, .$$

(c) *Per ogni $x \in H$, $||x - P(x)|| = \min\{||x - y|| \mid y \in M\}$.*
(d) *Se N è una base hilbertiana di M allora*

$$P = \sum_{u \in N} u \, (u| \ ) \, .$$

Prova. (a) Sappiamo già che $Q := I - P$ è proiettore (proposizione 2.4). Per (c) di proposizione 3.6 essendo anche $I^\dagger = I$ segue subito che $Q^\dagger = Q$ per cui Q è proiettore ortogonale.
(b) È sufficiente provare che $Q(H) = M^\perp$. A tal fine si noti che se $x \in Q(H)$ e $y \in M$ allora $(x|y) = (Qx|y) = (x|Qy) = (x|y - Py) = (x|y - y) = 0$ per cui $Q(H) \subset M^\perp$. Se non fosse $Q(H) = M^\perp$ dovrebbe essere $M^\perp = Q(H) \oplus Q(H)^{\perp'}$ dove $Q(H)^{\perp'} \neq \{0\}$ è l'ortogonale nello spazio di Hilbert $M^\perp$ rispetto al prodotto scalare ristretto a tale sottospazio chiuso. Ma allora per (d) di teorema 3.1 varrebbe la decomposizione diretta in spazi ortogonali $H = M \oplus Q(M) \oplus Q(H)^{\perp'}$ in contraddizione con $H =< M, Q(H) >$ e questo è assurdo.
(c) La tesi è immediata conseguenza di (d) del teorema 3.1. quando $K := M$ tenuto conto dell'unicità della decomposizione diretta.
(d) Completando N a base hilbertiana di H unendola con una bae hilbertiana N' di $M^\perp$, si verifica immediatamente che,

$$R = \sum_{u \in N} u \, (u| \ )$$

e

$$R' = \sum_{u \in N'} u \, (u| \ )$$

sono operatori limitati, soddisfano $RR = R$, $R(H) = M$, $R'R' = R'$, $R'(H) = M^\perp$ ed infine $R'R = RR' = 0$ e $R + R' = I$. Per la proposizione 2.4 R ed R' sono proiettori associati alla decompoizione diretta $M \oplus M^\perp$. Per l'unicità della decomposiazione su M e $M^\perp$ di ogni vettore, concludiamo che deve essere $R = P$ (e $R' = Q$). $\square$

Proposizione 3.10. *Sia H spazio di Hilbert e $M \subset H$ un sottospazio chiuso. I proiettori P e Q associati alla decomposizione diretta ortogonale $H = M \oplus M^\perp$ secondo la proposizione 2.5 (con $N := M^\perp$) e che proiettano rispettivamente su M e $M^\perp$ sono proiettori ortogonali.*

Prova. Bisogna solo provare che $P = P^\dagger$ e $Q = Q^\dagger$.
Se $x \in H$ allora si ha la decomposizione univoca $x = y + z$ con $y = P(x) \in M$ e $z = Q(x) \in M^\perp$. Sia $x' = y' + z'$ l'analoga decomposizione per $x' \in H$. Vale $(x'|Px) = (y' + z'|y) = (y'|y)$. Ma anche $(Px'|x) = (y'|y + z) = (y'|y)$ per cui $(x'|Px) = (Px'|x)$ ossia $((P^\dagger - P)x'|x) = 0$ per ogni $x, x' \in H$. Scegliendo $x = (P^\dagger - P)x'$ si vede che deve valere $Px' = P^\dagger x'$ per ogni x' e quindi $P = P^\dagger$. La prova per Q è la stessa ricordando che $(M^\perp)^\perp = M$. $\square$

Esercizi 3.2.
(1) Mostrare che se H è spazio di Hilbert, $U \in \mathfrak{B}(H)$ è operatore isometrico se e solo se $||Ux|| = ||x||$ per ogni $x \in H$.
Più in generale, provare che se $T : \mathcal{D}(T) \to X$ è un operatore sullo spazio vettoriale sul campo complesso e con prodotto scalare $(\,|\,)$, dove $\mathcal{D}(T)$ è un sottopazio (non necessariamente chiuso) di X, allora i due fatti seguenti sono equivalenti ($||\,||$ indica la norma indotta dal prodotto scalare):
(a) $||Tx|| = ||x||$ per ogni $x \in \mathcal{D}(T)$;
(b) $(Tx|Ty) = (x|y)$ per ogni coppia $x, y \in \mathcal{D}(T)$.
Suggerimento. (5) di Esercizi 3.1
(2) Sia H spazio di Hilbert e $T \in \mathfrak{L}(H)$ soddisfi $(u|Tu) = 0$ per ogni $u \in H$. Si dimostri che T è l'operatore nullo (cioè che manda ogni vettore in 0).
Soluzione. Vale

$$0 = (u + v|T(u + v)) = (u|Tu) + (v|Tv) + (u|Tv) + (v|Tu) = (u|Tv) + (v|Tu)\,.$$

Similmente

$$0 = i(u + iv|T(u + iv)) = i(u|Tu) + i(v|Tv) - (u|Tv) + (v|Tu) = -(u|Tv) + (v|Tu)\,.$$

Quindi, sommando membro a membro $(v|Tu) = 0$ per ogni $u, v \in H$. Scegliendo $v := Tu$ si ha che $||Tu||^2 = 0$ per ogni $u \in H$ ossia $Tu = 0$ per ogni $u \in H$.
(3) Si consideri $L^2([0, 1], m)$ dove m è la solita misura di Lebesgue. Sia $f \in \mathfrak{L}^2([0, 1], m)$. Mostrare che l'operatore $T_f : L^2([0, 1], m) \ni g \mapsto f \cdot g$, dove $\cdot$ è l'ordinario prodotto di funzioni punto per punto, è ben definito, limitato con norma $||T_f|| = ||f||$ e normale. Infine T_f è autoaggiunto se e solo se f ammette solo valori reali. **(4)** Sia $T \in \mathfrak{B}(H)$ autoaggiunto. Per $\lambda \in \mathbb{R}$, si consideri la serie di operatori

$$U(\lambda) := \sum_{n=0}^{\infty} (\lambda i)^n \frac{T^n}{n!}\,,$$

dove $T^0 := I$, $T^1 := T$, $T^2 := TT$ e via di seguito e la convergenza è quella nella topologia uniforme. Si dimostri che la serie converge ad un operatore unitario.
(5) In riferimento all'esercizio precedente, provare che se $\lambda, \mu \in \mathbb{R}$ $U(\lambda)U(\mu) = U(\lambda + \mu)$.
(6) Mostrare che la serie nell'esercizio (5) converge per ogni $\lambda \in \mathbb{C}$ ad un operatore limitato e che $U(\lambda)$ è sempre normale.
(7) Mostrare che l'operatore $U(\lambda)$ dell'esercizio (7) è positivo se $i\lambda \in \mathbb{R}$. Ci sono valori di $\lambda \in \mathbb{C}$

per cui $U(\lambda)$ è un proiettore (non necessariamente ortogonale)?
(8) Calcolare esplicitamente $U(\lambda)$ dell'esercizio (5) se T è definito come T_f dell'esercizio (4) con $f = \overline{f}$.
(9) Se, in $l(\mathbb{N})$, $T : \{x_n\} \mapsto \{x_{n+1}/n\}$, provare che T è limitato e calcolare $T^\dagger$.

3.5 Radici quadrate di operatori positivi e decomposizione polare di operatori limitati.

In questa sezione piuttosto tecnica veremo alcune nozioni molto utili in teoria deglie operatori limitati in uno spazio di Hilbert. Il risultato più importante di questa sezione è il cosiddetto *teorema di decomposizione polare per operatori limitati.* Si tratta di una decomposizione che generalizza ad opreatori la decomposizione di un numero complesso z nel prodotto del suo valore assoluto e di un esponenziale immaginario: $z = |z|e^{i\,arg\,z}$. La nozione valore assoluto di un operatore è utile per introdurre una generalizzazione del concetto di "convergenza assoluta" di serie numeriche costruite a partire da operatori e basi hilbertiane. Useremo queste serie nel definire gli operatori di Hilbert-Schmidt e gli operatori di classe traccia che rappresentano gli stati in meccanica quantistica. Parte delle dimostrazioni sono tratte da [14] e [15].

Definizione 3.12. *Se $A \in \mathfrak{B}(H)$, con H spazio di Hilbert, $B \in \mathfrak{B}(H)$ si dice* **radice quadrata** *ovvero* **radice quadrata (positiva)** *di A e $B^2 = A$ (e rispettivamente $B \geq 0$).*

Mostreremo tra poco che ogni operatore limitato positivo ha una ed una sola radice quadrata positiva. Ci serve un lemma iniziale che diamo sotto forma di proposizione.

Proposizione 3.11 *Sia H spazio di Hilbert. Se $\{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{B}(H)$ è una successione monotona non decrescente (non crescente) di operatori autoaggiunti limitata superiormente (rispettivamente inferiormente) da $K \in \mathfrak{B}(H)$, allora esiste $A \in \mathfrak{B}(H)$ autoaggiunto con $A \leq K$ (rispettivamente $A \geq K$) e tale che:*

$$A = s\text{-}\lim_{n\to+\infty} A_n\,. \tag{3.36}$$

Prova. Dimostraimo la tesi nel caso non decrescente, l'altro caso si riporta a questo considerando la successione di termini $K - A_n$.
Posto $B_n := A_n + ||A_0||$ per ogni $n \in \mathbb{N}$, si verifica facilmente che la successione dei B_n è monotona non decrescente fatta di operatori positivi e vale ancora $B_n \leq K + ||A_0|| =: K_1$ e K_1 è positivo. Dato che che ogni operatore positivo T definisce un semiprodotto scalare per cui è valida la disuguaglianza di Schwartz

$$|(x|Ty)|^2 \leq (x|Tx)(y|Ty))\,, \tag{3.37}$$

avremo che, se $n \geq m$:

$$|(x|(B_n - B_m)y)|^2 \leq (x|(B_n - B_m)x)(y|(B_n - B_m)y) \leq (x|K_1 x)(y|K_1 y) \leq ||K_1||^2 ||x||^2 ||y||^2 .$$

Da tale disuguaglianza, se si pone $x = (B_n - B_m)y$ si ricava che $||B_n - B_m|| \leq ||K_1||$ Da questa e dalla (3.37) con $y = (B_n - B_m)x$ e $T = B_n - B_m$ si ricava che, per ogni $x \in H$, se $n \geq m$:

$$||(B_n - B_m)x||^4 = ((B_n - B_m)x|(B_n - B_m)x)^2 \leq (x|(B_n - B_m)x)((B_n - B_m)x|(B_n - B_m)^2 x) .$$

l'ultimo termine è maggiorato da

$$(x|(B_n - B_m)x)||B_n - B_m||^3 ||x||^2 \leq ||K_1||^3 ||x||^2 [(x|B_n x) - (x|B_m x)] .$$

La successione non decrescente limitata di numeri positivi $(x|B_k x)$ è necessariamente convergente per cui è di Cauchy. Concludiamo che deve essere di Cauchy anche quella di vettori $B_k x$. Esisterà dunque il limite per $k \to +\infty$ di tale successione. Definiamo

$$B : H \ni x \mapsto \lim_{n \to \infty} B_n x .$$

Si verifica facilmente che per costruzione B è lineare, soddisfa $(Bx|x) = (x|Bx) \leq (x|K_1 x)$. Dato che K_1 è limitato ed autoaggiunto (essendo positivo), da (a) di proposizione 3.8 concludiamo che B è limitato ed autoaggiunto. Quindi, $A := B - ||A_0||$ è un operatore limitato autoaggiunto e vale

$$Ax = \lim_{n \to +\infty} (B_n - ||A_0||)x) = \lim_{n \to +\infty} A_n x$$

inoltre $A \leq K$. $\square$

Come detto sopra questo risultato ci permette di dimostrare l'esistenza delle radici quadrate degli operatori limitati positivi.

Teorema 3.8. *Sia H spazio di Hilbert e $A \in \mathfrak{B}(H)$ un operatore positivo. Esiste un'unica radice quadrata positiva, che indichiamo con $\sqrt{A}$. Inoltre:*
(a) *$\sqrt{A}$ commuta con tutti gli operatori limitati che commutano con A;*
(b) *se A è biettivo è biettivo anche $\sqrt{A}$.*

Nota. Come d'uso comune nella teoria degli operatori e delle algebre, dire che due operatori A e B *commutano*, ovvero che A *commuta con* B ovvero che B *commuta con* A, significa semplicemente che

$$AB = BA . \tag{3.38}$$

Senza precisare null'altro si sottointende quindi che l'identità scritta valga sul dominio di A e di B che devono coincidere ed inoltre $Ran(A)$ deve essere incluso nel dominio di B e $Ran(B)$ deve essere incluso nel dominio di A. È chiaro che tutto ciò è banalmente vero usando operatori $A, B \in \mathfrak{L}(X)$ dove X è un qualsiasi spazio vettoriale.

Prova del teorema 3.8. Se $A = 0$ la tesi è ovviamente vera: 0 è l'unica radice quadrata positiva di A perchè $B^2 = 0$ implica $(x|B^2x) = 0$ e quindi $||Bx||^2 = 0$ per ogni x da cui $B = 0$. Sia quindi $A \neq 0$ e poniamo $A' := A/||A||$. Vale $A' \leq I$ essendo per (a) di proposizione 3.8 e tenendo conto del fatto che $(x|ax) \geq 0$:

$$(x|(I - A')x) = ||x||^2 - \frac{(x|Ax)}{||A||} \geq 0 \,.$$

Poniamo $B_0 := 0$ e, per ogni $n = 0, 1, 2, \ldots$

$$B_{n+1} = B_n + (A' - B_n^2)/2 \,. \tag{3.39}$$

È chiaro che se la successione dei B_n converge a qualche limite B, dovrà essere $A' - B^2 = 0$ cioè $A = B^2$. Mostriamo che in effetti la successione converge davvero in senso forte.
Ogni B_n è autoaggiunto e, per induzione, ogni B_n è un polinomio in A' per cui i B_n commutano fra loro. Si ha che:

$$I - B_n = (I - B_{n-1})^2/2 + (I - A')/2 \geq 0$$

e, per induzione dalla formula appena scritta e da

$$B_{n+1} - B_n = (B_n - B_{n-1})[I - (B_n + B_{n-1})/2]$$

i ha che $B_n \leq B_{n+1}$. Per la proposizione 3.11 esiste

$$B := \text{s-}\lim_{n\to+\infty} B_n$$

e B è limitato autoaggiunto e $0 \leq B \leq I$. Dalla (3.39), passando al limite si ha che $A' = B^2$. Quindi $B' := \sqrt{||A||}B$ è una radice quadrata positiva di A. Si osservi che B' è limite forte di polinomi in A, $P_n = \sqrt{||A||}B_n$, che quindi commutano con A e con tutti gli operatori che commutano con A, ne consegue che B' commuta con tutti gli operatori limitati (ossia continui) T che commutano con A:

$$B'Tx = \lim_{n\to+\infty} P_nTx = \lim_{n\to+\infty} TP_nx = T\lim_{n\to+\infty} P_nx = TB'x \,.$$

Proviamo l'unicità della radice quadrata positiva. Siano V e V' due radici quadrate positive di A e poniamo esplicitamente $V := B'$ che ha la proprietà di commutare con tutti gli operatori che commutano con A. Siccome

$$AV' = V'^3 = V'A \,,$$

V' e A commutano tra di loro e di conseguenza V commuta con V'. Prendiamo $x \in H$ arbitrario e poniamo $y := V'x - Vx$. Abbiamo:

$$(y|V'y) + (y|Vy) = (y|[V' + V][V' - V]x) = (y|[V'^2 - V^2]x) = 0 \,,$$

dove abbiamo usato il fatto che $V'^2 = V^2 = A$. Di conseguenza, poichè $(y|Vy) \geq 0$ e $(y|V'y) \geq 0$ deve essere

$$(y|Vy) = (y|V'y) = 0 \,.$$

Indichiamo con W una radice quadrata di V. Dato che

$$||Wy||^2 = (Wy|Wy) = (y|W^2y) = (y|Vy) = 0$$

dovrà essere $Wy = 0$ ed a maggior ragione $Vy = W(Wy) = 0$. Nello stesso modo si trova $V'y = 0$. Per concludere notiamo che allora

$$||V'x - Vx||^2 = (x|[V' - V]^2x) = (x|[V' - V]y) = 0$$

e quindi $Vx = V'x$ per ogni $x \in H$, che significa $V = V'$.
Non resta che provare che $\sqrt{A}$ è biettivo se lo è A. Se A è biettivo allora commuta con A^{-1} di coneguenza commuta con A^{-1} anche $\sqrt{A}$. Si verifica allora immediatamente che $A^{-1}\sqrt{A} = \sqrt{A}A^{-1}$ è invero destro e sinistro di $\sqrt{A}$ che è di conseguenza biettivo. □

Corollario. *Sia H spazio di Hilbert, se $A, B \in \mathfrak{B}(H)$ sono positivi e commutano tra di loro allora il loro prodotto è un operatore limitato positivo.*

Prova. $\sqrt{B}$ commuta con A, perciò

$$(x|ABx) = (x|A\sqrt{B}^2x) = (x|\sqrt{B}A\sqrt{B}) = (\sqrt{B}x|A\sqrt{B}x) \geq 0\,. \quad \square$$

Concludiamo la sezione mostrando che ogni operatore limitato A su uno spazio di Hilbert ammette una decomposizione in un prodotto di un operatore positivo P, univocamente determinato, e un operatore isometrico U definito sul rango di quello positivo ed ivi univocamente determinato: $A = UP$. Tale decomposizione si dice *decomposizione polare* dell'operatore ed ha molteplici applicazioni in fisica matematica. Noi ueremo la decomposizione polare di operatori limitati compatti nella prossima sezione. Una definizione preliminare è necessaria.

Definizione 3.12. *Sia H spazio di Hilbert e $A \in \mathfrak{B}(H)$. L'operatore limitato, positivo e autoaggiunto:*

$$|A| := \sqrt{A^\dagger A}\,, \tag{3.40}$$

è detto **valore assoluto** *di A.*

Nota. Valendo, per ogni $x \in H$: $||\,|A|\,x||^2 = (x|\,|A|^2x) = (x|A^\dagger Ax) = ||Ax||^2$, otteniamo

$$||\,|A|\,x|| = ||Ax||\,, \tag{3.41}$$

da cui segue che

$$Ker(|A|) = Ker(A) \tag{3.42}$$

e quindi $|A|$ è iniettivo se e solo se lo è A. Un'altra utile proprietà è:

$$\overline{Ran(|A|)} = (Ker(A))^\perp\,, \tag{3.43}$$

che vale in quanto $\overline{Ran(|A|)} = ((Ran(|A|)^{\perp})^{\perp} = (Ker(|A|)^{\perp} = (Ker(A))^{\perp}$.

Passiamo al teorema di decomposizione polare.

Teorema 3.9. *Sia H spazio di Hilbert e $A \in \mathfrak{B}(H)$.*
(a) *Esiste una sola coppia di operatori $P, U \in \mathfrak{B}(H)$ tali che valgano insieme le condizioni seguenti:*
(1) vale la decomposizione

$$A = UP\,, \tag{3.44}$$

(2) P è positivo,
(3) U è isometrico su $Ran(P)$ (ossia $||U(x)|| = ||x||$ per ogni $x \in Ran(P)$)

(4) $Ker(U) = Ker(P)$ (o equivalentemente (4)' $Ker(U) = Ker(A)$).
(b) *Risulta essere $P = |A|$, $Ker(U) = Ker(A)$, e U è isometrico su $(Ker(A))^{\perp} = \overline{Ran(|A|)}$.*
(c) *Se A e $A^{\dagger}$ sono biettivi, U coincide con l'operatore unitario $A|A|^{-1}$.*

Prova. (a) Cominciamo con il provare l'unicità. Se esiste la decomposizione (3.44) $A = UP$ con $P \geq 0$ (oltre che limitato) e U limitato, allora $A^{\dagger} = PU^{\dagger}$ essendo P autoaggiunto perchè positivo ((3) in esercizi 3.2) quindi

$$A^{\dagger}A = PU^{\dagger}UP\,. \tag{3.45}$$

La condizione che U sia isometrico su $Ran(P)$ si scrive $(UPx|UPy) = (Px|Py)$ per ogni $x, y \in H$, che equivale a $(x|[PU^{\dagger}UP - P^2]y) = 0$ per ogni $x, y \in H$. Quindi $PU^{\dagger}UP = P^2$. Sostituendo in (3.45) concludiamo che deve essere $P^2 = A^{\dagger}A$ e dunque, essendo P positivo ed estraendo l'unica radice quadrata positiva (teorema 3.8) ad ambo membri, troviamo $P = |A|$. Quindi, se esiste una decomposizione con i requisiti fisati in (a), necessariamente $P = |A|$. Mostriamo che anche U è fissato unicamente. Notiamo che $H = Ker(P) \oplus (Ker(P))^{\perp}$ ma, usando (c) di proposizione 3.6 vale anche $(Ker(P))^{\perp} = \overline{Ran(P^{\dagger})} = \overline{Ran(P)}$ per il fatto che P è autoaggiunto essendo positivo ((3) in esercizi 3.2). Quindi $H = Ker(P) \oplus \overline{Ran(P)}$. Per assegnare un operatore su H è sufficiente assegnarlo su ciascuno dei due addendi della somma diretta scritta sopra. $U = 0$ su $Ker(P)$ nelle ipotesi del teorema, mentre la condizione $UP(x) = A(x)$ per ogni $x \in H$ fissa univocamente U su $Ran(P)$. D'altra parte essendo per ipotesi U limitato lo sarà anche ristretto a $Ran(P)$. Un operatore limitato definito in un dominio denso è estendibile univocamente ad un operatore definito sulla chiusura del dominio che risulta ancora essere limitato (vedi (3) e (4) in esercizi 2.1). Quindi U è in realtà fissato dalle ipotesi fatte su tutto $\overline{Ran(P)}$ e quindi su tutto H. La dimostrazione di unicità è conclusa, passiamo a quella di esistenza.
In base a quanto visto sopra è sufficiente mostrare che la condizione $UP = A$ con $P = |A|$, riscritta equivalentemente come $U : |A|(x) \mapsto A(x)$ per ogni $x \in H$, definisce effettivamente un operatore, che indicheremo con U_0, su $Ran(|A|)$ (potrebbe non essere nemmeno una funzione) e che tale operatore è isometrico. Notiamo che da (3.41) se $Ax = Ay$ allora $|A|x = |A|y$ e quindi l'applicazione $U_0 : Ran(|A|) \ni |A|(x) \mapsto A(x)$ è ben definita (non è plurivoca). Il fatto che U_0 sia lineare è ovvio per costruzione come lo è il fatto che sia un'isometria preservando

le norme per (3.41) (si tenga conto di (1) in esercizi 3.2). Il fatto che U_0 sia un'isometria su $Ran(|A|)$ implica, per continuità, che si etenda in modo unico ad un'isometria sulla chiusura di $Ran(|A|)$. Indichiamo ancora con U_0 tale estensione. Infine definiamo un operatore $U : H \to H$ tale che rispetto alla somma diretta vista sopra $H = Ker(|A|) \oplus \overline{Ran(|A|)}$ si decomponga in $U \restriction_{Ker(|A|)} := 0$ e $U \restriction_{\overline{Ran(|A|)}} := U_0$. È immediato verificare che $U \in \mathfrak{B}(H)$ e che U soddisfa (3.44) per costruzione. Infine vale $Ker(U) \subset Ker(|A|)$ per costruzione. Proviamo per concludere, che i due nuclei in realtà coincidono. Ogni eventuale u con $Uu = 0$ si decompone come $u_0 + x$ con $u_0 \in Ker(|A|)$ su cui U i annulla e $x \in \overline{Ran(|A|)}$, per cui deve valere $U_0 x = 0$. Dato che su tale spazio U_0 è isometrico deve essere $x = 0$ e quindi $u = u_0 \in Ker(|A|)$. Quindi $Ker(U) \supset Ker(|A|)$ e tenendo conto anche dell'altra inclusione vale infine $Ker(U) = Ker(|A|) = Ker(A)$ per (3.42).
(b) è già stato provato completamente provando (a).
(c) Se A e $A^\dagger$ sono biettivi allora è biettivo $A^\dagger A$ e quindi lo è $|A| = \sqrt{A^\dagger A}$ per (b) del teorema 3.8 e infine lo è $U := A|A|^{-1}$ perchè composizione di operatori biettivi. Vale $U^\dagger = |A|^{-1}A^\dagger$, dove si è tenuto conto del fatto che $|A|^{-1}$ è positivo (come segue subito dal fatto che è biettivo e che $|A|$ è positivo per definizione). Allora vale anche

$$U^\dagger U = |A|^{-1}A^\dagger A|A|^{-1} = |A|^{-1}|A|^2|A|^{-1} = I$$

e quindi U è isometrico su tutto H. Ulteriormente $Ker(U) = \{0\} = Ker(|A|)$ perchè entrambi gli operatori sono iniettivi. Infine, per definizione: $A = U|A|$. Abbiamo provato che $U|A|$ è una, e quindi l'unica, decomposizione di A che soddisfa le richieste del teorema esplicitate in (a). Il fatto che U sia unitario, segue dal fatto che U è isometrico e suriettivo (perchè biettivo). □

Definizione 3.13. *Sia H spazio di Hilbert e $A \in \mathfrak{B}(H)$. Si dice* **decomposizione polare** *dell'operatore A la decomposizione*

$$A = UP\,,$$

dove $P, U \in \mathfrak{B}(H)$ soddisfano le proprietà: P è positivo, U è isometrico su $Ran(P)$ e $Ker(U) = Ker(P)$.

Un utile corollario del teorema di decomposizione polare è il seguente.

Corollario. *Nelle ipotesi del teorema 3.9, se $U|A| = A$ è la decomposizione polare di A, vale l'dentità:*

$$|A^\dagger| = U|A|U^\dagger\,. \tag{3.46}$$

Prova. Da $A = U|A|$ segue immediatamente che $A^\dagger = |A|U^\dagger = U^\dagger\, U|A|U^\dagger$, dove si è tenuto conto del fatto che $U^\dagger U|A| = |A|$ dato che U è isometrico su $Ran(A) = Ran(|A|)$. Quindi, per l'operatore autoaggiunto $AA^\dagger$ vale:

$$AA^\dagger = U|A|U^\dagger\, U|A|U^\dagger\,.$$

Dato che $U|A|U^\dagger$ è evidentemente positivo avremo, per l'unicità della radice quadrata:

$$|A^\dagger| = \sqrt{(A^\dagger)^\dagger A^\dagger} = \sqrt{AA^\dagger} = U|A|U^\dagger \,.$$

Questo prova la tesi. □

3.6 *La trasformata di Fourier-Plancherel.

Introduciamo molto brevemente la trasformata di Fourier-Plancherel.

Notazione 3.3. I punti di $\mathbb{R}^n$ saranno indicati con singole lettere e le componenti con la stessa lettera ed un indice, in tal modo $x = (x_1, \ldots, x_n)$, dx denoterà l'ordinaria misura di Lebesgue in $\mathbb{R}^n$.
Diremo **multiindice** ogli n-pla, $\alpha = (\alpha_1, \ldots, \alpha_n)$ con $\alpha_i = 0, 1, 2, \ldots$ e indicheremo con $|\alpha|$ la somma $|\alpha| := \sum_{i=1}^n \alpha_i$. Useremo ancora le seguenti notazioni

$$\partial_x^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \,,$$

e nello stesso modo

$$M^\alpha(x) := x_1^{\alpha_1} \cdots x_n^{\alpha_n} \,.$$

Con $\mathcal{D}(\mathbb{R}^n)$ indicheremo lo spazio, anche indicato con $C_c^\infty(\mathbb{R}^n)$, delle funzioni a valori complessi infinitamente differenziabili con supporto compatto. Con $\mathcal{S}(\mathbb{R}^n)$ indicheremo lo **spazio di Schwartz** su $\mathbb{R}^n$, ovvero lo spazio vettoriale complesso delle funzioni $C^\infty(\mathbb{R}^n)$ a valori complessi che godono dell'ulteriore proprietà: per ogni $f \in \mathcal{S}(\mathbb{R}^n)$ e per ogni coppia di multiindici α e β esiste $K < \infty$ (dipendente da f, α e β!) tale che

$$|M^\alpha(x)\partial_x^\beta f(x)| \leq K, \quad \text{per ogni } x \in \mathbb{R}^n \,. \tag{3.47}$$

Le norme $||\,||_1$, $||\,||_2$ e $||\,||_\infty$ in questa sezione denoteranno rispettivamente la norma di $L^1(\mathbb{R}^n, dx)$, $L^2(\mathbb{R}^n, dx)$ e $L^\infty(\mathbb{R}^n, dx)$ (vedi (6) e (8) in Esempi 2.1).

Ovviamente $\mathcal{D}(\mathbb{R}^n)$ e $\mathcal{S}(\mathbb{R}^n)$ sono invarianti sotto l'azione degli operatori M^α (visto come operatore moltiplicativo) e ∂^α: funzioni di ciascuno dei due spazi rimangono nel rispettivo spazio quando trasformate sotto l'azione di M^α e ∂^α.
È chiaro che $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{L}^p(\mathbb{R}, dx)$ come sottospazio per ogni $1 \leq p \leq \infty$ dato che i compatti su $\mathbb{R}^n$ hanno misura di Lebesgue finita e che ogni $f \in \mathcal{D}(\mathbb{R}^n)$ è continua e quindi limitata sui compatti.
Vale anche che $\mathcal{S}(\mathbb{R}^n) \subset \mathcal{L}^p(\mathbb{R}, dx)$ per ogni $1 \leq p < \infty$ come sottospazio. Infatti, se $C \subset \mathbb{R}^n$ è un compatto contenente l'origine, $f \in \mathcal{S}(\mathbb{R}^n)$ è limitata su C essendo ivi continua e, fuori da C vale $|f(x)| \leq C_n|x|^{-n}$ per ogni $n = 0, 1, 2, 3, \ldots$ pur di scegliere $C_n \geq 0$ sufficientemente grande. In definitiva $|f|$ è limitata su $\mathbb{R}^n$ e quindi appartiene a $\mathcal{L}^\infty$ ma, per ogni $p \in [1, +\infty)$ è anche limitata da una funzione che appartiene a $\mathcal{L}^p$, tale funzione è costante su C e vale $1/|x|^n$ con

$n > 1/p$ fuori da C.
Oltre al fatto ovvio che $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, ricordiamo un fatto notissimo che sarà utile tra poco: *gli spazi $\mathcal{D}(\mathbb{R}^n)$ e $\mathcal{S}(\mathbb{R}^n)$ sono sottospazi densi di ogni $\mathcal{L}^p(\mathbb{R}, dx)$ per ogni $1 \leq p < \infty$.*
(Vedi per esempio [17].)

Definizione 3.14. *Le trasformazioni lineari da $\mathcal{L}^1(\mathbb{R}^n, dx)$ in $\mathcal{L}^\infty(\mathbb{R}^n, dx)$ definite da*

$$(\mathcal{F}f)(k) = \int_{\mathbb{R}^n} \frac{e^{ip\cdot x}}{(2\pi)^{n/2}} f(x)\, dx\,, \quad \textit{per ogni } f \in \mathcal{L}^1(\mathbb{R}^n, dx) \textit{ e ogni } k \in \mathbb{R}^n \tag{3.48}$$

$$(\mathcal{F}_- g)(x) = \int_{\mathbb{R}^n} \frac{e^{-ip\cdot x}}{(2\pi)^{n/2}} g(k)\, dk\,, \quad \textit{per ogni } g \in \mathcal{L}^1(\mathbb{R}^n, dk) \textit{ e ogni } x \in \mathbb{R}^n \tag{3.49}$$

sono dette rispettivamente: **trasformata di Fourier** *e* **trasformata inversa di Fourier**.

Note.
(1) Si osservi che dk indica comunque la misura di Lebesgue in $\mathbb{R}^n$. Abbiamo usato un nuovo nome per la variabile di $\mathbb{R}^n$ (k invece di x) nello esplicitare la trasformata inversa di Fourier solo perchè tale notazione è quella tradizionale e ciò i rivelà comodo nei calcoli.
(2) È ovvio che per le proprietà elementari dell'integrale,

$$|(\mathcal{F}f)(k)| \leq \left| \int e^{ip\cdot k} f(x) \frac{dx}{(2\pi)^{n/2}} \right| \leq \int |e^{ip\cdot k}|\, |f(x)| \frac{dx^n}{(2\pi)^{n/2}} = \int |f(x)| \frac{dx}{(2\pi)^{n/2}} = \frac{||f||_1}{(2\pi)^{n/2}}$$

e similmente $|(\mathcal{F}_- g)(x)| \leq ||g||_1/(2\pi^{n/2})$ per ogni $x, k \in \mathbb{R}^n$, per cui ha senso definire la trasformata di Fourier e la trasformata inversa di Fourier come operatori a valori almeno in $L^\infty(\mathbb{R}^n, dx)$.

Vediamo nel seguito solo le proprietà più immediate della trasformata di Fourier che si connettono più direttamente con la trasformazione di Fourier-Plancherel. Tralasceremo importanti risultati come la continuità rispetto alla topologia naturale indotta da seminorme nello spazio di Schwartz per i quali si rimanda a qualsiasi testo di teoria delle funzioni ed analisi funzionale oppure teoria delle distribuzioni.

Proposizione 3.12. *La trasformazione di Fourier e la trasformazione inversa di Fourier godono delle seguenti proprietà.*
(a) *Sono continue rispetto alle norme naturali del dominio e codominio valendo le disuguaglianze*

$$||\mathcal{F}f||_\infty \leq \frac{||f||_1}{(2\pi)^{n/2}} \quad e \quad ||\mathcal{F}_- g||_\infty \leq \frac{||g||_1}{(2\pi)^{n/2}}.$$

(b) *Lo spazio di Schwartz è invariante sotto $\mathcal{F}$ e $\mathcal{F}_-$, cioè $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) \subset \mathcal{S}(\mathbb{R}^n)$ e $\mathcal{F}_-(\mathcal{S}(\mathbb{R}^n)) \subset \mathcal{S}(\mathbb{R}^n)$.*
(c) *Le due trasformazioni ritrette allo spazio invariante $\mathcal{S}(\mathbb{R}^n)$ sono una l'inversa dell'altra: se $f \in \mathcal{S}(\mathbb{R}^n)$*

$$g(k) = \int_{\mathbb{R}^n} \frac{e^{ip\cdot x}}{(2\pi)^{n/2}} f(x)\, dx$$

se e solo se

$$f(x) = \int_{\mathbb{R}^n} \frac{e^{-ip\cdot x}}{(2\pi)^{n/2}} g(k)\, dk$$

(d) *Le due trasformazioni ristrette allo spazio invariante* $\mathcal{S}(\mathbb{R}^n)$ *sono isometriche rispetto al semi prodotto scalare di* $\mathcal{L}^2(\mathbb{R}^n, dx^n)$*: se* $f_1, f_2, g_1.g_2 \in \mathcal{S}(\mathbb{R}^n)$*,*

$$\int \overline{(\mathcal{F}f_1)(k)}(\mathcal{F}f_2)(k)dk = \int \overline{f_1(x)} f_2(x)dx$$

e

$$\int \overline{(\mathcal{F}_- g_1)(x)}(\mathcal{F}_- g_2)(x)dx = \int \overline{g_1(k)} g_2(k)dk\,.$$

Prova. (a) Diamo la prova per $\mathcal{F}$, per $\mathcal{F}_-$ si pò procedere analogamente. Poniamo

$$g(k) := \int_{\mathbb{R}^n} \frac{e^{ip\cdot x}}{(2\pi)^{n/2}} f(x)\, dx$$

Si verifica facilmente che il secondo membro può essere derivato in k tramite un operatore ∂_k^α facendo passare la derivata sotto il segno di integrale. Infatti risulta immediatamente che

$$|\partial_k^\alpha e^{ik\cdot x} f(x)| = \left| i^{|\alpha|} M^\alpha(x) e^{ik\cdot x} f(x)| \le |M^\alpha(x) f(x) \right|\,.$$

La funzione $x \mapsto |M^\alpha(x)f(x)|$ è in L^1 in quanto $f \in \mathcal{S}(\mathbb{R}^n)$. Dato che il modulo della derivata dell'integrando è maggiorata uniformemente da una funzione positiva ed integrabile, noti teoremi sul passaggio del simbolo di derivata sotto quello di integrale (basati sul teorema della convergenza dominata di Lebesgue) permettono di concludere che

$$\partial_k^\alpha g(k) := i^{|\alpha|} \int_{\mathbb{R}^n} \frac{e^{ip\cdot x}}{(2\pi)^{n/2}} M^\alpha(x) f(x)\, dx$$

Usando la derivazione per parti e notando che f con tutte le sue derivate si annulla più rapidamente di ogni potensa inversa di $|x|$ per $|x| \to +\infty$ si trova che:

$$M^\beta(k)\partial_k^\alpha g(k) = i^{|\alpha|}(-i)^{|\beta|} \int_{\mathbb{R}^n} \frac{e^{ip\cdot x}}{(2\pi)^{n/2}} M^\alpha(x) \partial_x^\beta f(x)\, dx \tag{3.50}$$

Quindi da (a), per ogni $k \in \mathbb{R}^n$

$$|M^\beta(k)\partial_k^\alpha g(k)| \le \left|\left| M^\alpha \partial^\beta f \right|\right|_\infty\,.$$

esendo finito il secondo membro in quanto $f \in \mathcal{S}(\mathbb{R}^n)$ ed essendo α e β arbitrari, concludiamo che $g \in \mathcal{S}(\mathbb{R}^n)$.
(b) L'identità in (3.50) si scrive in altro modo

$$M^\alpha \partial^\beta\, \mathcal{F} = (-1)^{|\beta|} i^{|\alpha|+|\beta|}\, \mathcal{F} M^\alpha \partial^\beta\,,$$

dove $\mathcal{F}$ è in realtà la restrizione della trasformata di Fourier a $\mathcal{S}(\mathbb{R}^n)$. In modo analogo si trova che

$$M^\alpha \partial^\beta \, \mathcal{F}_- = (-1)^{|\alpha|} i^{|\alpha|+|\beta|} \, \mathcal{F}_- M^\alpha \partial^\beta \, .$$

Da queste identità troviamo in particolare che

$$\begin{aligned} \mathcal{F}\mathcal{F}_- \, M^\alpha &= M^\alpha \mathcal{F}\mathcal{F}_- \, , & (3.51) \\ \mathcal{F}_-\mathcal{F} \, M^\alpha &= M^\alpha \mathcal{F}_-\mathcal{F} \, , & (3.52) \end{aligned}$$

dove M^α è pensato come *operatore moltiplicativo* $(M^\alpha f)(x) := M^\alpha(x) f(x)$, e anche

$$\begin{aligned} \mathcal{F}\mathcal{F}_- \, \partial^\alpha &= \partial^\alpha \mathcal{F}\mathcal{F}_- \, , & (3.53) \\ \mathcal{F}_-\mathcal{F} \, \partial^\alpha &= \partial^\alpha \mathcal{F}_-\mathcal{F} \, . & (3.54) \end{aligned}$$

Mostreremo ora che, in virtù di tali relazioni di commutazione, gli operatori $J := \mathcal{F}\mathcal{F}_-$ e $J_- := \mathcal{F}\mathcal{F}_-$ devono essere l'operatore identità di $\mathcal{S}(\mathbb{R}^n)$. Per prima cosa proviamo che, fissato $x_0 \in \mathbb{R}^n$, e $f \in \mathcal{S}(\mathbb{R}^n)$ il valore di $(Jf)(x_0)$ dipende solo da $f(x_0)$. Se $f \in \mathcal{S}(\mathbb{R}^n)$ possiamo sempre scrivere

$$f(x) = f(x_0) + \int_0^1 \frac{df(x_0 + t(x - x_0))}{dt} dt = f(x_0) + \sum_{i=1}^n (x_i - x_{0i}) g_i(x) \, ,$$

dove le funzioni g_i (che sono $C^\infty(\mathbb{R}^n)$ come si verifica facilmente) sono definite da

$$g_i(x) := \frac{\partial}{\partial x_i} \int_0^1 f(x_0 + t(x - x_0)) dt$$

e dunque e $f_1, f_2 \in \mathcal{S}(\mathbb{R}^n)$ e $f_1(x_0) = f_2(x_0)$, risulta

$$f_1(x) = f_2(x) + \sum_{i=1}^n (x_i - x_{0i}) h_i(x) \, , \tag{3.55}$$

dove, per differenza, la funzione $x \mapsto \sum_{i=1}^n (x_i - x_{0i}) h_i(x)$ e quindi le funzioni h_i sono in $\mathcal{S}(\mathbb{R}^n)$. Applicando J ad ambo membri di (3.55), tenendo conto del fatto che J commuta con i polinomi in x per (3.51) otteniamo

$$(Jf_1)(x) = (Jf_2)(x) + \sum_{i=1}^n (x_i - x_{0i})(Jh_i)(x) \, .$$

Prendendo $x = x_0$ si vede che $(Jf_1)(x_0) = (Jf_2)(x_0)$ sotto l'ipotesi iniziale di $f_1(x_0) = f_2(x_0)$. Quindi come dicevamo, $(Jf)(x_0)$ è una funzione di $f(x_0)$ unicamente. Tale funzione deve anche essere lineare dato che J è lineare per costruzione. Ne consegue che sarà $(Jf)(x_0) = j(x_0) f(x_0)$. Per l'arbitrarietà di x_0, abbiamo provato che J agisce come la moltiplicazione per una funzione j. Si osservi che tale funzione deve essere C^∞. Per provare ciò si scelga $f \in \mathcal{S}(\mathbb{R}^n)$ che valga costantemente 1 in un intorno $I(x_0)$ di x_0. Se $x \in I(x_0)$ vale $(Jf)(x) = j(x)$. Dato che il primo

membro è $C^\infty(I(x_0))$ lo deve essere anche il secondo. Valendo ciò nell'intorno di ogni punto di $\mathbb{R}^n$ vale anche $j \in C^\infty(\mathbb{R}^n)$. (3.53) implica allora che, per ogni $f \in \mathcal{S}(\mathbb{R}^3)$ ed ogni $x \in \mathbb{R}^3$:

$$j(x)\frac{\partial}{\partial x^i} f(x) = \frac{\partial}{\partial x^i} j(x) f(x) \,.$$

Scegliendo come prima f con valore costantemente uguale a 1 in un intorno, si ha dall'identità di sopra che in quell'intorno tutte le derivate di j devono annullarsi. Dato che ciò vale nell'intorno di ogni punto di $\mathbb{R}^n$ che è connesso, la funzione continua j deve essere una funzione costante su tutto $\mathbb{R}^n$. Il valore della costante ovviamente non dipende dall'argomento di J e quindi può essere calcolato valutando J su una funzione arbitraria di $\mathcal{S}(\mathbb{R}^n)$. Un utile esercizio è quello di calcolare J sulla funzione $x \mapsto e^{-x^2}$ e si vede che il valore della costante è proprio 1. La prova per J_- si eegue nello stesso modo.
(d) Usando (c) la tesi si prova immediatamente. Facciamo la dimostrazione per $\mathcal{F}$, quella per $\mathcal{F}_-$ è essenzialmente la stessa. Se $f_1, f_2 \in \mathcal{S}(\mathbb{R}^n)$, poniamo per $i = 1, 2$

$$g_i(k) = \int_{\mathbb{R}^n} \frac{e^{ip\cdot x}}{(2\pi)^{n/2}} f_i(x)\, dx \,.$$

È immediato verificare che nelle nostre ipotesi posiamo usare il teorema di Fubini-Tonelli ottenendo

$$\int \overline{g_1(k)} g_2(k) dk = \int \overline{g_1(k)} \int \frac{e^{ik\cdot x}}{(2\pi)^{n/2}} f_2(x) dx dk = \int\int \frac{e^{ik\cdot x}}{(2\pi)^{n/2}} \overline{g_1(k)} f_2(x) dx \otimes dk \,.$$

Usando nuovamente il teorema di Fubini-Tonelli per l'ultimo integrale opportunamente riscritto abbiamo

$$\int \overline{g_1(k)} g_2(k) dk = \int\int \overline{\frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} g_1(k)} f_2(x) dx \otimes dk = \int f_2(x) \int \overline{\frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} g_1(k)} dk dx = \int \overline{f_1(x)} f_2(x) dx,$$

dove abbiamo usato il risultato in (c). Queta è la tesi che volevamo provare. □

Passiamo alla trasformata di Fourier-Plancherel. Dato che $\mathcal{S}(\mathbb{R}^n)$ è denso in $\mathcal{L}^2(\mathbb{R}^n)$, passando alle classi di equivalenza di funzioni i potrà dire che $\mathcal{S}(\mathbb{R}^n)$ individua un sottopazio denso, che indicheremo ancora con lo stesso simbolo $\mathcal{S}(\mathbb{R}^n)$, nello spazio di Hilbert $L^2(\mathbb{R}^n)$. Gli operatori $\mathcal{F}$ e $\mathcal{F}_-$ possono essere pensati come definiti su tale sottospazio denso in $L^2(\mathbb{R}, dx)$. Il punto (d) della proposizione 3.12 implica in particolare che tali operatori sono limitati con norma pari a 1 dato che sono operatori isometrici. Come sappiamo dall'esercizio (3) di esercizi 2.1, $\mathcal{F}$ e $\mathcal{F}_-$ individueranno univocamente due operatori lineari limitati su tutto $L^2(\mathbb{R}^n, dx)$. Per esempio, l'operatore che estende $\mathcal{F}$ su $L^2(\mathbb{R}^n, dx)$ è definito come, e $f \in L^2(\mathbb{R}^n, dx)$

$$\widehat{\mathcal{F}} f := \lim_{n\to+\infty} \mathcal{F} f_n$$

dove $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ è una qualunque successione che converge a f nella topologia di $L^2(\mathbb{R}^n, dx)$ (come provato nell'esercizio citato sopra, il limite esiste sempre e non dipende dalla successione

scelta). Per la continuità del prodotto scalare, l'operatore esteso $\widehat{\mathcal{F}}$ continuerà a conservare il prodotto scalare di $L^2(\mathbb{R}^2, dx)$ e pertanto $\widehat{\mathcal{F}}$ sarà iniettivo su $L^2(\mathbb{R}^n, dx)$. In realtà sarà anche suriettivo per il seguente motivo elementare. Accanto all'operatore $\widehat{\mathcal{F}}$ possiamo similmente costruire l'operatore che estende su $L^2(\mathbb{R}^n, dx)$ la trasformata inversa di Fourier $\widehat{\mathcal{F}_-}$. Su $\mathcal{S}(\mathbb{R}^n, dx)$ vale

$$\mathcal{F}\mathcal{F}_- = I_{\mathcal{S}(\mathbb{R}^n)}$$

Passando alle estensioni su L^2 per linearità e continuità, tenendo conto del fatto che l'unica estensione lineare dell'identità su $\mathcal{S}(\mathbb{R}^n, dx)$ (costruita con la procedura generale detta sopra) è l'identità su $L^2(\mathbb{R}^n, dx)$, otteniamo che deve valere

$$\widehat{\mathcal{F}}\widehat{\mathcal{F}}_- = I\ ,$$

dove I è l'identità su $L^2(\mathbb{R}^n, dx)$. Questo fatto implica immediatamente la suriettività di $\widehat{\mathcal{F}}$.

Definizione 3.15 *L'unico operatore $\widehat{\mathcal{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dx)$ che estende per linearità e continuità la trasformata di Fourier ristretta a $\mathcal{S}(\mathbb{R}^n)$ è detto trasformazione di* **Fourier-Plancherel**.

Teorema 3.10 (Plancherel). *La trasformata di Fourier-Plancherel $\widehat{\mathcal{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dx)$ è un operatore lineare biettivo ed isometrico.*

Prova. La dimostrazione è stata data immediatamente sopra la definizione 3.15. □

Rimane aperta una questione. Se $f \in L^1(\mathbb{R}^n, dx) \cap L^2(\mathbb{R}^n, dx)$ a priori $\mathcal{F}f$ e $\widehat{\mathcal{F}}f$ sono diversi perchè per definire $\widehat{\mathcal{F}}$ non abbiamo esteso $\mathcal{F}$ partendo da $L^1(\mathbb{R}^n, dx)$, ma siamo partiti dal suo sottopazio $\mathcal{S}(\mathbb{R}^n)$. Questa era l'unica possibilità visto che $L^1(\mathbb{R}^n, dx) \not\subset L^2(\mathbb{R}^n, dx)$.
La seguente propoizione fa luce sulla questione dando anche un metodo pratico per calcolare la traformata di Fourier-Plancherel in termini di limiti di quella di Fourier.

Nota. Ricordiamo che se $K \subset \mathbb{R}^n$ è un insieme di misura finita ed in particolare se K è un compatto (i compatti hanno misura di Lebesgue finita), allora:
(1) $L^2(K, dx) \subset L^1(K, dx)$;
(2) se $\{f_n\}_{n\in\mathbb{N}} \subset L^2(K, dx)$ converge nella norma $|| \, ||_2$ a $f \in L^2(\mathbb{R}^n, dx)$ allora converge alla stessa funzione anche nella norma $|| \, ||_1$;
(3) $L^\infty(K, dx) \subset L^p(K, dx)$ per $1 \leq p < \infty$;
(4) se $\{f_n\}_{n\in\mathbb{N}} \subset L^\infty(K, dx)$ converge nella norma $|| \, ||_\infty$ a $f \in L^\infty(\mathbb{R}^n, dx)$ allora converge alla stessa funzione anche nella norma $|| \, ||_p$.
Queste quattro affermazioni si provano immediatamente come segue tenendo conto, per le prime due, che la funzione che vale costantemente 1 su un compatto, che quindi ha misura finita, è integrabile. Per la prima affermazione si osservi che vale

$$2|f(x)| \leq |f(x)|^2 + 1\ .$$

Per cui l'integrale (eventualmente infinito) del primo membro è maggiorato da quello del secondo. Per la seconda affermazione si noti che, dalla disuguaglianza di Cauchy-Schwartz

$$\left(\int_K |g(x)|\, 1\, dx\right)^2 \leq \left(\int_K |g(x)|^2 dx\right)\left(\int_K 1 dx\right) .$$

sostituendo $f(x) - f_n(x)$ a $g(x)$ si ha la prova della seconda affermazione. Riguardo alla terza e quarta affermazione è sufficiente notare che si osservi che, per la stessa definizione di integrale di Lebesgue:

$$\int_K |g|^p dx \leq \operatorname{ess\,sup}_K |f|^p \int_K dx = (||g||_\infty)^p \int_K dx$$

per ogni funzione misurabile g definita su K.

Proposizione 3.13. *La trasformazione di Fourier-Plancherel gode delle due seguenti proprietà.*
(a) *Se $f \in L^2(\mathbb{R}^n, dx) \cap L^1(\mathbb{R}^n, dx)$ la sua trasformata di Fourier-Plancherel si riduce alla trasformata di Fourier $\mathcal{F}f$ calcolata direttamente con la formula integrale (3.48).*
(b) *Se $f \in L^2(\mathbb{R}^n, dx)$ la sua trasformata di Fourier-Plancherel può essere calcolata come*

$$\hat{\mathcal{F}}f = \lim_{n\to\infty} g_n$$

dove il limite è nel senso di $L^2(\mathbb{R}^n, dk)$,

$$g_n(k) := \int_{K_n} \frac{e^{ik\cdot x}}{(2\pi)^{n/2}} f(x) dx \,, \tag{3.56}$$

e gli insiemi $K_n \subset \mathbb{R}^n$ sono tutti compatti e $K_{m+1} \supset K_m$ per $m = 1, 2, \ldots$ ed infine vale $\cup_{m=1}^\infty K_m = \mathbb{R}^n$.

Prova. (a) Cominciamo a provare la tesi per le funzioni $f \in L^2(\mathbb{R}^n, dx)$ tali che f sia differente da 0 su un insieme di misura nulla fuori da un compatto K_0. In questo caso f appartiene anche a $L^1(\mathbb{R}^n, dx)$. Sia quindi $\{s_n\} \subset \mathcal{S}(\mathbb{R}^n)$ una successione di funzioni che tendono a f nella norma di $L^2(\mathbb{R}^n, dx)$. Se B e B' sono due palle aperte di raggio finito tali che $B \supset \overline{B'} \supset B' \supset K_0$, possiamo costruire una funzione $h \in \mathcal{D}(\mathbb{R}^n)$ tale che vale costantemente 1 su B' e si annulla fuori da B. È chiaro che se $f_n := h \cdot s_n$, la successione $\{f_n\}$ è costituita da funzione di $\mathcal{D}(\mathbb{R}^n)$, e quindi $\mathcal{S}(\mathbb{R}^n)$ con supporti contenuti nel compatto $K := \overline{B}$. Quindi tutte le f_n sono anche in $L^1(\mathbb{R}^n, dx)$ e la successione $\{f_n\}$ tende a f sia norma $L^2(\mathbb{R}^n, dx)$ che quella di $L^1(\mathbb{R}^n, dx)$.
Per definizione valendo $f_n \to f$ nella norma $|| \, ||_2$ sarà

$$||\mathcal{F}f_n - \hat{\mathcal{F}}f||_2 \to 0 \tag{3.57}$$

per $n \to +\infty$. D'altra parte, dato che $f_n \to f$ anche nella norma $||\,||_1$ per (a) di proposizione 3.12, allora $||\mathcal{F}f_n - \mathcal{F}f||_\infty \to 0$ per $n \to \infty$. Dato che, sugli insiemi di misura finita, la convergenza nella norma $|| \; ||_\infty$ implica quella nella norma $|| \; ||_2$ avremo che

$$||\mathcal{F}f_n - \mathcal{F}f||_2 \to 0 \tag{3.58}$$

e quindi, valendo (3.57) e per l'unicità del limite dovrà essere $\widehat{\mathcal{F}}f = \mathcal{F}f$.
Supponiamo ora che $f \in L^2(\mathbb{R}^n, dx) \cap L^1(\mathbb{R}^n, dx)$ senza altre restrizoni. Consideriamo una classe di compatti $\{K_n\}$ che soddisfino le ipotesi nel punto (b). Definiamo le funzione $f_n := \chi_{K_n} \cdot f$ dove χ_E è la funzione caratteristica dell'insieme E (definita in mod che $\chi_E(x) = 0$ se $x \notin E$ e $\chi_E(x) = 1$ se $x \in E$). È chiaro che vale $f_n \to f$ puntualmente per $n \to +\infty$. Inoltre $|f(x) - f(x)_n|^p \leq |f(x)|^p$ per $p = 1, 2, \ldots$. Allora, per il teorema della convergenza dominata di Lebesgue, $f_n \to f$ per $n \to +\infty$ in particolare rispetto alla norma $|| \; ||_1$ e a quella $|| \; ||_2$. D'altra parte, per quanto dimostrato sopra,

$$\mathcal{F}f_n = \widehat{\mathcal{F}}f_n \, .$$

Facendo il limite per $n \to +\infty$ troviamo che $||\mathcal{F}f - \mathcal{F}f_n||_\infty \to 0$ ed allo stesso tempo vale anche $||\widehat{\mathcal{F}}f - \mathcal{F}f_n||_2 \to 0$. Questi fatti varranno anche per le funzioni che si ottengono restringendo $\widehat{\mathcal{F}}f$, $\mathcal{F}f$, $\mathcal{F}f_n$ ad un qualsiasi compatto K. Tenendo conto che per funzioni nulle fuori da un compatto la convergenza uniforme implica quella in L^2, si ha che se x appartiene ad un compatto qualsiasi $(\mathcal{F}f)(x) = \widehat{\mathcal{F}}f)(x)$ quasi ovunque. Ma ogni $x \in \mathbb{R}^n$ appartiene ad un compatto per cui $\mathcal{F}f = \widehat{\mathcal{F}}f$ come elementi di $L^2(\mathbb{R}^n, dx)$.
(b) La prova è stata data nelal parte finale della dimostrazione del punto (a). $\square$

Esempi 3.3
(1) Un fatto piuttosto importante che distingue nettamente lo spazio $\mathcal{D}(\mathbb{R}^n)$ da $\mathcal{S}(\mathbb{R}^n)$ è che il primo spazio, al contrario del secondo, *non* è invariante sotto la trasformazione di Fourier (e la trasformata inversa di Fourier). Infatti vale il eguente risultato:
se $f \in \mathcal{D}(\mathbb{R}^n)$ allora $\mathcal{F}f \in \mathcal{S}(\mathbb{R}^n)$ ma, se $\mathcal{F}f \in \mathcal{D}(\mathbb{R}^n)$ allora f e $\mathcal{F}f$ sono la funzione nulla. Lo stesso fatto vale per la trasformata inversa di Fourier.
La prova è semplice, la svolgiamo per $\mathcal{F}$, ma la stessa prova vale per $\mathcal{F}_-$. e vale

$$g(k) = \int_{\mathbb{R}^n} \frac{e^{ip \cdot x}}{(2\pi)^{n/2}} f(x) \, dx \, ,$$

dove f ha supporto compatto allora, l'integrale di sopra converge anche per $k \in \mathbb{C}^n$. Inoltre facendo uso del teorema della convergenza dominata di Lebesgue, si può passare la derivata nelle componenti di k, e nelle loro parti reale ed immaginaria, sotto il segno di integrale. Dato che $k \mapsto e^{ik \cdot x}$ è analitica (cioè analitica in ogni variabile separatamente), soddisferà le condizioini di Cauchy-Riemann in ogni variabile. Di conseguenza tali condizioni saranno soddisaftte anche per la funzione g in ogni variabile k_i. Si conclude che g è una funzione analitica su tutto $\mathbb{C}^n$. La restrizione di g ad $\mathbb{R}^n$ definirà, tramite la parte reale e quella immaginaria, due funzioni analitiche di variabile reale definite in tutto $\mathbb{R}^n$. Se g ha supporto compatto esisterà un aperto non vuoto di $\mathbb{R}^n$ in cui $Re\, g$ e $Im\, g$ si annullano. È un fatto ben noto nella teoria delle funzioni analitiche (reali di variabile reale) che se una funzione analitica definita su un aperto connesso (in questo caso tutto $\mathbb{R}^n$) si annulla su un aperto non vuoto incluso nel suo dominio allora si annulla ovunque nel dominio. Dobbiamo quindi concludere che, nelle nostre ipotesi su f, se g ha supporto compatto deve necessariamente essere la funzione nulla. Tale dovrà anche essere f in quanto $\mathcal{F}$ è invertibile su $\mathcal{S}(\mathbb{R}^n)$

(2) Un fatto connesso con (1) è il noto *teorema di Paley-Wiener* (vedi per es. [17]):
Sia $a > 0$ e si consideri $L^2([-a,a],dx)$ come sottospazio di $L^2(\mathbb{R},dx)$. Lo spazio $\hat{\mathcal{F}}(L^2([-a,a],dx))$ contiene tutte e sole le funzioni $g = g(k)$ tali che ciascuna di esse possa essere estesa analiticamente ed univocamente a tutto il piano complesso nella variabile $k \in \mathbb{C}$ in una funzione analitica che soddisfa

$$|g(k)| \leq Ce^{2\pi a|Imk|}\,, \quad \textit{per ogni } k \in \mathbb{C}$$

per qualche costante $C \geq 0$ dipendente da g.

Visto che deve essere $\hat{\mathcal{F}}(L^2([-a,a],dx)) \subset L^2(\mathbb{R},dk)$ per il teorema di Plancherel, il teorema di Paley-Wiener implica che le funzioni analitiche g che sono limitate secondo la disuguaglianza di sopra per qualche cotante $C \geq 0$, individuano elementi di $L^2(\mathbb{R},dk)$ quando k è ristretta a $\mathbb{R}$.
(3) Consideriamo lo spazio $L^2((a,b),dx)$ dove $-\infty \leq a < b \leq +\infty$ e dx denota l'usuale misura di Lebesgue du $\mathbb{R}$. Sussiste la seguente utilissima proposizione usata in (4) di esempi 3.2:
Sia $f : (a,b) \to \mathbb{C}$ misurabile tale che (1) l'insieme $\{x \in (a,b) \mid f(x) = 0\}$ ha misura nulla, (2) esistono $C, \delta > 0$ per cui $|f(x)| < Ce^{-\delta|x|}$ per ogni $x \in (a,b)$.
In questo caso lo spazio lineare finitamente generato da tutte le funzioni $x \mapsto x^n f(x) = f_n(x)$ per $n = 0,1,2,\ldots$ è denso in $L^2((a,b),dx)$.
Dimostriamo questa affermazione. Sia $S := \{f_n\}_{n\in\mathbb{N}}$. È sufficiente provare che $S^\perp = \{0\}$ dato che è $S^\perp \oplus \overline{< S >} = L^2((a,b),dx)$ per il teorema 3.1. Sia allora $h \in L^2((a,b),dx)$ tale che

$$\int_a^b x^n f(x)\overline{h(x)}dx = 0$$

per ogni $n = 0,1,2,\ldots$. Possiamo estendere la funzione h a tutta la retta reale definendola nulla fuori da (a,b). La condizione di sopra si riscrive

$$\int_{\mathbb{R}} x^n f(x)\overline{h(x)}dx = 0\,, \tag{3.59}$$

per ogni $n = 0,1,2,\ldots$. Valgono inoltre i seguenti fatti:
(i) $f \cdot \overline{h} \in L^1(\mathbb{R},dx)$ dato che entrambe le funzioni sono in $L^2(\mathbb{R},dx)$ per cui il loro prodotto è in $L^1(\mathbb{R},dx)$;
(ii) $f \cdot \overline{h} \in L^2(\mathbb{R},dx)$ perchè $|f(x)|^2 < C^2e^{-2\delta|x|} < C^2 < +\infty$ e $|h|^2$ è integrabile per ipotesi;
(iii) La funzione che manda $x \in \mathbb{R}$ in $e^{\delta'|x|}f(x)\overline{h}(x)$ è in $L^1(\mathbb{R},dx)$ per ogni $\delta' < \delta$. Infatti, essendo $x \mapsto |e^{\delta'|x|}f(x)| \leq Ce^{-(\delta-\delta')|x|}$ la funzione $x \mapsto e^{\delta'|x|}f(x)$ è in $L^2(\mathbb{R},dx)$ inoltre $h \in L^2(\mathbb{R},dx)$ per ipotesi, per cui il loro prodotto è in $L^1(\mathbb{R},dx)$.
Per (i) possiamo calcolare la trasformata di Fourier

$$g(k) = \int_{\mathbb{R}} \frac{e^{ik\cdot x}}{\sqrt{2\pi}}\, f(x)\overline{h}(x)\, dx\,.$$

Questa coinciderà con la trasformata di Fourier-Plancherel di $f \cdot \overline{h}$ per il punti (i) e (ii) insieme tenuto conto di (a) della proposizione 3.13. Per (iii), se k è complesso e $|Imk| < \delta$, la funzione

$g = g(k)$ di sopra è ben definita ed analitica nella banda aperta $B \subset \mathbb{C}$ definita da $Rek \in \mathbb{R}$, $|Imk| < \delta$, la prova si costruisce analogamente a quanto fatto nell'esempio (1). Usando il teorema della convergenza dominata di Lebesgue e passando sotto il segno di integrale l'operatore di derivazione, è infine facile provare che, per ogni $n = 0, 1, \ldots$

$$\frac{d^n g}{dk^n}|_{k=0} = \frac{i^n}{\sqrt{2\pi}} \int_{\mathbb{R}} x^n f(x) \overline{h(x)} dx \, .$$

Tutte queste derivate sono nulle per (3.59), di conseguenza lo sviluppo di Taylor di g nell'origine è nullo e quindi g si annulla in un disco aperto contenuto nella banda B. Essendo h analitica, si annullerà identicamente nell'aperto conesso B ed in particolare sull'asse k reale. Quindi, in particolare, avremo che la trasformata di Fourier-Plancherel della funzione $f \cdot \overline{h}$ è il vettore nullo di $L^2(\mathbb{R}, dk)$. Dato che la trasformata di Pourier-Plancherel è un isomorfismo isometrico dobbiamo concludere che $f \cdot \overline{h} = 0$ quasi ovunque su $\mathbb{R}$ ed in particolare su (a, b) dove per ipotesi l'insieme dei punti su cui f *non* è nulla ha misura strettamente positiva. Ma allora deve essere $h = 0$ quasi ovunque su (a, b), ovvero ogni $h \in S^{\perp}$ coincide con l'elemento nullo di $L^2((a, b), dx)$. Questo conclude la prova.

Capitolo 4

Proprietà elementari degli operatori compatti, di Hilbert-Schmidt e di classe traccia nello spazio di Hilbert.

In questo capitolo introdurremo alcuni tipi di operatori che saranno necessari quando definiremo il concetto di stato quantistico. Tali operatori sono noti in letteratura come *operatori di classe traccia* oppure *operatori nucleari*. Per arrivare a definirli è necessario introdurre qualche nozione in più riguardante in particolare gli *operatori compatti*, detti anche *completamente continui*.

4.1 Generalità sugli operatori compatti in spazi normati, di Banach e Hilbert.

Ricordiamo che in uno spazio topologico un insieme *compatto* Kè definito dalla proprietà seguente:
K è detto **compatto** *se da ogni ricoprimento di aperti di di K (cioè una classe di aperti $\{A_i\}_{i\in I}$ con $\cup_{i\in I} A_i \supset K$) è possibile estrarre un sotto ricoprimento finito,(ossia esiste un isieme finito $J \subset I$ tale che $\cup_{i\in J} A_i \supset K$).*
Una nozione connessa è quella di *compattezza sequenziale*:
K è detto **sequenzialmente compatto** *se per ogni successione $\{x_n\}_{n\in\mathbb{N}} \subset K$ esiste una sottosuccessione $\{x_{n_p}\}_{p\in\mathbb{N}}$ convergente ad un punto di K.*
Ricordiamo che i sottoinsiemi chiusi dei compatti sono compatti e, negli spazi di Hausdorff (in particolare gli spazi vettoriali normati come gli spazi di Hilbert) i compatti sono chiusi. Negli spazi topologici che siano anche spazi metrici (e quindi ancora in particolare gli spazi vettoriali normati come gli spazi di Hilbert), la proprietà di compattezza è *equivalente* a quella di compattezza sequenziale. Un'altra proprietà utile che vale in spazi metrici è la seguente.

Proposizione 4.1. *Sia X spazio metrico. Se $A \subset X$ è tale che ogni successione di punti di A*

ammette una sottosuccessioone convergente (non necessariamente in A), allora $\overline{A}$ è compatto.

Prova. L'unica cosa da mostrare è che se $\{y_n\} \subset \overline{A} \setminus A$ allora esiste una sottosuccessione che converge (in $\overline{A}$ essendo esso chiuso). Per ipotesi ci saranno delle successioni, una per ogni k, $\{x_n^{(k)}\} \subset A$ con $x_n^{(k)} \to y_k$ per $n \to \infty$. Fissando k e prendendo un corrispondente n_k sufficientemente grande, possiamo allora costruire termine per termine una nuova successione $\{z_k := x_{n_k}^{(k)}\} \subset A$ tale che $||y_k - z_k|| < 1/k$. Nelle ipotesi fatte su A, esisterà una sottosuccessione di $\{z_k\}$, $\{z_{k_p}\}$ che converge a qualche $y \in \overline{A}$.

$$||y_{k_p} - y|| \leq ||y_{k_p} - z_{k_p}|| + ||z_{k_p} - y||$$

Dato che $1/k_p \to 0$ per $p \to \infty$, fissato $\epsilon > 0$ esisterà P tale che, se $p > P$ allora $||z_{k_p} - y|| < \epsilon/2$ insieme a $1/k_p < \epsilon/2$, per cui $||y_{k_p} - y|| < \epsilon$. In altre parole $\{y_{k_p}\} \to y$ per $p \to \infty$. □

Un fatto importante e notevole, che distingue nettamente il caso di uno spazio normato finito dimensionale da quello infinito dimensionale è il seguente che si applica più in generale a spazi normati.

Proposizione 4.2. *Sia $(X, || \, ||)$ uno spazio normato.*
(a) *Se X ha dimensione infinita la palla chiusa unitaria $\{x \in X \mid ||x|| \leq 1\}$ non può essere compatta.*
(b) *Se X ha dimensione infinita la chiusura della palla aperta unitaria $\{x \in X \mid ||x|| < 1\}$ non può essere compatta.*

Note.
(1) Nel caso di dimensione infinita, non è detto che la chiusura della palla aperta coincida con la palla chiusa! In ogni caso $\overline{\{x \in X \mid ||x|| < 1\}} \subset \{x \in X \mid ||x|| \leq 1\}$. La palla chiusa è un insieme chiuso perchè controimmagine di un chiuso secondo una funzione continua.
(2) Nel caso di dimensione finita la chiusura della palla aperta coincide con la palla chiusa ed entrambe sono compatte. Ciò segue dal fatto che in spazi normati di dimensione finita tutte le metriche inducono la stessa topologia (e rendono lo spazio di Banach, lo si dimostri per esercizio) per cui ci si può sempre ridurre ad usare la metrica euclidea identificando lo spazio con un $\mathbb{R}^n$. In tal caso la chiusura della palla aperta coincide banalmente con la palla chiusa che è compatta perchè insieme chiuso e limitato di $\mathbb{R}^n$ (teorema di Heine-Borel).

Prova della proposizione 4.2. Facciamo la dimostrazione nel caso (a), l'altro caso si dimostra analogamente. Indichiamo con B la palla aperta e supponiamo che $\overline{B}$ sia compatta. Allora possiamo ricoprire $\overline{B}$, e quindi B, con un numero $N > 0$ di sfere aperte B_k, tutte di raggio $1/2$ e centrate rispettivamente in x_k, $k = 1, \ldots, n$. Consideriamo un sottospazio X_n di X di dimensione finita n tale che contenga i vettori x_k. Dato che la dimensione di X è infinita, possiamo scegliere $n > N$ arbitrariamente grande definendolo come lo spazio generato dai vettori x_k unitamente ad altri $n - N$ vettori x_k arbitrari in modo tale che l'insieme di vettori sia linearmente l indipendente. Definiamo le ulteriori sfere di raggi 1 e $1/2$ rispettivamente $P := B \cap X_n$

e $P_k := B_k \cap X_n$ per ogni $k = 1, \ldots, N$. Sia m è la misura di Lebesgue su X_n identificato con $\mathbb{R}^n$ (se il campo di X è $\mathbb{R}$) o $\mathbb{R}^{2n}$ (se il campo di X è $\mathbb{C}$) tramite la base x_k, $k = 1, \ldots, n$. Se normalizziamo m in modo che $m(P) = 1$, allora $m(P_k) = (1/2)^n$. Dato che $B \subset \cup_{k=1}^N B_k$ e dunque $P \subset \cup_{k=1}^N P_k$, deve essere $1 \leq N(1/2)^n$. Questo è impossibile se n è abbastanza grande (N è fissato). $\square$

Ricordiamo che in uno spazio normato $(X, || \, ||)$ un insieme M è detto **limitato** (rispetto alla norma $|| \, ||$ se $sup\{||x|| \mid x \in M\} < \infty$).

Definizione 4.1. *Siano X, Y spazi normati sullo stesso campo $\mathbb{R}$ o $\mathbb{C}$ (in particolare $X = Y = H$ spazio di Hilbert). $T \in \mathfrak{L}(X, Y)$ è detto* **compatto** *o* **completamente continuo** *se vale una delle due condizioni equivalenti:*
(a) *per ogni sottoinsieme $M \subset X$ limitato, $\overline{T(M)}$ è compatto in Y;*
(b) *ogni successione $\{x_n\}_{n\in\mathbb{N}}$ ammette una sottosuccessione $\{x_{n_k}\}_{k\in\mathbb{N}} \subset X$ tale che $\{Tx_{n_k}\}_{k\in\mathbb{N}}$ converge in Y.*

Note.
(1) È chiaro che (a) $\Rightarrow$ (b). L'implicazione inversa, (b) $\Rightarrow$ (a), è immediata conseguenza della proposizione 4.1.
(2) Ogni operatore compatto è sicuramente limitato in quanto la palla chiusa di raggio 1 centrata nell'origine di X è mappata in un insieme a chiusura compatta (e quindi limitato perché ricopribile da un numero finito di palle di raggio finito), per cui in particolare $||T(x)|| \leq K < +\infty$ per $||x|| = 1$.

4.2 Operatori compatti in spazi di Hilbert.

D'ora in avanti ci restringeremo a considerare operatori compatti in spazi di Hilbert, anche se molte delle proprietà che enunceremo valgono anche in ambiente meno sofisticato come spazi normati o spazi di Banach. Nel teorema che segue l'ipotesi di completezza dello spazio viene in realtà usata solo nell'ultima proprietà.

Teorema 4.1. *Sia H spazio di Hilbert e $\mathfrak{B}_0(H) \subset \mathfrak{B}(H)$ il sottoinsieme degli operatori compatti. valgono i seguenti fatti.*
(a) *$\mathfrak{B}_0(H)$ è un sottospazio vettoriale di $\mathfrak{B}(H)$.*
(b) *$\mathfrak{B}_0(H)$ è uno $*$-ideale bilatero di $\mathfrak{B}(H)$, ossia, oltre ad essere un sottospazio, $\mathfrak{B}(H)$ gode della proprietà che se $T \in \mathfrak{B}_0(H)$, allora $TK, KT \in \mathfrak{B}_0(H)$ per ogni $K \in \mathfrak{B}(H)$ ed inoltre $T^\dagger \in \mathfrak{B}_0(H)$.*
(c) *$\mathfrak{B}_0(H)$ è chiuso rispetto alla topologia uniforme.*

Prova. (a) segue immediatamente da (b) della definizione 4.1 (b) Data la successione $\{x_n\}$

limitata esiste una sottosuccessione di $\{Kx_n\}$ convergente, e quindi anche di $\{TKx_n\}$ dato che T è continuo. Quindi TK è compatto. Data la successione $\{x_n\}$ limitata , anche $\{Tx_n\}$ dato che T è limitato ($||Tx_n|| \leq ||T||||x_n|| \leq ||T||C \leq \infty$). Dalla compattezza di K segue che $\{KTx_n\}$ ammette una sottosuccessione convergente. Quindi KT è compatto. Per dimostrare la chiusura rispetto all'operazione di coniugazione hermitiana notiamo che, $|T|$ è compatto se e solo se T è compatto come segue immediatamente da (3.41). Decomponendo polarmente $T = U|T|$ secondo il teorema 3.9, segue che $T^\dagger = |T|U^\dagger$ dove si è tenuto conto del fatto che $|T| \geq 0$ per cui $|T|$ è autoaggiunto. La limitatatezza di $U^\dagger$ insieme alla compattezza di $|T|$ implicano che $T^\dagger = |T|U^\dagger$ è compatto per quanto provato all'inizio di questa parte di dimostrazione.
(c) Sia $\mathfrak{B}(H) \ni A = lim_{n\to+\infty}A_n$ con $A_n \in \mathfrak{B}_0(H)$. Sia $\{x_n\}$ una successione limitata di vettori di H: $||x_n|| \leq C$ per ogni n. Vogliamo provare che esiste una sottosuccesione di $\{Ax_n\}$ convergente. Con una procedura ricorsiva, costruiamo una famiglia di sottosuccessioneo

$$\{x_n\} \supset \{x_n^{(1)}\} \supset \{x_n^{(2)}\} \supset \cdots \tag{4.1}$$

tale che, per ogni $i = 1, 2, \ldots$, $\{x_n^{(i+1)}\}$ è sottosuccessione di $\{x_n^{(i)}\}$ tale che $\{A_{i+1}x_n^{(i+1)}\}$ è convergente. Questo è sempre possibile in quanto ogni $\{x_n^{(i)}\}$ è limitata da C essendo sottosuccessione di $\{x_n\}$ ed inoltre A_{i+1} è compatto per ipotesi. La sottosuccessione di $\{Ax_n\}$ che proveremo convergere è quella di elementi $Ax_i^{(i)}$. Usando la disuguaglianza triangolare abbiamo facilmente che

$$||Ax_i^{(i)} - Ax_k^{(k)}|| \leq ||Ax_i^{(i)} - A_nx_k^{(k)}|| + ||A_nx_i^{(i)} - A_nx_k^{(k)}|| + ||A_nx_i^{(i)} - Ax_k^{(k)}|| \,.$$

Usando la stima scritta, nelle nostre ipotesi si ha che:

$$||Ax_i^{(i)} - Ax_k^{(k)}|| \leq ||A - A_n||(||x_i^{(i)}|| + ||x_k^{(k)}||) + ||A_n(x_i^{(i)} - x_k^{(k)})|| \leq 2C||A - A_n|| + ||A_nx_i^{(i)} - A_nx_k^{(k)}||.$$

Fissato $\epsilon > 0$, se n è grande a sufficienza varrà $2C||A - A_n|| \leq \epsilon/2$ dato che $A_n \to A$ per ipotesi. Fissato n e se $r \geq n$, $\{A_n(x_p^{(r)})\}_p$ è sottosuccessione della successione convergente $\{A_n(x_p^{(n)})\}_p$. È facile vedere che la successione $\{A_n(x_p^{(p)})\}_p$ costruita con i termini "diagonali" di tutte queste sottosuccessioni, *ciascuna sottosuccessione della precedente per (4.1)*, è ancora sottosuccessione della successione convergente $\{A_n(x_p^{(n)})\}_p$, per cui è anche essa convergente (allo stesso limite). Concludiamo che, se $i, k \geq n$ sono grandi a sufficienza vale anche $||A_nx_i^{(i)} - A_nx_k^{(k)}|| \leq \epsilon/2$. Quindi, se n è grande a sufficienza $||Ax_i^{(i)} - Ax_k^{(k)}|| \leq \epsilon/2 + \epsilon/2 = \epsilon$. Ciò prova la tesi. □

Esempi 4.1.
(1) Se X, Y sono spazi normati e $T \in \mathfrak{B}(X, Y)$ è tale che $Ran(T) < +\infty$, allora T deve essere compatto. Proviamolo. Sia $V \subset X$ è limitato, cioè $||x|| < K$ se $x \in V$ per un fissato $K < \infty$ indipendente da x, allora $||T(V)|| \leq K||T|| < +\infty$ per cui $T(V)$ è limitato. $\overline{T(V)}$ è chiuso e limitato in uno spazio normato di dimensione finita che è sempre omeomorfo ad un $\mathbb{R}^n$ (vedi (2) in Note dopo proposizione 4.2). Per il teorema di Heine-Borel $\overline{T(V)}$ è compatto. Quindi T è compatto.

Come ulteriore sottocaso, se H è spazio di Hilbert si consideri un operatore $T_x \in \mathfrak{L}(H)$ della forma

$$T_x : u \mapsto (x|u)x$$

dove $x \in H$ è un fissato vettore. È chiaro che tale operatoreè compatto: se $||u_n|| < C < \infty$ per $n \in \mathbb{N}$, $|(x|u_n)| \leq ||x||C$ per cui esiste una sottosuccessione di numeri $\{(x|u_{n_k})\}_k$ convergente e dunque esiste una sottosuccessione di $\{Tu_n\}$, $\{Tu_{n_k} = (x|u_{n_k})x\}_k$, convergente in H.
(2) Se $\{x_n\}_{n\in\mathbb{N}}$ e $\{y_n\}_{n\in\mathbb{N}}$ sono sottoinsiemi ortogonali di H e $\mathfrak{B}(H) \ni T = \sum_{n\in\mathbb{N}}(x_n|\ )y_n$ è un operatore limitato ottenuto interpretando la serie nella topologia uniforme, allora T è compatto per (a) e (c) del teorema 4.1.
(3) In $l^2(\mathbb{N})$ consideriamo l'operatore $A : \{x_n\} \mapsto \{x_{n+1}/n\}$. Tale operatore è compatto in quanto è limite nella topologia uniforme degli operatori, per $n = 1, 2 \ldots,$

$$A_n : \{x_n\} \mapsto \{x_2/2, x_3/3, \ldots, x_n/n, 0, 0, \ldots\} .$$

Si verifica infatti facilmente che (lo si provi per esercizio)

$$||A - A_n|| \leq 1/(n+1) .$$

(4) Si consideri lo spazio X misura μ sulla σ-algebra Σ in X e sia μ σ-finita in modo da definire la misura prodotto $\mu \otimes \mu$. Consideriamo quindi $K \in L^2(X \times X, \mu \otimes \mu)$. Si può provare che

$$T_K : L^2(\mu) \ni f \mapsto \int K(x,y)f(y)d\mu(y)$$

definisce un operatore $T_K \in \mathfrak{B}(L^2(X,\mu))$ che è anche compatto qunado μ è anche separabile. Per prima cosa notiamo che, indipendentemente dall'ipotesi di separabilità, se $f \in L^2(\mu)$:

$$\int K(\cdot, y)f(y)d\mu(y) \in L^2(\mu) .$$

e

$$\left|\left|\int K(\cdot, y)f(y)d\mu(y)\right|\right|_{L^2(X,\mu)} \leq ||K||_{L^2(X\times X,\mu\otimes\mu)}||f||_{L^2(X,\mu)} ,$$

che equivale a dire:

$$||T_K|| \leq ||K||_{L^2(X\times X,\mu\otimes\mu)} . \tag{4.2}$$

(La prova di questo fatto è interamente basata sul teorema di Fubini-Tonelli: Se $K \in L^2(\mu \otimes \mu)$ allora, per Fubini-Tonelli:
(1) $|K(x,\cdot)|^2 \in L^1(\mu)$, μ-quasi ovunque ,
(2) $\int |K(x,y)|^2 d\mu(y) \in L^1(\mu)$.
Da (1) segue che $K(x,\cdot) \in L^2(\mu)$ quasi ovunque e quindi $K(x,\cdot)f \in L^1(\mu)$ e per la disuguaglianza di Cauchy-Schwarz:
(3) $\int |K(x,y)||f(y)|d\mu(y) \leq ||K(x,\cdot)||_{L^2}||f||_{L^2}$.
Posto $F(x) := \int K(x,y)f(y)d\mu(y)$, F è misurabile ed è da (3):

(4) $|F(x)|^2 \leq ||f||^2_{L^2} \int |K(x,y)|^2 d\mu(y)$.
Per (2) si ha quindi che $|F|^2 \in L^2(\mu)$ ed è quindi vero che vale

$$\int K(\cdot, y) f(y) d\mu(y) \in L^2(\mu) \, .$$

Da (4) e dal teorema di Fubini-Tonelli segue infine anche

$$\left|\left| \int K(\cdot, y) f(y) d\mu(y) \right|\right|_{L^2(\mu)} \leq ||K||_{L^2(\mu\otimes\mu)} ||f||_{L^2(\mu)} \, . \,)$$

Facciamo l'ipotesi ulteriore che μ sia separabile in modo che sia separabile $L^2(X, \mu)$. Per esempio X può essere un qualsiasi intervallo (o un boreliano) di $\mathbb{R}$ e μ la misura di Lebesgue su $\mathbb{R}$. Se $\{u_n\}_{n\in\mathbb{N}}$ è una base Hilbertiana di $L^2(X, \mu)$ allora $\{u_n \cdot u_m\}_{n,m\in\mathbb{N}}$ è una base Hilbertiana di $L^2(X \times X, \mu \otimes \mu)$ ($\cdot$ è l'ordinario prodotto di funzioni punto per punto) e allora, nella topologia di $L^2(X \times X, \mu \otimes \mu)$,

$$K = \sum_{n,m} k_{nm} u_n \cdot u_m \, ,$$

dove i numeri $k_{nm} \in \mathbb{C}$ dipendono da K. Quindi posto

$$K_p := \sum_{n,m\leq p} k_{nm} u_n \cdot u_m$$

si ha che $K_p \to K$ per $p \to \infty$. Tenuto conto di (4.2) applicata agli operatori $T_{K_p - K} = T_{K_p} - T_K$, dove T_{K_p}è indotto dal nucleo integrale K_p, segue che, per $p \to \infty$

$$||T_K - T_{K_p}|| \to 0 \, ,$$

per cui T_K è compatto in quanto gli operatori T_{K_p} sono compatti essendo somme finite del tipo di quelli discussi nell'esempio (1).

Gli operatori compatti in spazi di Hilbert godono di notevoli proprietà, in particolare gli operatori compatti e autoaggiunti hanno caratteristiche che generalizzano al caso infinito dimensionale le proprietà della matrici hermitiane finito dimensionali. Oltre a ciò sono uno strumeto utilissimo per estendere ulteriormente la teoria.

Teorema 4.2. *Sia H spazio di Hilbert e $T \in \mathfrak{B}_0(H)$ con $T = T^\dagger$, allora vale quanto segue.*
(a) *L'insieme $\sigma_p(T)$ degli autovalori di T, ha le seguenti caratteristiche:*
(1) è non vuoto,
(2) è reale,
(3) è al più numerabile,
(4) ha al più un solo punto di accumulazione dato da 0,
(5) non contiene due autovalori distinti con lo stesso valore assoluto.

(6) Vale l'identità

$$\sup\{|\lambda| \mid \lambda \in \sigma_p(T)\} = ||T|| \,.$$

Più precisamente, l'estremo superiore suddetto è anche massimo ed è raggiunto in $\Lambda \in \sigma_p(T)$ *tale che*

$$\Lambda = ||T|| \quad se \quad \sup_{||x||=1}(x|Tx) = ||T|| \,, \tag{4.3}$$

oppure

$$\Lambda = -||T|| \quad se \quad \inf_{||x||=1}(x|Tx) = -||T|| \,. \tag{4.4}$$

(7) T *coincide con l'operatore nullo se e solo* 0 *è l'unico autovalore.*

(b) *Ogni autospazio di* T *con autovalore non nullo ha dimensione finita.*

Prova di parte del teorema 4.2. (a) Proveremo tutti i punti eccetto (3), (4) e (5) che saranno provati nella dimostrazione del teorema successivo. $\sigma_p(T)$ se è non vuoto è fatto di elementi reali per (c) di proposizione 3.8 essendo T autoaggiunto. Per (a) in proposizione 3.8 ci possono essere solo le seguenti due possibilità: $\sup_{||x||=1}(x|Ux) = ||T||$ oppure $\inf_{||x||=1}(x|Ux) = -||T||$. Supponiamo vero il primo fatto ed eseguiamo la dimostrazione in quasto caso. L'altro caso si tratta con la stessa dimostrazione scambiando T con $-T$. Assumeremo che $||T|| > 0$ altrimenti la prova della tesi è ovvia. Per ipotesi esisterà una successione di punti x_n con $||x_n|| = 1$ tali che $(x_n|Tx_n) \to ||T|| =: \Lambda > 0$. Dato che T è compatto potremo estrarre una sottosuccessione, che per comodità indichiamo ancora nello stesso modo, per cui Tx_n converge a qualche punto di H. Vale (dove usiamo anche il fatto che $||Tx_n|| \leq ||T||||x_n|| = ||T||$)

$$||Tx_n - \Lambda x_n||^2 = ||Tx_n||^2 - 2\Lambda(x_n|Tx_n) + \Lambda^2 \leq ||T||^2 + \Lambda^2 - 2\Lambda(x_n|Tx_n) \,.$$

Essendo $||T|| = \Lambda$, nel limite a $n \to +\infty$ si ha che

$$Tx_n - \Lambda x_n \to 0 \,. \tag{4.5}$$

Ma allora anche

$$x_n = \frac{1}{\Lambda}[Tx_n - (Tx_n - \Lambda x_n)]$$

converge a qualche $x \in H$, in quanto combinazione lineare di successioni convergenti. Dato che T è continuo e $x_n \to x$, (4.5) implica che

$$Tx = \lambda x \,.$$

Notiamo che $x \neq 0$ perchè $||x|| = \lim_{n\to\infty} ||x_n|| = 1$. Abbiamo provato che x è un autovettore con autovalore Λ.

(7) è immediata conseguenza di (6).

(b) Sia H_Λ l'autospazio di T con autovalore $\lambda \neq 0$. Se $B \subset H_\lambda$ è la palla aperta di raggio 1 centrata nell'origine, allora si può scrivere $B = T(\lambda^{-1}B)$ e $\lambda^{-1}B$ è limitato a sua volta per costruzione. Dato che T è compatto, $\overline{B}$ è compatto. Quindi, nello spazio di Hilbert H_λ, la

chiusura della palla aperta unitaria è un insieme compatto e quindi $dimH_\lambda < \infty$ per la proposizione 4.2. $\square$

Passiamo al teorema di Hilbert sullo sviluppo degli operatori compatti autoaggiunti su una base di hilbertiana autovettori.

Teorema 4.3 (Hilbert). *Sia $(H, (\,|\,))$ spazio di Hilbert e $T \in \mathfrak{B}_0(T)$. Se P_λ è il proiettore ortogonale sull'autospazio con autovalore $\lambda \in \sigma_p(T)$, dove $\sigma_p(T)$ è l'insieme degli autovalori di T, ordinato in modo tale che $|\lambda_0| > |\lambda_1| > |\lambda_2| \ldots$, vale*

$$T = \sum_{\lambda \in \sigma_p(T)} \lambda P_\lambda \,, \tag{4.6}$$

dove la somma della serie è nel senso della topologia uniforme.

Prova del teorema 4.3 e della parte rimanente del teorema 4.2. Sia λ un autovalore con autospazio H_λ. Sia P_λ il proiettore ortogonale su H_λ e $Q_\lambda := I - P_\lambda$ il proiettore ortogonale su $H_\lambda^\perp$. Vale

$$TP_\lambda = P_\lambda T = \lambda_\lambda T \,. \tag{4.7}$$

Questa identità si dimostra notando che se $x \in H$, $P_\lambda x \in H_\lambda$ e allora $TP_\lambda x = \lambda P_\lambda x$ da cui $TP_\lambda = \lambda P_\lambda$. Prendendo l'aggiunto ad ambo membri e tendo conto del fatto che $\lambda \in \mathbb{R}$, $T = T^\dagger$, $P_\lambda = P_\lambda^\dagger$ si ha subito che $P_\lambda T = \lambda P_\lambda = TP_\lambda$. Come ulteriore conseguenza troviamo che vale anche, direttamente dalla definizione di Q_λ e da quanto appena provato,

$$Q_\lambda T = TQ_\lambda \,. \tag{4.8}$$

Notiamo ancora che essendo $I = P_\lambda + Q_\lambda$ abbiamo che $T = P_\lambda T + Q_\lambda T$ ossia

$$T = \lambda T + Q_\lambda T \,. \tag{4.9}$$

Si osservi che $Q_\lambda T$ è sicuramente ancora compatto per (b) del teorema 4.1 Nel seguito useremo ripetutamente queste identità anche senza citarle esplicitamente.
Cominciamo a scegliere $\lambda = \lambda_0$ l'autovalore con valore assoluto più alto. Avremo che

$$T = \lambda_0 P_0 + Q_0 T = \lambda_0 T + T_1$$

dove d'ora in poi scriveremo P_n, Q_n, H_n per, rispettivamente P_{λ_n} Q_{λ_n}, H_{λ_n}. Se $T_1 := Q_0 T = 0$ la dimostrazione è finita, se non lo è sappiamo comunque che T_1 è compatto per cui possiamo reiterare la procedura appena usata usando T_1 in luogo di T, trovando

$$T = \lambda_0 P_0 + \lambda_1 P_1 + Q_1 T_1 = \lambda_0 P_0 + \lambda_1 P_1 + T_2 \,.$$

Notiamo che sopra λ_1 è un autovalore, quello con modulo maggiore, di T_1 e P_1 è il proiettore ortogonale sull'autospazio associato. Tuttavia λ_1 è anche autovalore di T con lo stesso autovalore: se $u \in H_1 \subset H_0^\perp$ allora

$$Tu = \lambda_0 P_0 u + \lambda_1 P_1 u + T_1 Q_1 u = 0 + \lambda_1 u + 0 = \lambda_1 u \,.$$

In particolare sarà dunque

$$|\lambda_1| \leq |\lambda_0|.$$

Si osservi ancora che ogni autovettore u di T con autovalore λ_1 deve essere tale anche per $T_1 = Q_0 T$: infatti applicando u ad ambo membri si trova

$$T_1 u = Q_0 \lambda u = \lambda u - \lambda P_0 u = \lambda u$$

dove abbiamo usato che autovettori di autovalori distinti per operatori autoaggiunti (T) sono ortogonali. In definitiva, H_1 coincide con l'autospazio di T con autovalore λ_1 e P_1 è il proiettore su tale autospazio. Non potrà quindi essere $\lambda_0 = \lambda_1$ perchè H_0 e H_1 sono stati costruiti in modo che $H_1 \subset H_0^{\perp}$. In particolare

$$|\lambda_1| < |\lambda_0|.$$

Questo fatto ha una consegunza importante. Dato che $||T|| = |\lambda_0|$ e $||T_1|| = \lambda_1$ per il teorema precedente, avremo che

$$||T_1|| < ||T|| \, .$$

Se risulta $T_2 = 0$ la dimostrazione è finita, altrimenti procederemo nello stesso modo ottenendo che

$$T - \sum_{k=0}^{n} \lambda_k P_k = T_n \tag{4.10}$$

dove

$$|\lambda_0| > |\lambda_1| > \cdots > |\lambda_k| > \ldots$$

e

$$||T_k|| = |\lambda_k| \tag{4.11}$$

Se nessuno degli operatori T_k è l'operatore nullo, la procedura continua indefinitamente. Mostriamo in tal caso la serie decrescente dei numeri positivi $|\lambda_k|$ deve tendere a 0 (non ci può essere un punto di accumulazione precedente a 0). Supponiamo che $|\lambda_0| > |\lambda_1| > \cdots > |\lambda_k| > \ldots \geq \epsilon > 0$ Scegliamo un vettore $x_n \in H_n$ per ogni n ed in modo tale che $||x_n|| = 1$. Dunque la successione degli x_n è limitata, per cui la successione dei Tx_n o una sua sottosuccessione deve convergere essendo T compatto. Ma questo è impossibile perchè, sviluppando il quadrato della norma e tenendo conto che x_n e x_m sono perpendicolari ((ii) in (b) di proposizione 3.8)

$$||Tx_n - Tx_m||^2 = ||\lambda_n x_n - \lambda_m x_m|| = |\lambda_n|^2 + |\lambda_m|^2 \geq \epsilon \, ,$$

per ogni n e m, per cui nè la successione dei Tx_n nè una sua sottosuccessione può convergere. Questo è assurdo e quindi la successione dei λ_n (se effettivamente sono infiniti) deve convergere a 0. Come conseguenza di (4.10) e (4.11), quanto appena provato implica che vale anche

$$T = \sum_{k=0}^{\infty} \lambda_k P_k \, . \tag{4.12}$$

Questa identità è (4.6) se la classe dei λ_k esaurisce la classe degli autovalori non nulli di T. Per provare che è vero, notiamo che se $\lambda \neq \lambda_n$ per ogni n è un autovalore non nullo con proiettore P_λ, valendo $P_n P_\lambda = 0$ per ogni n (perchè autospazi di autovalori distinti di operatori autoaggiunti sono ortogonali per (ii) in (b) di proposizione 3.8) (4.12) implicherebbe

$$TP_\lambda = \sum_{k=0}^{\infty} \lambda_k P_k P_\lambda = 0 \,,$$

da cui, se $u \in H_\lambda$

$$Tu = 0$$

per cui $\lambda = 0$ caso che abbiamo escluso. □

Il teorema di Hilbert insieme al teorema di decomposizione polare 9.3 consentono di estendere la formula (4.6) dello sviluppo di un operatore compatto autoaggiunto ad una analoga formula valida nel caso non autoaggiunto. Abbiamo bisogno di una proposizione preliminare prima di endunciare e provare il teorema.

Proposizione 4.3. *Sia H spazio di Hilbert. $A \in \mathfrak{B}(H)$ è compatto se e solo se $|A|$ (definizione 3.13) è compatto.*

Prova. Supponiamo A compatto. Sia $\{x_k\}_k$ una successione limitata in H e sia $\{Ax_{k_n}\}_n$ una sottosuccessione di $\{Ax_k\}_k$ convergente in virtù della compattezza di A. Essendo tale sottosuccessione successione di Cauchy, in virtù di (3.41), la sottosuccessione di $\{|A|x_k\}_k$, $\{|A|x_{k_n}\}_n$ è successione di Cauchy e quindi converge. Quindi $|A|$ è compatto. Scambiando il ruolo di A e $|A|$ ed usando la stessa dimostrazione si prova che se $|A|$ è compatto è tale anche A.

Diamo una definizione conclusiva che sarà utile nel seguito.

Definizione 4.2. *Sia H spazio di Hilbert e $A \in \mathfrak{B}_0(H)$. Gli autovalori λ non nulli di $|A|$ vengono detti* **valori singolari** *di A e il loro insieme si indica con $sing(A)$. La dimensione (finita) m_λ dell'autospazio corrispondente a $\lambda \in sing(A)$ viene detta* **molteplicità** *di λ.*

Teorema 4.4. *Sia $(H, (\,|\,))$ spazio di Hilbert e $A \in \mathfrak{B}_0(H)$ con $A \neq 0$ in modo che l'insieme dei valori singolari di A, $sing(A)$, sia non vuoto. Ordinando $sing(A)$ in modo decrescente: $\lambda_0 > \lambda_1 > \lambda_2 > \ldots$, ed usando tale ordinamento nella somme scritte sotto vale:*

$$A = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda \, (u_{\lambda,i}| \quad) \, v_{\lambda,i} \,, \tag{4.13}$$

$$A^\dagger = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda \, (v_{\lambda,i}| \quad) \, u_{\lambda,i} \,, \tag{4.14}$$

dove le somme, se infinite, sono intese nel senso della convergenza della topologia uniforme e, per ogni $\lambda \in sing(A)$, l'insieme degli elementi $u_{\lambda,1}, \ldots, u_{\lambda,m_\lambda}$ è una base ortonormale nell'autospazio di $|A|$ associato all'autovalore λ. L'insieme ortonormale di vettori $\{v_{\lambda,i}\}$ è definito da

$$v_{\lambda,i} := \frac{1}{\lambda} A u_{\lambda,i} \,, \tag{4.15}$$

per ogni $\lambda \in sing(A)$ e $i = 1, 2, \ldots, m_\lambda$. Per gli stessi valori di λ e i risulta infine

$$v_{\lambda,i} = U u_{\lambda,i} \,, \tag{4.16}$$

dove U è l'operatore che appare nella decomposizione polare $A = U|A|$.

Prova. Notiamo che $|A|$ è autoaggiunto positivo e compatto. I suoi autovalori saranno quindi reali positivi e distribuiti come specificato in (a) del teorema 4.2 Studieremo il caso in cui l'insieme degli autovalori è infinito (e quindi numerabile) lasciando il sottocaso banale al lettore. Il teorema 4.3 permette di sviluppare $|A|$ come

$$|A| = \sum_{\lambda \in \sigma_p(|A|)} \lambda P_\lambda \,,$$

dove la convergenza è intesa rispetto alla topologia uniforme. È chiaro che ci si può ridurre a considerare solo gli autovalori non nulli dato che quello nullo non fornisce contributo alla serie :

$$|A| = \sum_{\lambda \in sing(A)} \lambda P_\lambda \,,$$

Se non ci fossero autovalori non nulli ci sarebbe l'unico autovalore nullo e sarebbe $|A| = 0$ per il teorema 4.2 e quindi dalla decomposizione polare avremmo $A = 0$ e la prova del teorema sarebbe ovvia.
Si prova facilmente che se $\mathfrak{B}(H) \ni T_n \to T \in \mathfrak{B}(H)$ nella topologia uniforme e se $U \in \mathfrak{B}(H)$, dalla limitatezza di U vale anche $UT_n \to UT$ nsempre nel senso della topologia uniforme degli operatori. Dato che l'operatore U della decomposizione polare di A, $A = U|A|$ è limitato, avremo subito che, nella convergenza della topologia uniforme:

$$A = U|A| = \sum_{\lambda \in sing(A)} \lambda U P_\lambda \,. \tag{4.17}$$

Il teorema 4.2 nel punto (b) precisa che il sottospazio chiuso su cui proietta ogni P_λ (con $\lambda > 0$) ha dimensione finita m_λ. Possiamo sempre trovare una base ortonormale di tale autospazio: $\{u_{\lambda,i}\}_{i=1,\ldots,m_\lambda}$. Si osservi che $(u_{\lambda,i}|u_{\lambda',j}) = \delta_{\lambda\lambda'}\delta_{ij}$ per costruzione, siccome autovettori con autovalori distinti sono ortogonali (essendo $|A|$ normale perchè positivo) per (ii) in (b) di proposizione 3.8 Dato che $u_{\lambda,i} = A(u_{\lambda,i}/\lambda)$ segue che $u_{\lambda,i} \in Ran(|T|)$. Quindi U agisce sugli $u_{\lambda,i}$ in modo

isometrico per cui i vettori a secondo membro di (4.16) sono ancora ortonormali. (4.16) equivale a (4.15) per la decomposizione polare di A:

$$Au_{\lambda,i} = U|A|u_{\lambda,i} = U\lambda u_{\lambda,i} = \lambda v_{\lambda,i} \ .$$

È un facile esercizio provare che proiettore ortogonale P_λ ($\lambda > 0$) si può scrivere come

$$P_\lambda = \sum_{i=1}^{m_\lambda} (u_{\lambda,i}| \quad) \, u_{\lambda,i} \ .$$

Di conseguenza

$$UP_\lambda = \sum_{i=1}^{m_\lambda} (u_{\lambda,i}| \quad) \, Uu_{\lambda,i} = \sum_{i=1}^{m_\lambda} (u_{\lambda,i}| \quad) \, v_{\lambda,i} \ .$$

Sostituendo in (4.17) si trova (4.13). (4.14) si ottiene da (4.13) tenendo conto dei seguenti fatti. (i) L'operazione di coniugazione hermitiana $A \mapsto A^\dagger$ è antilineare, per cui trasforma combinazioni lineari di operatori in combinazioni lineari dei loro aggiunti e con i corrisondenti coefficienti sostituiti dai complessi coniugati dei coefficienti della combinazione lineare iniziale.
(ii) L'operazione di coniugazione hermitiana è continua nella topologia uniforme di $\mathfrak{B}(H)$ in quanto $||A^\dagger|| = ||A||$ per (a) di proposizione 3.6
In conseguenza di ciò da (4.13) troviamo (si tenga conto del fatto che $\lambda \in \mathbb{R}$):

$$A^\dagger = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda \, [(u_{\lambda,i}| \quad) \, v_{\lambda,i}]^\dagger \ ,$$

dove la serie converge nella topologia uniforme. È un facile esercizio verificare che

$$[(u_{\lambda,i}| \quad) \, v_{\lambda,i}]^\dagger = (v_{\lambda,i}| \quad) \, u_{\lambda,i} \ ,$$

da cui segue immediatamente la (4.14). □

Il teorema provato consente di introdurre la nozione di operatore di Hilbert-Schmidt e di operatore di classe traccia che vedremo nelle prossime sezioni.

4.3 Operatori di Hilbert-Schmidt.

Una classe particolare di operatori compatti sono quelli di Hilbert-Schmidt. Questi trovano applicazioni in vasti rami della fisica matematica oltre che nella meccanica quantistica.

Definizione 4.3. *Sia $(H, (\,,|\,))$ spazio di Hilbert e sia $||\ ||$ la norma indotta dal prodotto scalare. $A \in \mathfrak{B}(H)$ è detto operatore di* **Hilbert-Schmidt** *se esiste una base hilbertiana $\mathfrak{U}$ per cui l'insieme $\{||Au||^2\}_{u \in \mathfrak{U}}$ ammette somma finita nel senso della definizione 3.4.*

La classe degli operatori di Hilbert-Schmidt su H sarà indicata con $\mathfrak{B}_2(H)$ e, se $A \in \mathfrak{B}_2(H)$ allora

$$||A||_2 := \sqrt{\sum_{u \in \mathcal{U}} ||Au||^2}\,, \tag{4.18}$$

dove $\mathcal{U}$ è la base hilbertiana suddetta.

Per prima cosa mostriamo che la particolare base scelta nella definizione è in realtà generica e che anche $||A||_2$ non dipende dalla scelta di tale base.

Proposizione 4.4. *Sia $(H, (\,,|))$ spazio di Hilbert, sia $||\;||$ la norma indotta dal prodotto scalare e siano $\mathcal{U}$ e $\mathcal{V}$ due basi hilbertiane e $A \in \mathfrak{B}(H)$.*
(a) *$\{||Au||^2\}_{u \in \mathcal{U}}$ ammette somma finita se e solo se $\{||Av||^2\}_{v \in \mathcal{V}}$. In tal caso vale*

$$\sum_{u \in \mathcal{U}} ||Au||^2 = \sum_{v \in \mathcal{U}} ||Av||^2\,. \tag{4.19}$$

(b) *$\{||Au||^2\}_{u \in \mathcal{U}}$ ammette somma finita se e solo se $\{||A^\dagger v||^2\}_{v \in \mathcal{U}}$ ammette somma finita. In tal caso vale*

$$\sum_{u \in \mathcal{U}} ||Au||^2 = \sum_{v \in \mathcal{U}} ||A^\dagger v||^2\,. \tag{4.20}$$

Prova. Proviamo prima (b). Consideriamo che per (d) del teorema 3.4

$$||Au||^2 = \sum_{v \in V} |(v|Au)|^2 < \infty\,,$$

per cui, fissato u, solo una quantità al più numerabile di coefficienti $|(v|Au)$ è non nulla per (b) di proposizione 3.2. Ciò determina un insieme al più numerabile $\mathcal{V}(u) \subset \mathcal{V}$ tale che

$$\sum_{u \in \mathcal{U}} ||Au||^2 = \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}(u)} |(v|Au)|^2 < \infty\,. \tag{4.21}$$

Questo significa in particolare, riutilizzando (b) di proposizione 3.2, che solo un insieme al più numerabile di elementi $u \in \mathcal{U}$ fornisce somme $\sum_{v \in \mathcal{V}(u)} |(v|Au)|^2$ non nulle. In definitiva i coefficienti $(v|Au)$ sono non nulli solo per un insieme Z al più numerabile di coppie $(u, v) \in \mathcal{U} \times \mathcal{V}$. Aggiungendo coppie (u, v) per cui $|(v|Au)| = 0$ possiamo sempre vedere l'insieme Z come un sottoinsieme di $U \times V$ dove U e V sono sottoinsiemi numerabili di $\mathcal{U}$ e $\mathcal{V}$ rispettivamente. Possiamo mettere su U e V le misure μ e ν rispettivamente che contano gli elementi di U e V e le serie sopra considerate si trascrivono in termini di integrali rispetto alle misure dette (proposizione 3.2). In particolare (4.21) si riscrive:

$$\sum_{u \in \mathcal{U}} ||Au||^2 = \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}(u)} |(v|Au)|^2 = \int_V d\nu(v) \int_U d\mu(u) |(v|Au)|^2 < +\infty\,. \tag{4.22}$$

μ e ν sono σ-finite per costruzione per cui ha senso definire la misura prodotto $\mu \otimes \nu$ e applicare il teorema di Fubini-Tonelli. Il teorema di Fubini-Tonelli assicura che, per l'ultima parte di (4.22), la funzione $(v, u) \mapsto |(v|Au)|^2$ è allora integrabile rispetto alla misura prodotto e che possiamo interscambiare gli integrali:

$$\sum_{u \in \mathcal{U}} ||Au||^2 = \int_{U \times V} |(v|Au)|^2 d\mu(u) \otimes d\nu(v) = \int_U d\mu(u) \int_V d\nu(u) |(v|Au)|^2 < +\infty \, .$$

Notiamo che $(v|Au) = (A^\dagger v|u)$ per cui solo una quantità numerabile di elementi $(A^\dagger v|u)$ (con $(u, v) \in U \times V$) sarà non nulla e avremo in particolare

$$\sum_{v \in \mathcal{V}} \sum_{u \in \mathcal{U}} |(A^\dagger v|u)|^2 = \int_U d\mu(u) \int_V d\nu(u) |(A^\dagger v|u)|^2 = \sum_{u \in \mathcal{U}} ||Au||^2 < +\infty \, .$$

Ma il primo membro è proprio $\sum_{v \in \mathcal{V}} ||A^\dagger v||^2$. Abbiamo provato parte della proposizione (b): *se* $\{||Au||^2\}_{u \in \mathcal{U}}$ *ammette somma finita allora* $\{||A^\dagger v||^2\}_{v \in \mathcal{V}}$ *ammette somma finita, ed in tal caso vale (4.20).*
Applicando questa parte della proposizione (a) all'operatore $A^\dagger$ invece che A, scambiando il nome di $\mathcal{U}$ con quello di $\mathcal{V}$ (ciò è possibile in quanto sono arbitrarie per ipotesi) e tenendo conto del fatto che $(A^\dagger)^\dagger = A$ si ha la seconda parte di (b):
se $\{||A^\dagger v||^2\}_{v \in \mathcal{V}}$ *ammette somma finita allora* $\{||Au||^2\}_{u \in \mathcal{U}}$ *ammette somma finita, ed in tal caso vale (4.20).*
La dimostrazione di (a) segue immediatamente dalla proposizione (b) e dalla stessa (b) in cui l'operatore A è rimpiazzato da $A^\dagger$ e $A^\dagger$ da $(A^\dagger)^\dagger = A$ e le basi $\mathcal{U}$ e $\mathcal{V}$ si scambiano nome. $\square$

Possiamo eneunciare e provare alcune delle numerose ed interessanti proprietà matematiche degli operatori di classe traccia. La propietà più interessante dal punto di vista matematico è la (b) del teorema scritto sotto: gli operatori di Hilbert-Schmidt formano uno spazio di Hilbert con un prodotto scalare che induce una norma che coincide, per ogni operatore A di Hilbert-Schimdt proprio con il numero che avevamo indicato con $||A||_2$. Un'altro fatto importante è che gli operatori di Hilbert-Schmidt sono compatti e formano un ideale all'interno dello spazio degli operatori limitati.

Teorema 4.5. *Sia H spazio di Hilbert. Gli operatori di Hilbert-Schmidt godono delle seguenti proprietà.*
(a) *$\mathfrak{B}_2(H)$ è un sottospazio di $\mathfrak{B}(H)$ e più fortemente uno $*$-ideale di $\mathfrak{B}(H)$, inoltre:*
(i) $||A||_2 = ||A^\dagger||_2$ *per ogni* $A \in \mathfrak{B}_2(H)$*;*
(ii) $||AB||_2 \leq ||B|| \, ||A||_2$ *e* $||BA||_2 \leq ||B|| \, ||A||_2$ *per ogni* $A \in \mathfrak{B}_2(H)$ *e* $B \in gB(H)$*;*
(iii) $||A|| \leq ||A||_2$ *per ogni* $A \in \mathfrak{B}_2(H)$*.*
(b) *Se $A, B \in \mathfrak{B}_2(H)$ e se $\mathcal{B}$ è una base hilbertiana in H, si definisca*

$$(A|B)_2 := \sum_{x \in \mathcal{B}} (Ax|Bx) \, . \tag{4.23}$$

L'applicazione

$$(\mid)_2 : \mathfrak{B}_2(H) \times \mathfrak{B}_2(H) \to \mathbb{C}$$

è ben definita (la somma di sopra si riduce sempre ad una serie assolutamente convergente), definisce un prodotto scalare su $\mathfrak{B}_2(H)$ che per ogni $A \in \mathfrak{H}_2(H)$ soddisfa:

$$(A|A)_2 = ||A||_2^2 \tag{4.24}$$

(c) $(\mathfrak{B}_2(H), (\mid)_2)$ *è spazio di Hilbert.*
(d) $\mathfrak{B}_2(H) \subset \mathfrak{B}_0(H)$. *Più precisamente, $A \in \mathfrak{B}_2(H)$ se e solo se è compatto e l'insieme dei numeri positivi $\{m_\lambda \lambda^2\}_{\lambda \in sing(A)}$ (essendo m_λ la molteplicità di λ) ha somma finita. In tal caso risulta*

$$||A||_2 = \sqrt{\sum_{\lambda \in sing(A)} m_\lambda \lambda^2}\,. \tag{4.25}$$

Prova. (a) Ovviamante $\mathfrak{B}_2(H)$ è chiuso rispetto alla moltiplicazione per scalare. Mostriamo che lo è rispetto a quello di somma di operatori. Se $A, B \in \mathfrak{B}_2(H)$ e $\mathcal{B}$ è una qualsiasi base hilbertiana:

$$\sum_{u \in \mathcal{B}} ||(A+B)u||^2 \leq \sum_{u \in \mathcal{B}} (||Au|| + ||Bu||)^2 \leq 2 \left[\sum_{u \in \mathcal{B}} ||Au||^2 + \sum_{u \in \mathcal{B}} ||Bu||^2 \right].$$

Quindi $\mathfrak{B}_2(H)$ è sottospazio di $\mathfrak{B}(H)$. Il fatto che sia chiuso rispetto alla coniugazione hermitiana è una immediata conseguenza di (b) della proposizione 4.4 che implica anche (i). Mostriamo che vale (ii) e contestualmente che $\mathfrak{B}(H)$ è chiuso rispetto alla composizione destra e sinistra con operatori limitati. Se $A \in \mathfrak{B}_2(H)$ e $B \in \mathfrak{B}_2(H)$ allora:

$$\sum_{u \in \mathcal{B}} ||BAu||^2 \leq \sum_{u \in \mathcal{B}} ||B||^2 ||Au||^2 = ||B||^2 \sum_{u \in \mathcal{B}} ||Au||^2\,.$$

Questo prova insieme che $\mathfrak{B}(H)$ è chiuso rispetto alla composizione sinistra con operatori limitati e che vale la prima disuguaglianza in (ii). La chiusura rispetto alla moltiplicazione a destra è immediata conseguenza di quella a sinistra e della chiusura rispetto alla coniugazione hermitiana: $B^\dagger A^\dagger \in \mathfrak{B}_2(H)$ se $A \in \mathfrak{B}_2(H)$ e $B \in \mathfrak{B}_2(H)$ e quindi $(B^\dagger A^\dagger)^\dagger \in \mathfrak{B}_2(H)$ ossia $AB \in \mathfrak{B}_2(H)$. Da (i) abbiamo anche che se $A \in \mathfrak{B}_2(H)$ e $B \in \mathfrak{B}_2(H)$ allora $||AB||_2 = ||(AB)^\dagger||_2 = ||B^\dagger A^\dagger||_2 \leq ||B^\dagger|| ||A^\dagger||_2 = ||B|| ||A||_2$ che completa la dimostrazione di (ii). Per provare (iii) notiamo che:

$$||A|| = \sup_{||x||=1} ||Ax|| = \sup_{||x||=1} (||Ax||^2)^{1/2} = \sup_{||x||=1} \left(\sum_{u \in \mathcal{B}} |(u|Ax)|^2 \right)^{1/2} = \sup_{||x||=1} \left(\sum_{u \in \mathcal{B}} |(A^\dagger u|x)|^2 \right)^{1/2}$$

dove abbiamo usato (d) del teorema 3.4 rispetto alla base hilbertiana $\mathcal{B}$. Utilizzando la disuguaglianza di Cauchy-Schwartz nell'ultimo termine della catena di uguaglianze di sopra otteniamo:

$$||A|| \leq \sup_{||x||=1} \left(\sum_{u \in \mathcal{B}} ||Au||^2 ||x||^2 \right)^{1/2} = \left(\sum_{u \in \mathcal{B}} ||Au||^2 \right)^{1/2} = ||A^\dagger||_2 = ||A||_2\,.$$

(b) Se $A, B \in \mathfrak{B}_2(H)$ e $\mathcal{B}$ è una base hilbertiana allora la quantità di termini Au e Bu non nulli, per $u \in \mathcal{B}$ è al più numerabile per la definizione 4.3 e (b) di proposizione 3.2 Valendo

$$|(Au|Bu)| \leq ||Au||\,||Au|| \leq \frac{1}{2}(||Au||^2 + ||Bu||^2)$$

concludiamo che la quantità di termini non nulli $(Au|Bu)$ per $u \in \mathcal{B}$ è anch'essa al più numerabile e che la serie dei termini non nulli $(Au|Bu)$ è assolutamente convergente per cui non conta l'ordinamento scelto nel calcolare (4.23). Mostreremo tra poco che la scelta della base hilbertiana è irrilevante. Prima notiamo che (4.24) è vera banalmente e che $(\ |\ )_2$ soddisfa gli assiomi di un semiprodotto scalare come è immediato verificare. L'assioma di positività **H3** è immediatamente conseguenza di (iii) quindi $(\ |\ )_2$ è un prodotto scalare che induce la norma $||\ ||_2$. Quindi vale la formula in (1) di esercizi 3.1 che sussiste per ogni prodotto scalare:

$$4(A|B)_2 = ||A+B||_2^2 + ||A-B||_2^2 - i||A+iB||_2^2 + i||A-iB||_2^2\,.$$

Dato che il secondo membro non dipende dalla base hilbertiana usata per la proposizione 4.4, non dipenderà dalla scelta della base nemmeno il primo membro.
(c) Bisogna solo provare che lo spazio è completo. Sia $\mathcal{B}$ una base hilbertiana di H. Sia $\{A_n\}_n$ una successione di Cauchy di operatori di Hilbert-Schmidt rispetto alla norma $||\ ||_2$. Da (iii) la successione è di Cauchy anche nella topologia uniforme per cui, essendo $\mathfrak{B}(H)$ completo per il torema 3.2, esisterà $A \in \mathfrak{B}(H)$ con $||A - A_n|| \to 0$ per $n \to +\infty$. La proprietà di Cauchy dice che preso $\epsilon > 0$, esiste N_ϵ tale che, $||A_n - A_m||_2^2 \leq \epsilon^2$ se $n, m > N_\epsilon$. Tenendo conto della definizione di $||\ ||_2$, varrà anche che, per lo stesso ϵ, per ogni sottoinsieme finito $I \subset \mathcal{B}$:

$$\sum_{u \in I} ||(A_n - A_m)u||^2 \leq ||A_n - A_m||_2^2 \leq \epsilon^2$$

quando $n, m > N_\epsilon$. Passando al limite per $m \to \infty$ si trova che vale ancora:

$$\sum_{u \in I} ||(A_n - A)u||^2 \leq \epsilon$$

per ogni sottoinsieme finito $I \subset \mathcal{B}$ se $n > N_\epsilon$. In definitiva, per l'arbitrarietà di I, abbiamo che:

$$||A - A_n||_2 \leq \epsilon \quad \text{se } n > N_\epsilon\,. \tag{4.26}$$

Da questo fatto segue in particolare che $A - A_n \in \mathfrak{B}_2(H)$ e quindi $A = A_n + (A - A_n) \in \mathfrak{B}_2(H)$ ed inoltre, per l'arbitrarietà di $\epsilon > 0$ in (4.26), vale anche che A_n tende a A rispetto alla norma $||\ ||_2$. Quindi ogni successione di Cauchy nella norma $||\ ||_2$ converge in $\mathfrak{B}_2(H)$ che di conseguenza è completo.
(c) Sia $A \in \mathfrak{B}_2(H)$ mostriamo che è compatto e che soddisfa (4.25). Sia $\mathcal{B}$ una base hilbertiana. Vale $\sum_{u \in \mathcal{B}} ||Au||^2 < +\infty$ dove solo una quantità numerabile di addendi è non nulla ed infine sappiamo che la somma si può scrivere come una serie come be noto considerando solo i valori u_n per cui $||Au||^2 > 0$. Di conseguenza varrà che: per ogni $\epsilon > 0$ esiste N_ϵ tale che

$$\sum_n^{+\infty} ||Au_n||^2 \leq \epsilon^2\,.$$

La stessa proprietà si può quindi trascrivere in termini di tutti gli elementi di $\mathcal{B}$: per ogni $\epsilon > 0$ esiste un sottoinsieme finito $I \subset \mathcal{B}$ tale che

$$\sum_{u \in \mathcal{B} \setminus I} ||Au||^2 < \epsilon^2 \, .$$

Sia quindi A_I l'operatore definito come: $A_I u := Au$ se $u \in I$ e $A_I u := 0$ se $u \in \mathcal{B} \setminus I$. Per costruzione, se P_I è il proiettore ortogonale sul sottospazio generato dagli u con $u \in I$, $||A_I x|| = ||AP_I x|| \leq ||A|| \, ||P_I|| \, ||x|| \leq$, in particolare A_I è quindi limitato. Il rango di A_I è generato dai vettori Au con $u \in I$ per cui A_I è compatto ((1) in esempi 4.1). Inoltre per (iii) di (a) abbiamo che

$$||A - A_I|| \leq ||A - A_I||_2 = \left(\sum_{u \in \mathcal{B} \setminus I} ||Au||^2 \right)^{1/2} \leq \epsilon \, .$$

Quindi A è un punto di accumulazione dello spazio degli operatori compatti rispetto alla topologia uniforme. Tenendo conto del fatto che l'ideale degli operatori compatti è chiuso nella topologia uniforme ((c) di teorema 4.1), il risultato ottenuto prova che A è compatto e che vale (4.25). Consideriamo ora l'operatore positivo compatto $|A| = \sqrt{A^\dagger A}$ e sia $\{u_{\lambda,i}\}_{\lambda \in sing(A), i=\{1,2,\ldots,m_\lambda\}}$ una base hilbertiana di $Ker(|A|)^\perp$ costruita come nel teorema 4.4. Possiamo completarla a base dello spazio di Hilbert aggiungendo una base hilbertiana per $Ker(|A|)$, che coincide con $Ker(A)$ per la nota dopo la definizione 3.13. Usando la base hilbertiana detta per esprimere $||A||_2$ abbiamo:

$$||A||_2^2 = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} (Au_{\lambda,i} | Au_{\lambda,i}) = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} (A^\dagger A u_{\lambda,i} | u_{\lambda,i}) = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} m_\lambda \lambda^2 \, . \quad (4.27)$$

Dove si è tenuto conto del fatto che, la parte di base riguardante $Ker(A)$ per costruzione non fornisce contributo, $|A| u_{\lambda,i} = \sqrt{A^\dagger A} u_{\lambda,i} = \lambda u_{\lambda,i}$ e $(u_{\lambda,i} | u_{\lambda',j}) = \delta_{\lambda\lambda'} \delta_{ij}$. Se viceversa A è compatto e $\{m_\lambda \lambda^2\}_{\lambda \in sing(A)}$ ha somma finita, allora da (4.27) segue che $||A||_2 < +\infty$ per cui $A \in \mathfrak{B}_2(H)$. $\square$

Esempi 4.2.
(1) Torniamo all'esempio (4) in esempi 4.1 Consideriamo gli operatori

$$T_K : L^2(X, \mu) \to L^2(X, \mu)$$

indotti dai nuclei integrali $K \in L^2(X \times X, \mu \otimes \mu)$ (con μ misura σ-finita separabile),

$$(T_K f)(x) := \int_X K(x, y) f(y) d\mu(y) \quad \text{per ogni } f \in L^2(X, \mu) \, .$$

Abbiamo provato che tali operatori sono compatti. Mostriamo ora che sono anche di Hilbert-Schmidt.

Usando le stesse definizione dell'esempio in questione, se $f \in L^2(X,\mu)$ avevamo visto che (vedi punto (3) nell'esempio (4)), per ogni $x \in X$,

$$F(x) = \int_X |K(x,y)|\,|f(y)| d\mu(y) < \infty\,.$$

Dato che $F \in L^2(X,\mu)$, per ogni $g \in L^2(X,\mu)$ la funzione $x \mapsto g(x)F(x)$ è una funzione integrabile (in modo da poter definire il prodotto scalare di g e F). Il teorema di Fubini-Tonelli assicura quindi che la funzione $(x,y) \mapsto g(x)K(x,y)f(y)$ è in $L^2(X \times X, \mu \otimes \mu)$ e che

$$\int_{X\times X} \overline{g(x)} K(x,y) f(y)\, d\mu(x) \otimes \mu(y) = \int_X d\mu(x)\, \overline{g(x)} \int_X K(x,y) f(y) d\mu(y) = (g|T_k f)\,. \quad (4.28)$$

Consideriamo allora una base hilbertiana di $L^2(X \times X, \mu \otimes \mu)$ del tipo $\{u_i \cdot \overline{u_j}\}_{i,j}$ dove $\{u_k\}_k$ è una base hilbertiana per $L^2(X,\mu)$ (e quindi lo è anche $\{\overline{u_k}\}_k$ come si prova immediatamente). Dato che $K \in L^2(X \times X, \mu \otimes \mu)$ varrà lo sviluppo

$$K = \sum_{i,j} \alpha_{ij} u_i \cdot \overline{u_j}\,,\ , \quad (4.29)$$

per cui

$$||K||^2_{L^2} = \sum_{i,j} |\alpha_{ij}|^2 < \infty\,. \quad (4.30)$$

D'altra parte, da (4.28) e (4.29)

$$(u_i|T_K u_j) = \int_{X\times X} \overline{u_i(x)} u_j(y) K(x,y)\, d\mu(x) \otimes \mu(y) = (u_i \cdot \overline{u_j}|K) = \alpha_{ij}\,,$$

per cui (4.30) si può riscrivere

$$||K||^2_{L^2} = \sum_{i,j} |(u_i|T_K u_j)|^2 < \infty\,.$$

Per definizione di operatore di Hilbert-Schmidt, T_K è quindi un operatore di Hilbert-Schmidt e vale anche

$$||T_K||_2 = ||K||_{L^2}\,.$$

(2) È abbastanza facile provare che se $H = L^2(X,\mu)$ con μ σ-finita e separabile, allora $\mathfrak{B}_2(H)$ consiste di tutti e soli gli operatori T_K con $K \in L^2(X \times X, \mu \otimes \mu)$ e che quindi l'applicazione che identifica K con T_K è un isomorfismo di spazi con prodotto scalare dallo spazio di Hilbert $L^2(X \times X, \mu \otimes \mu)$ allo spazio di Hilbert $\mathfrak{B}_2(H)$. Lasciamo la prova al lettore.

4.4 Operatori di classe traccia (o nucleari).

Passiamo ad introdurre gli operatori di *classe traccia* detti anche *operatori nucleari*. Seguiremo essenzialmente l'approccio in [14] (un altro approccio si trova in [16]).

Proposizione 4.5. *Sia H spazio di Hilbert e $A \in \mathfrak{B}(H)$. I seguenti tre fatti sono equivalenti.*
(a) *Esiste una base hilbertiana N per H per cui $\{(u||A|u)\}_{u\in N}$ ha somma finita:*

$$\sum_{u\in N}(u||A|u) < +\infty\,.$$

(a)' *$\sqrt{|A|}$ è operatore di Hilbert-Schmidt.*
(b) *A è compatto ed inoltre l'insieme indiciato $\{m_\lambda\lambda\}_{\lambda\in sing(A)}$, dove m_λ è la molteplicità di λ, ha somma finita.*

Prova. Si osservi che (a)' è una semplice trascrizione di (a) tenendo conto del fatto che $\sqrt{|A|}\sqrt{|A|} = |A|$.
In tutti i casi A è compatto (nel caso (a) perchè $|A| = (\sqrt{|A|})^2$ è compatto essendo $\sqrt{|A|}$ compatto in quanto ogni operatore di Hilbert-Schmidt è compatto ((d) teorema 4.5), il prodotto di operatori compatti è compatto ((b) teorema 4.1), ed infine $|A|$ è compatto se e solo se lo è A (proposizione 4.3).
Costruiamo una base hilbertiana di H usando in $Ker(|A|)^\perp$ autovettori di $|A|$: $u_{\lambda,i}$ con $i = 1,\ldots,m_\lambda$ e $|A|u_{\lambda,i} = \lambda u_{\lambda,i}$, e completando tale base in $Ker(|A|)$ con una qualsiasi base hilbertiana. Rispetto a tale base:

$$\left|\left|\sqrt{|A|}\right|\right|_2^2 = \sum_{\lambda,i}\left(\sqrt{|A|}u_{\lambda,i}\middle|\sqrt{|A|}u_{\lambda,i}\right) = \sum_{\lambda,i}\left(u_{\lambda,i}\middle|(\sqrt{|A|})^2u_{\lambda,i}\right) = \sum_{\lambda,i}(u_{\lambda,i}||A|u_{\lambda,i}) = \sum_\lambda m_\lambda\lambda\,.$$

La tesi è evidentemente vera in seguito all'identità ottenuta. □

Definizione 4.4. *Sia H spazio di Hilbert. $A \in \mathfrak{B}(H)$ è detto operatore di* **classe traccia** *o equivalentemente* **operatore nucleare** *se soddisfa una delle tre condizioni (a), (a)' o (b) in proposizione 4.5. La classe degli operatori di classe traccia su H sarà indicata con $\mathfrak{B}_1(H)$ e, se $A \in \mathfrak{B}_1(H)$ allora*

$$||A||_1 := \left|\left|\sqrt{|A|}\right|\right|_2^2 = \sum_{\lambda\in sing(A)} m_\lambda\lambda\,, \tag{4.31}$$

dove abbiamo usato le stesse notazioni che in proposizione 4.5.

Note.
(1) Il nome "classe traccia" deriva evidentemente almeno dal seguente fatto. Per un operatore A di classe traccia, $||A||_1$, che generalizza al caso infinitodimensionale la nozione di traccia di una matrice (che corrisponde ad $|A|$ *non* A) è finita. In realtà le analogie non finiscono qui come

vedremo tra poco.
(2) È chiaro che valgono le inclusioni:

$$\mathfrak{B}_1(H) \subset \mathfrak{B}_2(H) \subset \mathfrak{B}_0(H) \subset \mathfrak{B}(H)\,.$$

ossia

op. classe traccia $\subset$ op. Hilbert-Schmidt $\subset$ op. compatti $\subset$ op. limitati

(3) Ciscuno degli insiemi scritti è sottospazio vettoriale dello spazio degli operatori limitati ed in più $*$-ideale bilatero (per gli operatori di classe traccia lo proveremo tra poco) infine ciscuno di questi sottospazi è anche spazio di Hilbert o di Banach rispetto ad una naturale struttura: gli operatori compatti formano un sottospazio chiuso rispetto alla topologia uniforme in $\mathfrak{B}(H)$ e quindi uno spazio di Banch rispetto alla norma operatoriale, gli operatori di Hilbert-Schmidt definiscono uno spazio di Hilbert rispetto al prodotto scalare di Hilbert-Schmidt e gli operatori di classe traccia definiscono uno spazio di Banach rispetto alla norma $||||_1$ come diremo più avanti.

Prima di passare ad estendere il concetto di traccia al caso infinitodimensionale, vediamo le proprietà più importanti degli operatori nucleari o di classe traccia.

Teorema 4.6. *Sia H spazio di Hilbert. Gli operatori nucleari in H godono delle seguenti proprietà.*
(a) *$A \in \mathfrak{B}_1(H)$ se e solo se esistono due operatori $B, C \in \mathfrak{B}_2(H)$ tali che $A = BC$. In tal caso:*

$$||BC||_1 \leq ||B||_2 ||C||_2\,. \tag{4.32}$$

(b) *$\mathfrak{B}_1(H)$ è un sottospazio di $\mathfrak{B}(H)$ e più fortemente è uno $*$-ideale, inoltre:*
(i) $||AB||_1 \leq ||B||\,||A||_1$ e $||BA||_1 \leq ||B||\,||A||_1$ per ogni $A \in \mathfrak{B}_1(H)$ e $B \in gB(H)$;
(ii) $||A||_1 = ||A^\dagger||_1$ per ogni $A \in \mathfrak{B}_1(H)$;
(c) *$||\ ||_1$ definisce una norma su $\mathfrak{B}_1(H)$.*

Nota. È possibile provare che $\mathfrak{B}_1(H)$ è uno spazio di Banach rispetto alla norma $||\ ||$ (vedi referenze in [14]).

Prova della proposizione 4.6 eccetto (ii) in (b). (a) Se A è di classe traccia, usando la decomposizione polare di $A = U|A|$ (teorema 3.9), B e C possono sempre essere definiti come $B := U\sqrt{|A|}$ e $C := \sqrt{|A|}$. Per definizione di operatore di classe traccia $\sqrt{|A|}$ è operatore di Hilbert-Schmidt dunque C è Hilbert-Schmidt. Lo è anche B in quanto $U \in \mathfrak{B}(H)$ e gli operatori di Hilbert-Schmidt formano un ideale in $\mathfrak{B}(H)$ come provato nel teorema 4.5. Viceversa, se B e C sono di Hilbert-Schmidt, mostriamo che $A := BC$ è di classe traccia. Notiamo che A è compatto per il teorema 4.5. dobbaimo solo provare che $\sum_{\lambda \in sing(A)} m_\lambda \lambda < +\infty$. Se $BC = 0$ la tesi è ovvia. Assumiamo $BC \neq 0$ e sviluppiamo l'operatore compatto BC come nel teorema 4.4

$$A = BC = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda (u_{\lambda,i}|\ \ )v_{\lambda,i}\,.$$

Per non caricare troppo le notazioni poniamo $\Gamma = \{(\lambda, i) | \lambda \in sing(A),\ \ i = 1, 2, \ldots, m_\lambda\}$ e λ_j indichi il primo elemento della coppia $j = (\lambda, i)$. Allora è chiaro che

$$\sum_{j\in\Gamma} \lambda_j = \sum_{\lambda\in singA} m_\lambda \lambda\,.$$

Se $S \subset \Gamma$ è finito:

$$\sum_{j\in S} \lambda_j = \sum_{j\in S} (v_j|BCu_j) = \sum_{j\in S} (B^\dagger v_j|Cu_j) \leq \sum_{j\in S} j||B^\dagger v_j||\,||Cu_j|| \leq \sqrt{\sum_{j\in S} ||B^\dagger v_j||^2}\sqrt{\sum_{j\in S} j||Cu_j||^2}.$$

Dato che gli insiemi ortonormali di elementi $u_j = u_{\lambda,i}$ e $v_j = v_{\lambda,i}$ possono essere separatamente completati a base hilbertiana, l'ultimo termine della catena di disuguaglianze di sopra è maggiorato da:

$$||B^\dagger||_2||C||_2 = ||B|_2||C||_2\,.$$

Prendendo l'estremo supreiore su tutti gli S finiti, concludiamo che:

$$||BC||_1 = \sum_{\lambda\in sing(A)} m_\lambda \lambda \leq ||B|_2||C||_2$$

ed in particolare $A = BC \in \mathfrak{B}_1(H)$.
(b) e (c). La chiusura di $\mathfrak{B}_1(H)$ rispetto al prodotto per scalare si prova immediatamente dalla stessa definizione di operatore nucleare. Mostriamo che $\mathfrak{B}_1(H)$ è chiuso rispetto alla somma. Siano $A, B \in \mathfrak{B}_1(H)$. Se $A + B = 0$ allora $A + B$ è nucleare banalmente. Assumiamo pertanto $A+B \neq 0$ che è comunque compatto. Decomponendo polarmente abbiamo: $A = U|A|$, $B = V|B|$, $A + B = W|A + B|$. Usando la solita decomposizione sui valori singolari come nella prova di (a) avremo:

$$A + B = \sum_{\beta\in sing(A+B)} \sum_{i=1}^{m_\beta} \beta(u_{\beta,i}|\ \ )v_{\beta,i}\,.$$

Se $\Gamma = \{(\beta, i) | \beta \in sing(A + B),\ \ i = 1, 2, \ldots, m_\beta\}$ e $S \subset \Gamma$ è finito e β_j è il primo elemento della coppia $j \in \Gamma$:

$$\sum_{j\in S} \beta_j = \sum_{j\in S} (v_j|(A + B)u_j) = \sum_{j\in S} (v_j|Au_j) + \sum_{j\in S} (v_j|Bu_j)\,.$$

La somma delle due somme ottenuta alla fine può essere riscritta comealtrimenti ottenendo

$$\sum_{j\in S} \beta_j \leq \sum_{j\in S} (\sqrt{|A|}U^\dagger v_j|\sqrt{|A|}u_j) + \sum_{j\in S} (\sqrt{|B|}V^\dagger v_j|\sqrt{|B|}u_j)\,.$$

Procedendo come per l'analoga dimostrazione in (a) otteniamo

$$\sum_{j\in S} \beta_j \leq ||\sqrt{|A|}U^\dagger||_2\,||\sqrt{|A|}||_2 + ||\sqrt{|B|}V^\dagger||_2\,||\sqrt{|B|}||_2 \leq ||\sqrt{|A|}||_2^2 + ||\sqrt{|B|}||_2^2\,,$$

(nell'ultimo passaggio abbiamo tenuto conto del fatto che, per esempio per A vale la disuguaglianza ((a) teorema 4.5) $||\sqrt{|A|}U^\dagger||_2 \leq ||\sqrt{|A|}||_2\, ||U^\dagger||$ dato che $\sqrt{|A|}$ è di Hilbert-Schmidt e ancora $||U^\dagger|| \leq 1$ come è immediato provare visto che è un operatore isometrico su $ker(|A|)^\perp$ e si annulla su $Ker(|A|)$)
Notiamo infine che

$$||\sqrt{|A|}||_2^2 + ||\sqrt{|B|}||_2^2 = ||A||_1 + ||B||_1 \,.$$

Abbiamo in definitiva provato che $A + B \in \mathfrak{B}_1(H)$ e che in $\mathfrak{B}_1(H)$ vale la disuguaglianza triangolare:

$$||A + B||_1 + ||A||_1 + ||B||_1 \,.$$

Quindi $||\ \ ||_1$ è una seminorma. In realtà è una norma in quanto $||A||_1 = 0$ implica che gli autovalore di $|A|$ siano tutti nulli e quindi, essendo $|A|$ compatto, $|A| = 0$ per (7) in (a) di teorema 4.2. Per la decomposizione polare di $A = U|A|$, segue infine che $A = 0$. Fino ad ora abbiamo provato che $\mathfrak{B}_1(H)$
e sottospazio vettoriale di $\mathfrak{B}(H)$ e che $||\ \ ||_1$ è una norma. Mostriamo ora che $\mathfrak{B}_1(H)$ è chiuso rispetto alla composizione con operatori limitati a destra ed a sinistra. Siano $A \in \mathfrak{B}_1(H)$ e $B \in \mathfrak{B}(H)$ e $A = U|A|$. Allora $BA = (BU\sqrt{|A|})\,\sqrt{|A|}$ dove i due fattori sono operatori di Hilbert-Schmidt e quindi per (a), $BA \in \mathfrak{B}_1(H)$. Usando (ii) di (a) in teorema 4.5, (4.31) e (a) del presente teorema:

$$||BA||_1 \leq ||BU\sqrt{|A|}||_2\, ||\sqrt{|A|}||_2 \leq ||BU||\, ||\sqrt{A}||_2\, ||\sqrt{A}||_2 \leq ||B||\, ||\sqrt{A}||_2^2 = ||B||\, ||A||_1 \ :$$

Inoltre $AB = (U\sqrt{|A|})\,\sqrt{|A|}B \in \mathfrak{B}_1(H)$ perchè entrambi i fattori sono di Hilbert Schmidt e vale (a). In modo analogo al caso trattato sppra si tova che $||AB||_1 \leq ||B||\, ||A||_1$. La prova di (ii) in (b) sarà data nella prova della proposizione 6.5. □

Per concludere introduciamo il concetto di traccia di un operatore nucleare e mostriamo come la traccia abbia le stesse proprietà formali della traccia di matrici.

Proposizione 4.6. *Se $(H, (\,|\,))$ è uno spazio di Hilbert, $A \in \mathfrak{B}_1(H)$ e N è una base hilbertiana di H, è ben definita*

$$trA := \sum_{u \in N} (u|Au)\,, \tag{4.33}$$

inoltre:
(a) *trA non dipende dalla base hilbertiana scelta;*
(b) *per ogni coppia ordinata (B, C) di operatori di Hilbert-Schmidt tali che $A = BC$ vale:*

$$trA = (B^\dagger|C)_2\,; \tag{4.34}$$

(c) *$|A| \in \mathfrak{B}_1(H)$ e vale:*

$$||A||_1 = tr|A|\,. \tag{4.35}$$

Prova. (a) e (b). Ogni operatore di classe traccia può decomporsi come prodotto di due operatori di Hilbert-Schmidt come provato in (a) di teorema 4.6 Cominciamo con il notare che se $A = BC'$ con B, C operatori di Hilbert-Schmidt allora

$$(B^\dagger|C)_2 = \sum_{u\in N}(B^\dagger u|Cu) = \sum_{u\in N}(u|BCu) = \sum_{u\in N}(u|Au) = trA\,.$$

Questo prova che trA è ben definito essendo un prodotto scalare di Hilbert-Schmidt e che vale (4.34). (Quanto scritto sopra mostra anche che $(B^\dagger|C)_2 = (B'^\dagger|C')_2$ se $BC = B'C'$ essendo B, B', C, C' operatori di Hilbert-Schmidt.) Si osservi inoltre che il risultato ottenuto prova anche l'invarianza al veriare della base hilbertiana visto che $(\ |\ )_2$ non dipende dalla scelta della base hilbertiana.
La prova di (c) è immediata notando prima di tutto che dall'unicità della decomposizione polare vale $|\,(|A|)\,| = |A|$ e quindi $|A| \in \mathfrak{B}_1(H)$ se $A \in \mathfrak{B}_1(H)$ perchè $\sqrt{|\,(|A|)\,|} = \sqrt{|A|}$. Infine, scegliendo una base hilbertiana ottenuta completando in $Ker(|A|)$ la solita base hilbertiana di $\overline{Ran(|A|)}$ fatta di autovettori di $|A|$, $\{u_{\lambda,i}\}$, si ha

$$tr|A| = \sum_{\lambda\in sing(A)}\sum_{i=1}^{m_\lambda}(u_{\lambda,i}|\ |A|u_{\lambda,i}) = \sum_{\lambda\in sing(A)}\sum_{i=1}^{m_\lambda}\lambda = \sum_{\lambda\in sing(A)} m_\lambda\lambda = ||A||_1\,.$$

Questo conclude la dimostrazione. □

Definizione 4.5. *Sia H spazio di Hilbert e $A \in \mathfrak{B}_1(H)$. trA è detta* **traccia** *dell'operatore A.*

La seguente proposizione enuncia altre utili proprietà degli operatori nucleari, in particolare, esattamente come nel caso finitodimensionale, la ciclicità della traccia. Si osservi che, e ciò è importante per la applicazioni in fisica, non è richiesto che tutti gli operatori nella traccia siano di classe traccia.

Proposizione 4.6. *Sia H spazio di Hilbert. La traccia gode delle seguenti proprietà.*
(a) *Se $A, B \in \mathfrak{B}_1(H)$ e $\alpha, \beta \in \mathbb{C}$ allora:*

$$trA^\dagger = \overline{trA}\,, \tag{4.36}$$

$$tr(\alpha A + \beta B) = \alpha\, trA + \beta\, trB\,. \tag{4.37}$$

(b) *Se A è di classe traccia e B limitato oppure A e B sono emntrambi di classe traccia, allora*

$$trAB = trBA\,. \tag{4.38}$$

(c) *Se $A_1, A_2, \ldots A_n$ sono in $\mathfrak{B}(H)$ con almeno un operatore di classe traccia oppure almeno due degli A_k sono di Hilbert-Schmidt, allora vale la proprietà di* **ciclicità** *della traccia*

$$tr\,(A_1A_2\cdots A_n) = tr\,(A_{\sigma(1)}A_{\sigma(2)}\cdots A_{\sigma(n)})\,, \tag{4.39}$$

dove $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ *è una permutazione ciclica di* $(1, 2, \ldots, n)$.

Prova della proposizione 4.6 e di (ii) in (b) della proposizione 4.5.
(a) la prova è immediata dalla stessa definizione di traccia.
(b) Per dimostrare la tesi cominciamo con il provare che vale se A e B sono entrambi di Hilbert-Schmidt. Per (b) di proposizione 4.5 (4.38) è equivalente a dire che

$$(A^\dagger|B)_2 = (B^\dagger|A)_2 \,. \tag{4.40}$$

La prova di (4.40) è immediata usando la relazione, valida per tutti i prodotti scalari e le norme generate:

$$4(X|Y) = ||X+Y||_2^2 + ||X-Y||^2 - i||X+iY||^2 + i||X-iY||^2 \,,$$

e ricordando che, nel caso della norma di Hilbert-Schmidt, $||Z||_2 = ||Z^\dagger||_2$ ((i) in (a) di teorema 4.5).
Supponiamo ora che A sia di classe traccia e B sia limitato. Allora $A = CD$ con C e D di Hilbert-Schmidt per (a) del teorema 4.6. Inoltre DB e BC sono di Hilbert-Schmidt in quanto $\mathfrak{B}_2(H)$ è un ideale di $\mathfrak{B}(H)$. Quindi usando la proprietà di interscambio provata valida nel caso di due operatori di Hilbert-Schmidt

$$trAB = trC\,DB = trDB\,C = trD\,BC = trBC\,D = trBA \,.$$

(c) È chiaro che essendo $\mathfrak{B}_1(H)$ un ideale di $\mathfrak{B}(H)$ è sufficiente che un solo operatore tra $A_1, \ldots, A_n$ sia di classe traccia perchè lo sia il loro prodotto. In particolare, usando (a) del teorema 4.6 unitamente al fatto che anche $\mathfrak{B}_2(H)$ è un ideale di $\mathfrak{B}(H)$, si vede subito che se due operatori tra $A_1, \ldots, A_n$ sono di Hilbert-Schmidt allora il prodotto totale è di classe traccia. È chiaro che la (4.39) è equivalente a

$$tr\,(A_1 A_2 \cdots A_n) = tr\,(A_2 A_3 \cdots A_n A_1) \,, \tag{4.41}$$

infatti continuando a permutare un passo alla volta si ottengono tutte le permutazioni cicliche. Proviamo la (4.41). Supponiamo dapprima il caso in cui due operatori, A_i e A_j con $i < j$ siano di Hilbert-Schmidt. Se $i = 1$ la tesi segue subito da (b) con la coppia di operatori di Hilbert-Schmidt $A = A_1$ e $B = A_2 \cdots A_n$. Se $i > 1$ allora i 4 operatori (i) $A_1 \cdots A_i$, (ii) $A_{i+1} \cdots A_n$, (iii) $A_{i+1} \cdots A_n A_1$, (iv) $A_2 \cdots A_i$ sono necessariamente di Hilbert-Schmidt perchè contengono ciascuno A_i oppure A_j (mai entrambi), quindi vale

$$tr(A_1 \cdots A_n) = tr(A_1 \cdots A_i A_{i+1} \cdots A_n) = tr(A_{i+1} \cdots A_n A_1 A_2 \cdots A_i) = tr(A_2 \cdots A_i A_{i+1} \cdots A_n A_1),$$

che è quanto volevamo provare. Proviamo la proprietà di ciclicità nell'ipotesi che A_i sia di classe traccia. Se $i = 1$ la tesi segue immediatamente da (b) con $A = A_1$ e $B = A_2 \cdots A_n$. Supponiamo allora che $i > 1$. Allora $A_1 \cdots A_i$ e $A_2 \cdots A_i$ sono di classe traccia perchè ciascuno dei due contiene A_i, quindi

$$tr(A_1 \cdots A_n) = tr(A_1 \cdots A_i A_{i+1} \cdots A_n) = tr(A_{i+1} \cdots A_n A_1 A_2 \cdots A_i) = tr(A_2 \cdots A_i A_{i+1} \cdots A_n A_1).$$

La proprietà di ciclicità della traccia consente di provare (ii) in (b) della proposizione 5.5. In base a (4.35) si tratta di provare che $tr|A| = tr|A^\dagger|$. Dal corollario del teorema di decomposizione polare (teorema 3.9) sappiamo che $|A^\dagger| = U|A|U^\dagger$ dove $U|A| = A$ è la decomposizione polare di A. Vale allora

$$||A^\dagger||_1 = tr|A^\dagger| = tr(U|A|U^\dagger) = tr(U^\dagger U|A|) = tr|A| = ||A||_1 \ ,$$

dove abbiamo usato il fatto che $U^\dagger U|A| = |A|$ dato che U è isometria su $Ran(|A|)$ (teorema 3.9). $\square$

Parte II

Il formalismo della Meccanica Quantistica e la Teoria Spettrale

Capitolo 5

Brevi cenni di fenomenologia dei sistemi quantistici e di Meccanica Ondulatoria.

Nella prossima sezione daremo qualche idea generale di che cosa si intenda per sistema quantistico. Le parti successive di questo capitolo possono essre tralasciati dal lettore più matematico non interessato alla genesi dei concetti fisici della MQ. In tali paragrafi passeremo rapidamente ad enunciare in modo pittoto sommario alcuni dei fatti sperimentali e dei modelli teorici "protoquantistici" che hanno infine condotto alla formulazione prima della *meccanica ondulatoria* e poi della MQ vera e propria. Molti dettagli fisici si possono trovare in [12]. Ometteremo completamente ogni discussione su alcune importanti tappe di questo percorso storico: spettroscopia atomica, modelli atomici (Rutherford,Bohr, Bohr-Sommerfeld), esperimento di Franck-Hertz, per i quali si rimanda a testi di fisica (es. [12]). Il nostro succinto sommario metterà solo a giustificare il modello di base per la teoria matematica della MQ che svilupperemo partire dalla sezione 5.5.

5.1 Generalità sui sistemi quantistici.

Usiamo qui la parola *sistema fisico* in senso molto generico e discorsivo. È molto difficile definire dal punto di vista fisico cosa sia un *sistema quantistico*. Possiamo cominciare a dire che più di *sistema fisico quantistico* più opportuno parlare di sistema fisico dal *comportamento quantistico* facendo questa distinzione da un punto di vista più fenomenologico-sperimentale che teorico. All'interno della formulazione teorica della meccanica quantistica non esiste una precisa demarcazione tra sistemi classici e sistemi quantistici, se non una demarcazione imposta del tutto "a mano" come vedremo più avanti, e tale argomento è oggi, ancora di più che in passato, oggetto di discussione ed attività di ricerca teorica e sperimentale.
Parlando in modo del tutto generico, possiamo dire che hanno comportamento quantistico i sistemi della *microfisica*, cioè *molecole*, *atomi*, *nuclei* e *particelle subnucleari* quando presi in-

dividualmente. Sistemi fisici costituiti da più copie di tali sottosistemi (per esempio *cristalli*) possono avere a loro volta comportamento quantistico. Sistemi macroscopici hanno comportamento tipicamente quantistico solo in determinate condizioni difficili da realizzare (es. *condensazione di Bose-Einstein*, *L.A.S.E.R.*). In modo leggermente più preciso della rozza distinzione tra microsistemi e macrosistemi esposta sopra, possiamo dire che, generalmente parlando, quando un qualsiasi sistema fisico si comporta quantisticamente l'azione caratteristica del sistema, cioè il numero dalle dimensioni fisiche di *Energia* $\times$ *Tempo* (equivalentemente *Quantità di Moto* $\times$ *Lunghezza* ovvero *Momento angolare*) ottenuto combinando opportunamente le dimensioni fisiche caratteristiche del sistema (es. massa, velocita', dimensioni lineari ecc...) nei processi considerati è dell'ordine o inferiore al valore della **costante di Planck**

$$h = 6.6262 \cdot 10^{-34} Js.$$

La costante di Planck e la parola *quantum* che caratterizza il nome della MQ vennero introdotte per la prima volta da Planck nel 1900 nel suo lavoro sulla *teoria del corpo nero* per eliminare il problema dell'energia totale teoricamente infinita di un sistema fisico costituito da radiazione elettromagnetica in equilibrio termodinamico con le pareti di una cavità tenuta ad una temperatura fissata. La sua ipotesi teorica, rivelatasi poi esatta, prevedeva che la radiazione in potesse scambiare energia con le pareti in quantità proporzionali alle frequenza degli oscillatori atomici nelle pareti la cui costante universale di proporzionalità era la costante di Planck, tali quantità vennero chiamate in latino *quanta* o *quanti* in italiano. Tornando al criterio per discriminare i sistemi quantistici da quelli classici usando la costante h, consideriamo per esempio un elettrone legato ad un nucleo di idrogeno. Un'azione caratteristica dell'elettrone si ottiene come, per esempio il prodotto della massa dell'elettrone ($\sim 9 \cdot 10^{-31} Kg$), per una stima della velocità intorno al nucleo ($\sim 10^6 m/s$) per il valore del raggio di Bohr dell'atomo di idrogeno ($\sim 5 \cdot 10^{-11} m$). Il calcolo fornisce un'azione caratteristica dell'ordine di $4.5 \cdot 10^{-35} Js$ più piccolo dunque della costante di Planck. Ci si aspetta quindi che l'elettrone nell'atomo abbia comportamento quantistico, come infatti accade. Si può fare un calcolo analogo per sistemi fisici macroscopici come per esempio un pendolo di massa di qualche grammo e lunghezza dell'ordine del centimetro che oscilla sottoposto alla forza di gravità. Un'azione caratteristica di questo sistema può essere definita come il prodotto dell'energia cinetica massima per il periodo di oscillazione. Si trovano, per il sistema in questione valori dell'azione caratteristica di moltissimi ordini di grandezza superiori ad h.

Osservazione. Una delle proprietà peculiare dei sistemi quantistici è il fatto che lo spettro dei valori assumibili dalle grandezze fisiche attribuibili ad essi è generalmente diverso (almeno in alcuni casi importanti) dallo spettro dei valori delle stesse grandezze assumibili da sistemi macroscopici analoghi. In certi casi, la differenza è sorprendente perché si passa da uno spettro *continuo* di valori possibili nel caso classico, ad uno spettro *discreto*, nel caso quantistico. È importante però precisare la discretezza dello spettro dei valori assumibili da una generica grandezza fisica quantistica *non* è una caratteristica necessaria in MQ: esistono grandezze quantistiche con spettro di valori continuo anche in MQ. Questo fraintendimento è all'origine o a volte conseguenza di una delle interpretazioni più frequenti, ma riduttiva, dell'aggettivo *quantistica*

nella parola Meccanica Quantistica.

5.2 Alcune proprietà particellari delle onde elettromagnetiche.

Le onde elettromagnetiche (e quindi in particolare la luce) in particolari cisrcostanze sperimentali rivelano comportamenti tipici di quelli di *insiemi di particelle*. Nella descrizione matematica di tali comportamenti classicamente anomali è coinvolta la costante di Planck. Possiamo citare due esempi di comportamento classicamente anomalo che hanno svolto un ruolo fondamentale nello sviluppo iniziale della MQ per i corrispondenti modelli teorici protoquantistici costruiti nel tentativo di darne una spiegazione: l' *effetto fotoelettrico* e l'*effetto Compton*.

5.2.1 Effetto Fotoelettrico.

L'effetto fotoelettrico riguarda l'emissione di elettroni (corrente elettrica) da parte di metalli in seguito ad illuminazione con un'onda elettromagnetica. Tale effetto era noto dalla prima metà del secolo XIX e alcune sue caratteristiche risultavano del tutto inspiegabili alla luce della teoria classica dell'interazione onde elettromagentica e materia [12]. In particolare era inspiegabile la presenza di un valore minimo della frequenza della luce usata per irraggiare il metallo al di sotto del quale non si ha emissione elettronica. Il valore di tale soglia dipende dal metallo usato. Una volta raggiunta la soglia, l'istantaneità del processo di emissione era del tutto inspegabile.
La teoria classica implica che l'emissione di elettroni ci deve essere indipendentemente dalla frequenza usata pur di aspettare un tempo sufficientemente lungo in modo da fare assorbire agli elettroni del metallo sufficiente energia da superare l'energia di legame con gli atomi.
Nel 1905 A. Einstein propose un modello, per l'epoca molto ardito, che rendeva conto di tutte le proprietà anomale dell'effetto fotoelettrico[1] con una precisione sperimentale notevole. Seguendo Planck, l'ipotesi centrale di Einstein era che un onda elettromagnetica *monocromatica* cioé con frequenza definita, che indicheremo con ν, sarebbe in realtà costituita da particelle materiali, dette *quanti di luce*, ciascuna dotata di energia

$$E = h\nu \,. \tag{5.1}$$

In tale modello l'energia totale dell'onda sarebbe quindi pari alla somma delle energie dei singoli quanti di luce.
Tutto ciò era ed è in contrasto totale con la teoria dell'elettromagnetismo classico in cui un'onda elettromagnetica è un sistema continuo la cui energia è legata all'*ampiezza* dell'onda invece che alla *frequenza*. Quello che dunque accadrebbe nell'effetto fotoelettrico secondo il modello di Einstein è che, illuminando il metallo con un onda monocromatica, ciascuno dei pacchetti energetici associati all'onda possa essere assorbito da un elettrone esterno di ogni atomo del metallo. Più precisamente l'ipotesi di Einstein per spiegare la fenomenologia sperimentale, era

[1] Einstein ricevette il Premio Nobel proprio per tale lavoro.

che il pacchetto possa essere o assorbito *completamente* oppure *non* assorbito affatto, senza possibilità di assorbimenti parziali. Se e solo se l'energia del quanto è uguale o superiore all'energia dell'elettrone di legame con l'atomo E_0 (dipendente dal metallo e misurabile indipendentemente dall'effetto fotoelettrico), l'elettrone viene *istantaneamente* espulso, incamerando in energia cinetica l'eventuale energia in eccesso del quanto. La frequenza $\nu_0 := E_0/h$ individuerebbe così la soglia sulla frequenza osservata sperimentalmente. Tale ipotesi si rivelò in perfetto accordo con dati sperimentali

5.2.2 Effetto Compton.

L'effetto Compton risale al 1923. Esso riguarda la diffusione di onde elettromagnetiche di altissima frequenza – raggi X ($> 10^{17} Hz$) e raggi γ ($> 10^{18} Hz$) – da parte di materia estesa (gas, liquidi e solidi). Schematizzando al massimo l'effetto sperimentale consiste in questo. Irraggiando la sostanza (che diremo *ostacolo*) con un'onda elettromagnetica piana propagante nella direzione z e monocromatica di lunghezza d'onda λ, si osserva un'onda diffusa dall'ostacolo con diverse componenti. Una componente è diffusa in tutte le direzioni ed ha la stessa lunghezza d'onda della radiazione incidente, l'altra componente ha una lungezza d'onda leggermente superiore a λ, $\lambda(\theta)$ che dipende dall'angolo di osservazione θ. Se più precisamente, θ è definito come l'angolo tra la direzione di propagazione dell'onda incidente z e la direzione di propagazione dell'onda diffusa oltre l'ostacolo con lunghezza d'onda $\lambda(\theta)$, misurando θ a partire dall'asse z, vale la relazione:

$$\lambda(\theta) = \lambda + f\,(1 - cos\theta) \tag{5.2}$$

dove[2] la costante f ha le dimensioni di una lunghezza e si ottiene dai dati sperimentali. Il suo valore è $f = 0.024(\pm 0.001)$ Å.
La teoria elettromagnetica classica era ed è assolutamente incapace di spiegare un tale fenomeno. Tuttavia come Compton dimostrò, esso poteva essere spiegato assumendo le tre ipotesi seguenti del tutto incompatibili con la teoria classica ma in accordo con l'ipotesi dei quanti di luce di Einstein.
(a) L'onda elettromagnetica è costituita da particelle che trasportano energia, esattamente come aveva ipotizzato Einstein, secondo la legge (5.1).
(b) Ogni quanto di luce possiede anche un'impulso

$$\mathbf{p} := \hbar \mathbf{k}\,, \tag{5.3}$$

dove $\mathbf{k}$ è il *vettore d'onda* dell'onda elettromagnetica associata al quanto (vedi sotto).
(c) I quanti di luce interagiscano, in un processo d'urto (in generale in regime relativistico), con gli atomi e con gli elettroni più esterni degli atomi dell'ostacolo soddisfacendo le leggi di conservazione dell'impulso e dell'energia.

[2] Ricordiamo che 1 Å$= 10^{-10} m$

A commento di (c), ricordiamo che il vettore vettore d'onda $\mathbf{k}$ associato ad un onda piana monocromatica ha, per definizione, direzione e verso pari a quello di propagazione dell'onda e modulo dato da $2\pi/\lambda$, dove λ è la lunghezza d'onda dell'onda. Equivalentemente, se ν è la frequenza dell'onda,

$$|\mathbf{k}| = 2\pi/\lambda = 2\pi\nu/c\,, \tag{5.4}$$

dove si è usata la ben nota relazione

$$\nu\lambda = c\,, \tag{5.5}$$

valida per le onde elettromagnetiche ed essendo $c = 2.99792 \cdot 10^8 m/s$ è la velocità della luce.

Possiamo dare qualche dettaglio in più per il lettore interessato. Le leggi di conservazione dell'energia e dell'impulso da usarsi in regime relativistico, cioè quando (qualcuna) delle velocità coinvolte sono dell'ordine di c, si scrivono rispettivamente, nelle ipotesi fatte:

$$m_e c^2 + h\nu = \frac{m_e c^2}{\sqrt{1 - v^2/c^2}} + h\nu(\theta)\,, \tag{5.6}$$

$$\hbar\mathbf{k} = \frac{m_e \mathbf{v}}{\sqrt{1 - v^2/c^2}} + \hbar\mathbf{k}(\theta)\,. \tag{5.7}$$

A primo membro compaiono le quantità precedenti dell'urto a secondo membro quelle successive all'urto. $m_e = 9.1096 \cdot 10^{-31} Kg$ è la massa dell'elettrone. L'elettrone è pensato in quiete prima dell'urto con il quanto di luce. In seguito all'urto il quanto di luce viene diffuso nella direzione individuata da θ, mentre l'elettrone acquista velocità $\mathbf{v}$. Il vettore d'onda precedente all'urto, $\mathbf{k}$, è quindi nella direzione (generica ma fissata) z mentre il vettore d'onda del quanto di luce dopo l'urto, $\mathbf{k}(\theta)$, forma un angolo θ con $\mathbf{k}$. Dalla (5.7), tenendo conto della definizione del vettore d'onda, si ricava subito

$$\frac{m_e^2 c^2}{1 - v^2/c^2} = \frac{h^2\nu^2}{c^2} + \frac{h^2\nu(\theta)^2}{c^2} - 2\frac{h\nu}{c}\frac{h\nu(\theta)}{c}\cos\theta\,.$$

Eliminando ν tra la relazione appena ottenuta e la (5.6) si ricava immediatamente

$$\nu(\theta) = \nu - \frac{h\nu\nu(\theta)}{m_e c^2}\,(1 - cos\theta)\,. \tag{5.8}$$

Tenendo conto della relazione (5.1) e $\nu = c/\lambda$, si ricava facilemente la (5.2) nella forma

$$\lambda(\theta) = \lambda + \frac{h}{m_e c}\,(1 - cos\theta) \tag{5.9}$$

da cui si evince che $f = h/(m_e c)$ il cui valore numerico coincide proprio con quello ottenuto sperimentalmente quando si sostituiscono i valori numerici di h, m_e e c. Si osservi anche, che nel limite formale per $m_e \to +\infty$, la (5.9) fornisce $\lambda(\theta) \to \lambda$. Ciò porta anche a spiegare la componente dell'onda diffusa isotropicamente senza variazione della lunghezza d'onda (rispetto alla lunghezza d'onda della radiazione incidente), come dovuta a quanti di luce che interagiscono con particelle di massa molto più grande di quella elettronica (un atomo del materiale o tutto l'ostacolo nel suo complesso).

Osservazione. Il modello di Einstein e quello di Compton spiegano perfettamente l'effetto fotoelettrico e la formula (5.2) in termini quantitativi e qualitativi. Tuttavia devono considerarsi come modelli *ad hoc* slegati ed addirittura in contrasto con il corpus della fisica teorica dell'epoca: l'idea della costituzione corpuscolare delle onde elettromagnetiche e quindi della luce, come noto dai tempi di Newton e Huygens, non può spiegare effetti ondulatori notissimi della luce quali i fenomeni di *interferenza* e *diffrazione*. In qualche modo il modello ondulatorio e quello corpuscolare della luce (onde elettromagnetiche) devono *coesistere* nella realtà, ma ciò non è possibile nel paradigma della fisica classica. Ciò invece accade nella formulazione relativistica completa della MQ introducendo il concetto di *fotone*.

5.3 Cenni di Meccanica ondulatoria.

In queste dispense non ci occuperemo più delle proprietà quantistiche della luce che richiederebbero uno sviluppo ulteriore del formalismo della MQ che presenteremo. Le idee esposte sopra sui primi tentativi di descrizione quantistica della luce furono però utili, ribaltando il punto di vista, per arrivare alla formulazione della *meccanica ondulatoria*, il primo passo verso la formulazione della MQ.
La *meccanica ondulatoria* è una delle due prime versioni un pò rudimentali della MQ *per particelle dotate di massa*[3]. Nel seguito ci occuperemo molto succintamente di esporre alcune idee portanti della meccanica ondulatoria che mettono in luce alcuni punti fondamentali e generali che useremo per costruire il formalismo proprio della MQ. In particolare tralasceremo alcuni risultati storicamente connessi quali l'equazione stazionaria di Schrödinger e la spiegazione, tramite essa, dello spettro energetico dell'atomo di idrogeno. Torneremo più avanti su questi argomenti dopo avere costruito il formalismo.

5.3.1 Onde di de Broglie.

Un quanto di luce di Einstein e Compton è associato ad un'onda elettromagnetica piana monocromatica di numero d'onda $\mathbf{k} = \mathbf{p}/\hbar$ e *pulsazione* $\omega = E/\hbar$, dove lo ricordiamo, la pulsazione è semplicemente 2π volte la frequenza. Ogni componente dell'onda piana elettromagnetica (una componente del campo elettrico o magnetico che vibra perpendicolarmente a $\mathbf{k}$) ha allora la forma di un'onda scalare:

$$\psi(t, \mathbf{x}) = Ae^{i(\mathbf{k}\cdot\mathbf{x} - t\omega)} \,. \tag{5.10}$$

In realtà solo la parte reale dell'onda di sopra ha senso fisico. Un onda a valori complessi invece che reali del tipo di sopra ha comunque interesse fisico perché appare nella decomposizione di Fourier (vedi sezione 3.6) di una soluzione generale dell equazioni del campo elettromagnetico (equazioni di Maxwell) e più generalmente dell'equazione di d'Alembert. In termini di impulso

[3] L'altra versione sviluppata parallelamente da Heisenberg consisteva nella cosiddetta *meccanica delle matrici* [12] della quale non ci occuperemo.

e energia del quanto di luce, la stessa onda può riscriversi:

$$\psi(t, \mathbf{x}) = Ae^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x} - tE)} \,. \tag{5.11}$$

Dove lo si noti compaiono solo l'impulso e l'energia del quanto di luce. Nel 1924 de Broglie fece un'ipotesi estremamente rivoluzionaria: come alle onde elettromagnetiche si potevano associare, almeno in certi contesti sperimentali, delle particelle, si doveva poter associare delle onde alle particelle materiali della stessa forma (5.11) dove ora però, $\mathbf{p}$ e E sono rispettivamente l'impulso e l'energia (cinetica) della particella. Non era affatto chiara la natura di tali fantomatiche onde. La lunghezza d'onda associata ad una particella di impulso $\mathbf{p}$:

$$\lambda = h/|\mathbf{p}| \,, \tag{5.12}$$

viene detta *lungezza d'onda di de Broglie della particella.*
Nel 1927 si ebbe evidenza sperimentale di onde in qualche modo associate al comportamento di particelle materiali, più precisamente elettroni, da parete di due distinti esperimenti da parte di Davisson e Germer e G.P. Thompson.
Senza entrare nei dettagli degli esperimenti diremo solo quanto segue. È noto che quando un'onda (elettromagnetica, sonora ecc...) incontra un ostacolo con una struttura interna di dimensioni tipiche *dell'ordine o superiori* alla lunghezza d'onda dell'onda incidente, l'onda diffusa dall'ostacolo produce un fenomeno detto *diffrazione*: le varie parti interne dell'ostacolo agiscono sull'onda creando interferenza costruttuva e distruttiva, in modo che l'onda diffusa produca, su uno schermo su cui incide, delle figure costituite da zone d'ombra e zone di maggiore intensità (cioè più *chiare* nel caso di onde luminose). Tali figure si dicono *figure di diffrazione.* Se l'ostacolo è un cristallo, dalle figure di diffrazione si può risalire alla struttura interna di cristalli.

Osservazione. È importante sottolineare che il fenomeno della diffrazione è strettamente dovuto alla natura ondulatoria dell'onda (al fatto che ci sia qualcosa che oscilli ed al principio di sovrapposizione). Non è possibile produrre figure di diffrazione sparando delle particelle, che soddisfino le usuali leggi della meccanica classica, su qualsivoglia ostacolo.

Tornando agli esperimenti di Davisson e Germer e G.P. Thompson, essi ottennero figure di diffrazione prodotte da fasci di elettroni sparati su cristalli. Più precisamente le figure di diffrazione si produssero su uno schermo dall'*accumularsi delle tracce puntiformi degli elettroni* diffusi da un cristallo di passo reticolare dell'ordine di 1 Å. Il fatto ancora più notevole che corroborava l'ipotesi di de Broglie era il fatto che, negli esperimenti detti, le figure di diffrazione apparivano solo se la lunghezza di de Broglie degli elettroni era dell'ordine di 1Å o inferiore, esattamente come nel fenomeno della diffrazione prodotto da onde "usuali" su cristalli con passo reticolare di 1 Å.

5.3.2 Funzione d'onda di Schrödinger e interpretazione probabilistica di Born.

Nel 1926, in due famosi e geniali articoli, Schrödinger prese sul serio l'ipotesi di de Broglie e fece un'assunzione più precisa: quella di associare ad una particella materiale non un'onda

piana come (5.11), ma un pacchetto d'onde costituito dalla sovrapposizione (nel senso della trasformata di Fourier, vedi sezione 3.6) di onde piane di de Broglie. Nel caso di particelle *libere* cioé la cui energia è data dalla sola energia cinetica, l'onda di Schrödinger si scrive:

$$\psi(t,\mathbf{x}) = \int_{\mathbb{R}^3} \frac{e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-tE(\mathbf{p}))}}{(2\pi\hbar^2)^{3/2}}\,\hat{\psi}(\mathbf{p})\,d^3\mathbf{p}\,, \tag{5.13}$$

dove $E(\mathbf{p}) := \mathbf{p}^2/(2m)$ essendo m la massa della particella. Nei due articoli Schrödinger osservò che l'ottica geometrica si basa su un'equazione, detta equazione dell'iconale, che ha una struttura formalmente analoga all'equazione di Hamilton-Jacobi per una particella della meccanica classica. Schrödinger tentò allora di determinare un equazione fondamentale di una *meccanica ondulatoria* per la materia rispetto alla quale l'equazione di Hamilton-Jacobi della meccanica giocasse lo stesso ruolo approssimato che equazione dell'iconale (ottica geometrica) gioca rispetto all'equazione di d'Alembert (ottica ondulatoria) [12]. In questo modo pervenne alla notissima *equazione di Schrödinger* che gioca un ruolo centrale nella MQ. Noi ricaveremo tale equazione più avanti dopo avere costruito il formalismo. Nel caso generale di una particella soggetta ad una forza dovuta ad un potenziale U, $\mathbf{f}(t,\mathbf{x} = -\nabla U(t,\mathbf{x})$, l'equazione era

$$i\hbar\frac{\partial\psi(t,\mathbf{x})}{\partial t} = \left[-\frac{\hbar^2}{2m}\Delta + U(t,\mathbf{x})\right]\psi(t,\mathbf{x}) \tag{5.14}$$

dove $\Delta := \nabla\cdot\nabla$ è l'operatore di Laplace in $\mathbb{R}^3$.
L'onda di de Broglie-Schroedinger ψ venne chiamata *funzione d'onda* della particella a cui è associata.
L'interpretazione oggi accettata della funzione d'onda ψ – almeno nell'interpretazione standard ("di Copenhagen") del formalismo della MQ – venne data Born sempre nel 1926:
fissato $t \in \mathbb{R}$,

$$\rho(t,\mathbf{x}) := \frac{|\psi(t,\mathbf{x})|^2}{\int_{\mathbb{R}^3}|\psi(t,\mathbf{y})|^2 d^3\mathbf{y}}$$

è la densità di probabilità di trovare la particella nel punto $\mathbf{x}$ quando si esegua un esperimento per determinarne la posizione.
L'interpretazione di Born che si rivelò in accordo con l'esperienza, era già in sostanziale accordo con i risultati degli esperimenti di Davisson e Germer e G.P. Thompson in cui le tracce delle particelle sullo schermo si addensavano nelle regioni in cui $\rho(t,\mathbf{x}) > 0$ ed erano assenti nelle regioni in cui $\rho(t,\mathbf{x}) = 0$ dando luogo alle figure di diffrazione. (Il fatto che le onde piane di de Broglie non abbiano diretto significato fisico alla luce dell'interpretazione di Born avendo integrale del modulo quadro divergente è fisicamente irrilevante. Le onde piane monocromatiche usate per interpretare i risultati sperimentali secondo l'ipotesi di de Broglie possono essere approssimate a piacimento da pacchetti a quadrato sommabile con una distribuzione $\hat{\psi}(\mathbf{p})$ arbitrariamente stretta attorno ad un valore $\mathbf{p}_0$ che determina con approssimazione piccola a piacere la lunghezza d'onda $\lambda_0 = |\mathbf{p}_0|/h$ di de Broglie degli elettroni.)

Osservazioni.
(1) Dal punto di vista matematico, l'assunzione dell'interpretazione di Born richiede che abbiano senso fisico solo le funzioni d'onda a *quadrato sommabile* e *non quasi ovunque nulle*, cioè gli elementi *non nulli* di $L^2(\mathbb{R}^3 d^3\mathbf{x})$: uno *spazio di Hilbert* entra in gioco per la prima volta, nella costruzione della MQ.

(2) Assumendo l'interpretazione di Born si deve concludere che, quando non viene eseguito alcun esperimento per determinarne la posizione, la particella associata alla funzione d'onda ψ, non può evolvere temporalmente secondo le leggi della meccanica classica: se seguisse una traiettoria regolare come prescritto dalla meccanica classica, la funzione $|\psi|^2$ associata dovrebbe essere quasi ovunque nulla fuori dalla traiettoria. Tuttavia avendo ogni traiettoria regolare in $\mathbb{R}^3$ misura nulla, $|\psi|^2$ dovrebbe essere nulla quasi ovunque in $\mathbb{R}^3$ portando ad una contraddizione. In altre parole, quando non si esegue un esperimento per determinare la posizione della particella, la particella non deve considerarsi un oggetto classico, ma la sua evoluzione temporale è descritta dall'evoluzione dell'onda ψ (secondo l'equazione di Schrödinger).

(3) Se si ammette, come nell'interpretazione di Copenhagen, che la funzione d'onda ψ sia una descrizione *completa* dello stato fisico della particella, si deve concludere che la posizione della particella sia *fisicamente* indefinita prima di eseguire un esperimento che la determini e che venga fissata all'*atto dell'esperimento* in modo probabilistico. Non è corretto pensare che l'uso di una descrizione probabilistica sia dovuto ad insufficiente conoscenza dello stato del sistema: "la posizione esiste ma noi non la conosciamo". Nell'interpretazione di Copenhagen la posizione *non* esiste fino a quendo non si esegue un esperimento per determinarla e lo stato della particella (l'informazione massima sulle sue proprietà fisiche variabili nel tempo) è descritto dall'onda ψ. Nella meccanica ondulatoria la particella quantistica ha dunque una duplice natura di *onda-corpuscolo*, ma le due nature *non vengono mai in contrasto perché non si manifestano mai contemporaneamente.*

5.4 Principio di indeterminazione di Heisenberg.

Quando cerchiamo di valutare sperimentalmente il valore di qualsiasi grandezza definita per un sistema fisico interagendo con il sistema stesso, possiamo alterare lo stato del sistema. Tuttavia nella descrizione classica, in linea di principio possiamo rendere piccola a piacere la perturbazione apportata allo stato. Nel 1927 Heisenberg si rese conto che le ipotesi di Planck, Einstein, Compton, de Broglie avevano una conseguenza (anche epistemologica) notevole in questo contesto. In termini quantitativi il principio di Heisenberg invece dice che, considerado sistemi quantistici e particolari grandezze, non è sempre possibile rendere piccola a piacere la perturbazione apportata allo stato del sistema: la costante di Planck rappresenta un limite invalicabile per il prodotto delle indeterminazioni di alcune coppie di grandezze.
Più precisamente, esaminando vari esperimenti ideali in cui si assumeva qualcuna delle ipotesi

dei modelli di Planck, Einstein, Compton e de Broglie, Heisenberg arrivò alla conclusione che:

cercando di determinare la posizione o l'impulso di una particella lungo un fissato asse x, se ne altererà l'impulso o rispettivamente la posizione lungo lo stesso asse *in modo tale che il prodotto della minima larghezza possibile delle due incertezze Δx e Δp con cui alla fine risultano essere note le due grandezze soddisfi:*

$$\Delta x \Delta p \quad \sim \quad h \,, \tag{5.15}$$

senza poter andare sotto tale stima. Se la posizione e l'impulso sono valutiati lungo assi reciprocamente ortogonali il prodotto delle incertezze può essere arbitrariamente piccolo.

(5.15) esprime il *principio di indeterminazione di Heisenberg* per la posizione e l'impulso.
Una relazione analoga vale per l'incertezza ΔE con cui è determinata l'energia E di una particella e l'incertezza Δt con cui si determina il tempo t in cui viene fatta la misura dell'energia[4]

$$\Delta E \Delta t \quad \sim \quad h \,. \tag{5.16}$$

A titolo di esempio consideriamo l'esperimento ideale in cui si cerca di determinare la posizione X di un elettrone di momento inizialmente noto P illuminandolo con un fascetto di luce monocromatica di lughezza d'onda λ che si propaga nella direzione x. Immaginiamo di voler leggere la posizione su uno schermo parallelo all'asse X tramite una lente interposta tra l'asse X e lo schermo. Un quanto di luce che ha interagito con l'elettrone produrrà un'immagine X' colpendo schermo dopo avere attraversato la lente. A causa dell'apertura finita della lente è non possibile conoscere con esattezza la direzione lungo la quale è stato diffuso il quanto di luce che produce l'immagine X' sullo schermo. Dall'ottica ondulatoria sappiamo che in X' si avrà una figura di diffrazione che permetterà di raggiungere nella misura della coordinata X una accuratezza non più piccola di

$$\Delta X \sim \frac{\lambda}{\sin \alpha} \,,$$

dove α è il semi angolo sotto il quale da X si vede la lente. Al quanto di luce corrisponde l'impulso h/λ, per cui l'incertezza della componente P_x del quanto di luce diffuso sarà data approssimativamente da $h(\sin \alpha)/\lambda$. L'impulso totale del sistema particella, quanto di luce, microscopio, dovrà rimanere costante per cui l'incertezza nella componente x della quantità di moto della particella dopo la diffusione del quanto di luce dovrà essere uguale alla corrispondente incertezza per il quanto di luce stesso:

$$\Delta P_x \sim \frac{h}{\lambda} \sin \alpha \,.$$

Il prodotto delle due indeterminazioni della particella lungo l'asse x risulta essere dunque al minimo

$$\Delta X \Delta P_x \sim h \,.$$

[4] la relazione d'indeterminazione tempo-energia ha uno status ed una interpretazione molto più problematico delle analoghe relazioni posizione-impulso.

Osservazione. Il principio di Heisenberg, a questo livello, ha la stessa (in)consistenza logica dei modelli protoquantistici usati euristicamente per arrivarne alla formulazione. Deve essere visto più come una *ipotesi di lavoro* per costruire una nuova nozione di particella in cui le nozioni di posizione ed impulso classici abbiano senso solo entro limiti fissati dall principio stesso: *gli stati fisici di una particella quantistica devono essere tali per cui l'impulso o la posizione non siano definite e definibili contemporaneamente.*
Val la pena di ricordare che nella formulazione finale della MQ, come vedremo, il *principio* di Heisenberg diventa un *teorema*.

5.5 Le grandezze compatibili ed incompatibili.

Indipendentemente dal principio di Heienberg, la fenomenologia quantistica mostra che ci sono coppie grandezze A e B che risultano essere *incompatibili*. Questo significa che, se si misura prima la grandezza A sul sistema ottenendo il risultato a e quindi – immediatamente dopo – si misura B ottenendo il risultato b, una successiva misura di A – vicina temporalmente a piacere a quella di B (in modo da non poter imputare il risultato all'evoluzione temporale) – produce un risultato a_1 generalmente differente da a. Lo stesso fenomeno appare scambiando il ruolo di A e B.
Coppie di grandezze incompatibili sono la componente della posizione e dell'impulso lungo un fissato, tali grandezze soddisfano anche il principio di Heinsenberg. Questi due fatti sono connessi, ma il legame può essere spiegato chiaramente solo dopo avere costruito il formalimo completo.
In generale coppie di grandezze incompatibili non soddifano il principio di Heisenberg.
È importante sottolineare che la fenomenologia quantistica mostra che esitono anche coppie di grandezze A', B' *compatibili*. Questo significa che, se si misura prima la grandezza A sul sistema ottenendo il risultato a' e quindi – immediatamente dopo – si misura B' ottenendo il risultato b', una successiva misura di A' – vicina temporalmente a piacere a quella di B' (in modo da non poter imputare il risultato all'evoluzione temporale) – produce lo stesso risultato a'. Lo stesso fenomeno appare scambiando il ruolo di A' e B'.

Un esempio di coppie di grandezze incompatibili che non soddisfano il principio di Heisenberg sono due componenti diverse dello **spin** di una particella. Lo spin dell'elettrone (e poi di tutte le particelle nucleari e subnucleari) fu introdotto da Goudsmit e Uhlembeck nel 1925 per spiegare alcune proprietà "anomale", il cosiddetto *effetto Zeeman anomalo*, degli spettri energetici (righe spettrali) degli atomi di metalli alcalini. In senso semiclassico lo spin rappresenta il momento angolare intriseco dell'elettrone che, per certi aspetti, può essere pensato come dovuto ad un incessante moto rotatorio dell'elettrone attorno al proprio centro di massa. Tuttavia questa interpretazione è molto fuorviante è non deve essere presa alla lettera. Un signifcato profondo allo spin può solo essere dato nel contesto della definizione, dovuta a Wigner, di particella elementare quantistica come sistema elementare invariante sotto l'azione del gruppo di Poincaré.
Allo spin è associato un momento magnetico intrinseco che è alla fine il responsabile dell'effet-

to Zeeman anomalo osservato. Tuttavia lo spin è una grandezza vettoriale con caratteristiche peculiari quantistiche che lo differenziano completamente da un momento angolare classico e lo fanno rientrare nella categoria dei momenti angolari quantistici. La prima differenza è nei valori assumibili dal modulo dello spin. Nell'unità di misura $\hbar$ tali valori sono sempre numeri del tipo $\sqrt{s(s+1)}$ dove s è un intero fissato dipendente dal tipo di particella, $s = 1/2$ per l'elettrone. Ciascuna delle tre componenti dello spin rispetto ad una terna di assi destrorsa ortonormale può assumereciscuno dei $2s+1$ valori discreti $-s, -s+1, \ldots, s-1, s$. Le tre componenti dello spin risultano essere grandezze fisiche *incompatibili* nel senso detto sopra: la misurazione alternata e fatta a tempi vicinissimi di due differenti componenti dello spin fornisce valori sempre differenti per tali componenti dello spin. È importante precisare che la stessa incompatibilità si ha anche per le componenti del momento angolare orbitale e totale di una particella.

Un esempio di coppie di grandezze compatibili per una particella quantistica sono la componente x dell'impulso e la componente y del vettore posizione. Risulta anche che ogni grandezza fisica è compatibile con se stessa. Tale fatto non è per nulla banale.

Capitolo 6

La costruzione del formalismo matematico della MQ: il teorema di Gleason.

La Matematica in Fisica è come il maiale: non si butta mai via niente.

6.1 Le idee che stanno alla base dell'interpretazione standard della fenomenologia quantistica.

Cerchiamo di riassumere alcuni punti cruciali del comportamento dei sistemi quantistici.

QM1. (i) *I risultati delle misure di grandezze fisiche su sistemi quantistici, di cui è fissato lo stato, hanno esiti probabilistici. Non è possibile prevedere l'esito della misura, ma solo (almeno per quanto riguarda la posizione di una particella nell'interpretazione di Born) la probabilità che l'esito fornisca un certo risultato.*

(ii) Tuttavia se una grandezza è stata misurata ed ha prodotto un certo risultato, ripetendo la misura della stessa grandezza immediatamente dopo la prima misura (per evitare che nel frattempo lo stato del sistema evolva), il risultato della misura sarà lo stesso già ottenuto.

QM2. (i) *Esistono delle grandezze fisiche incompatibili nel senso che segue. Dette A e B due grandezze incompatibili, se si misura prima la grandezza A sul sistema (a stato fissato) ottenendo il risultato a e quindi – immediatamente dopo – si misura B ottenendo il risultato b, una successiva misura di A – vicina temporalmente a piacere a quella di B (in modo da non poter imputare il risultato all'evoluzione temporale dello stato) – produce un risultato a_1 generalmente differente da a. Lo stesso fenomeno appare scambiando il ruolo di A e B.*

(ii) Esistono anche grandezze fisiche compatibili nel senso che segue. Dette A' e B' due grandezze compatibili, se si misura prima la grandezza A' sul sistema (a stato fissato) ottenendo

il risultato a' e quindi – immediatamente dopo – si misura B' ottenendo il risultato b', una successiva misura di A' – vicina temporalmente a piacere a quella di B' (in modo da non poter imputare il risultato all'evoluzione temporale) – produce lo stesso risultato a' già ottenuto. Lo stesso fenomeno appare scambiando il ruolo di A' e B'. In particolare, per (ii) di **QM1**, *ogni grandezza fisica è compatibile con se stessa.*

Osservazioni.
(1) È chiaro che non è possibile avere la certezza assoluta che i sistemi fisici quantistici soddisfino (i) di Q1. Si potrebbe infatti sospettare che la stocasticità dell'esito delle misure derivi dal fatto che, in realtà, gli sperimentatori non conoscano completamente lo stato del sistema e se lo conoscessero completamente potrebbero prevedere con certezza gli esiti delle misure. Nell'interpretazione standard della MQ, la cosiddetta *interpretazione di Copenhagen*, la stocasticità degli esiti delle misure è assunta come un fatto fondamentale della natura dei sistemi quantistici. Esistono comunque interessanti tentativi di interpretazione della fenomenologia quantistica basati su formalismi alternativi (le cosiddette formulazioni in termini di *variabili nascoste*) [5]. In tali approcci la stocasticità è spiegata come dovuta ad una incompleta informazione da parte dello sperimentatore sul reale stato fisico del sistema che è descritto da più variabili (ed in modo molto diverso) rispetto a quelle usate nella formulazione standard. Nessuno di tali tentativi (malgrado alcuni siano veramente notevoli come la teoria di Bohm) è risultato fino ad oggi veramente soddisfacente tanto da essere un vero competitore dell'interpretazione e della formulazione standard della MQ (quando si considerino anche le teorie quantistiche relativistiche e la teoria dei campi quantistica relativistica in particolare).
Bisogna anche sottolineare che *non* è comuque possibile costruire una teoria fisica *completamente classica* (includendo le teorie relativistiche non quantistiche tra le teorie classiche) che possa spiegare completamente la fenomenologia quantistica. Le variabili nascoste devono soddisfare un requisito di *contestualità* piuttosto inusuale nelle teorie classiche. Infine ogni teoria che spieghi la fenomenologia quantistica, inclusa la stessa MQ, deve essere *non locale* [5]. In altre parole la fenomenologia sperimentale prova l'esistenza di correlazioni tra esiti di esperimenti che avvengono in regioni distinte dello spazio ed in periodi di tempo per cui non è possibile che vi sia trasmissione di informazione tramite qualunque mezzo fisico che si muova ad una velocità non superiore a quella della luce nel vuoto.
(2) È implicito in MQ1 e MQ2 che i sistemi fisici interessanti nella fenomenologia e teoria quantistica vengano divisi in due grandi categorie: gli *strumenti di misura* ed i *sistemi quantistici*. Nella formulazione di Copenhagen si assume che gli strumenti di misura siano sistemi che soddisfino alle leggi della fisica classica. Queste assunzioni che corrispondono a dati di fatto sperimentali e che sono a dir vero piuttosto rozze dal punto di vista teorico, sono alla base dell'interpretazione del formalismo e non si riesce a dire molto altro nel corpus della formulazione standard. Per esempio non è chiaro dove si situi la linea di confine tra sistemi classici e quantistici, nemmeno come questa linea possa essere descritta all'interno del formalismo e nemmeno se il sistema costituito dall'apparato di misura e dal sistema quantistico possa essere a sua volta considerato come un sistema quantistico più grande e trattato all'interno del formalismo. Infine, il processo fisico dell'interazione tra strumento di misura e sistema quantistico quantistico che

fornisce l'esito della misura non è descritto all'interno del formalismo quantistico standard come un processo dinamico. Vedremo meglio più avanti tale punto. Rimandiamo a [5, 6] per una discussione su tali interessanti e difficili argomenti.

6.2 Stati classici come misure di probabilità sulla σ-algebra delle proposizioni elementari.

6.2.1 Stati e misure di Borel

Consideriamo un sistema fisico classico arbitrario descritto in formulazione hamiltoniana. Questo significa che lo spazio ambiente, lo *spaziotempo delle fasi*, $\mathcal{H}_{n+1}$ sarà dato da una varietà differenziabile fibrata su $\mathbb{R}$, *l'asse del tempo*, le cui fibre $\mathcal{F}_t$ saranno varietà differenziabili (tra di loro diffeomorfe) $\mathcal{F}$ di dimensione $2n$ dotate di una struttura simplettica che permette di formulare le equazioni di Hamilton. Nell'intorno di ogni punto su ciascuna sottovarietà $\mathcal{F}_{t_0}$ avremo delle coordinate naturali $t, q^1, \ldots, q^n, p_1, \ldots, p_n$ rispetto alle quali la forma simplettica è $\omega = \sum_{i=1}^n dq^i \wedge dp_i$ e le equazioni di Hamilton prendono la solita forma per $k = 1, 2, \ldots, n$:

$$\frac{dq^k}{dt} = \frac{\partial H(t, q(t), p(t))}{\partial p_k} , \tag{6.1}$$

$$\frac{dp_k}{dt} = -\frac{\partial H(t, q(t), p(t))}{\partial p_k} . \tag{6.2}$$

essendo H la funzione hamiltoniana delle coordinate locali considerate ed essendo t la coordinata sull'asse del tempo. L'evoluzione del sistema sarà descritta quindi dalle curve integrali delle equazioni scritte sopra. Ad ogni fissato tempo $t \in \mathbb{R}$, ogni evoluzione del sistema determina un punto $(t, p(t))$ con $p(t) \in \mathcal{F}_t$ in cui la curva integrale considerata interseca $\mathcal{F}_t$. Nel caso si stia considerando, non un sistema fisico ma un insieme statistico di sistemi fisici identici, l'evoluzione del sistema è descritta assegnando una funzione densità positiva localmente rappresentabile come $\rho = \rho(t, q, p)$ che risolve l'*equazione di Liouville*:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^n \left(\frac{\partial \rho}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q^i} \frac{\partial \rho}{\partial p_i} \right) = 0 . \tag{6.3}$$

Si richiede ovviamente che, per ogni t

$$\int_{\mathcal{F}_t} \rho \, d\mu = 1 , \tag{6.4}$$

dove $d\mu$, localmente rappresentabile come $d\mu_t = dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n$, è la misura indotta dalla forma simplettica ω su $\mathcal{F}_t$.
È un fatto noto come *teorema di Liouville* la proprietà dell'integrale in (6.4) di essere invariante al variare di t, quando ρ soddisfa (6.3) [12].

$\rho(t,p)$ con $p \in \mathcal{F}_t$ rappresenta la densità probabilità che il sistema fisico si trovi in p al tempo t.
In definitva, sia nel caso di un unico sistema fisico, che nel caso di una trattazione statistica, l'evoluzione del sistema è descritta assegnando una classe di misure di Borel di probabilità[1] $\{\nu_t\}_{t\in\mathbb{R}}$, ciascuna definita sul corrispondente spazio $\mathcal{F}_t$ dove:

(i) se si lavora con un insieme statistico $\nu_t = \rho d\mu_t$;

(ii) se si lavora con un solo sistema si può definire una misura ν_t caso limite di (i) come la misura di Dirac concentrata in $p(t)$.

Ricordiamo che la misura di Dirac δ_x concentrata in un punto x di uno spazio con misura X è definita come $\delta_x(E) = 0$ se $x \notin E$ e $\delta_x(E) = 1$ se $x \in E$, dove E è un qualsiasi elemento della σ-algebra, Σ, dello spazio X.

Lo *stato del sistema al tempo t* (eventualmente pensato come stato *statistico*) sarà quindi individuato dalla misura ν_t.

Il significato operativo della misura ν_t nel caso essa sia una misura di Dirac sarà chiarito tra poco.

6.2.2 Proposizioni e insiemi.

Se ammettiamo che la descrizione hamiltoniana del nostro sistema fisico contenga tutte le proprietà fisiche del sistema, tutte le proposizione sul sistema che al tempo t sono vere o false, o sono vere con una certa probabilità, devono poter essere descritte in $\mathcal{F}_t$ in qualche modo. Inoltre, il loro valore di verità ovvero la probabilità che siano vere deve poter essere ottenuta una volta noto lo stato ν_t del sistema.

Mettiamoci inizialmente nel caso in cui si considera un unico sistema e quindi ν_t è una misura di Dirac. Ogni proposizione P individua un insieme di $\mathcal{F}_t$ che contiene tutti e soli i punti di $\mathcal{F}_t$ che la rendono verificata. Indicheremo tale insieme ancora con P.

Noto lo stato ν_t, la proposizione sarà vera se e solo se il punto $p(t)$ che descrive lo stato al tempo t appartiene all'insieme P.

Se assegnamo convenzionalmente il valore 0 ad una proposizione falsa e 1 ad una vera, in termini di misure ν_t, il valore di verità della proposizione P sullo stato ν_tè dato allora da $\nu_t(P)$ dove ora $P \subset \mathcal{F}_t$ è l'insieme associato alla proposizione P.

Questo fatto chiarice il significato operativo della misura di Dirac ν_t.

La stessa intepretazione può essere usata nel caso lo stato abbia un'interpretazione statistica: *$\nu_t(P)$ rappresente la probabilità che P sia vera al tempo t quando lo stato è ν_t.*

Nota. Anche il primo caso, quello di una miura di Dirac, può interpretarsi nel senso delle misure di probabilità (e la misura di Dirac è una misura di probabilità) se si attribuisce al valore di verità 1 la *probabilità certa* e al valore di verità 0 la *probabilità nulla.* D'ora in poi adotteremo tale interpretazione unificante in entrambi i casi.

È chiaro che tutto quanto detto ha senso se gli insiemi P appartengono alla σ-algebra su cui

[1] Ricordiamo che una misura positiva μ sullo spazio X è detta misura di probabilità se $\mu(X) = 1$.

sono definite le misure ν_t. Tale σ-algebra è quella di Borel ed è pertanto ragionevolmente grande.

Nota. È chiaro che una stessa proposizione può essere formulata in modi differenti ma equivalenti. Nell'identificare proposizioni con insiemi di $\mathcal{F}_t$ noi stiamo esplicitamente assumendo che: *se due proposizioni individuano lo stesso sottoinsieme di $\mathcal{F}_t$ si devono considerare come la stessa proposizione.*

6.2.3 Interpretazione insiemistica dei connettivi logici.

Prese due proposizioni P e Q, possiamo comporle con connettivi logici per costruire altre proposizioni. In particolare possiamo considerare la proposizione costruita con la disgiunzione, $P \vee Q$, e la proposizione costruita con la congiunzione, $P \wedge Q$. Con una sola proposizione possiamo costruirne la negazione $\neg P$.
Possiamo intepretare queste proposizioni in termini di insiemi della σ-algebra di $\mathcal{F}_t$:
(i) $P \vee Q$ corrisponde a $P \cup Q$;
(ii) $P \wedge Q$ corrisponde a $P \cap Q$;
(iii) $\neg P$ corrisponde a $\mathcal{F}_t \setminus P$.

È possibile mettere un ordinamento parziale in ogni classe di sottoinsiemi di $\mathcal{F}_t$ usando l'inclusione insiemistica: $P \leq Q$ se e solo se $P \subset Q$.
L'interpretazione più naturale a livello di proposizioni del fatto che $P \subset Q$ è semplicemente che $P \Rightarrow Q$.

Note.
(1) La probabiltà che siano vere queste proposizioni composte può essere calcolata usando la misura ν_t in quanto le operazioni insiemistiche che corrispondono a $\vee$, $\wedge$, $\neg$ sono operazioni rispetto alle quali ogni σ-algebra è chiusa.
(2) Si verifica facilmente che se ν_t è una misura di Dirac, il valore di probabilità (che ricordiamo essere solo 0 o 1 in questo caso) assegnato a ciascuna delle espressioni composte di sopra, coincide con il valore di verità calcolato tramite le tavole di verità del connettivo logico usato. Per esempio, $P \vee Q$ è vera (cioè $\nu_t(P \cup Q) = 1$) se e solo se almeno una delle due proposizioni costituenti è vera (cioè se e solo se $\nu_t(P) = 1$ oppure, senza esclusione, $\nu_t(Q) = 1$); infatti il punto p su cui è concentrata la misura di Dirac $\delta_x = \nu_t$ cade in $P \cup Q$ se e solo se x cade in almeno uno dei due insiemi P e Q.

6.2.4 Proposizioni "infinite" e grandezze fisiche.

Solitamente nel calcolo proposizionale non si considerano delle proposizioni composte contenenti infinite proposizioni ed infiniti connettivi logici come $P_1 \vee P_2 \vee \ldots$. L'interpretazione delle proposizioni e dei connettivi in termini di elementi e operazioni in una σ-algebra permette però di considerare tali proposizioni.

Per rispondere alle domande dobbiamo precisare la natura delle (o di alcune delle) proposizioni di cui abbiamo discusso sopra. Possiamo legare almeno alcune di tali proposizioni *grandezze fisiche* misurabili sul sistema. In via del tutto generale possiamo considerare le *grandezze fisiche* definite sul nostro sistema hamiltoniano come qualche classe di funzioni regolari in qualche senso $f : \mathcal{H}_{n+1} \to \mathbb{R}$. Una scelta molto generale per il requisito di *regolarità* si ha prendendo tutta la classe delle funzione misurabili limitate. Tali funzioni definiscono *densità* di qualche grandezza che almeno localmente possono essere integrate. Si possono fare delle scelte meno radicali prendendo per esempio solo la classe delle funzioni continuo oppure C^1 oppure C^∞.
Se $f : \mathcal{H}_{n+1} \to \mathbb{R}$ è una grandezza fisica valutabile sul sistema fisico, possiamo costruire con essa delle proposizioni del tipo:

$$P_E^{(f)} = \text{“}\textit{Il valore di } f \textit{ valutato sullo stato del sistema cade nell'insieme di Borel } E \subset \mathbb{R}\text{”} ,$$

Il fatto di considerare insiemi di Borel E e non solo, per esempio, intervalli aperti, consente di trattare sullo stesso piano grandezze che assumono valori continui e grandezze che assumo valori discreti. Infatti nella classe degli insiemi di Borel di $\mathbb{R}$ ci sono anche gli insiemi chiusi, ci sono gli stessi punti di $\mathbb{R}$ ed insiemi contenenti quantità finite e numerabili di punti.
In termini insiemistici:

$$P_E^{(f)} = f^{-1}(A) .$$

Nota. Come al solito, una volta noto lo stato ν_t, la probabilità che $P_E^{(f)}$ sia vera al tempo t per il sistema è $\nu_t(P_E^{(f)})$. Se lo stato non è statistico, il valore suddetto è il valore di verità della stessa proposizione $P_E^{(f)}$.

Consideriamo un 'intervallo $[a, b)$ con $b \leq +\infty$. Possiamo decomporre tale intervallo nell'unione disgiunta dui una infinità di sottointervalli: $[a, b) = \cup_{i=1}^{\infty}[a_i, a_{i+1})$ dove $a_1 := a$ ed inoltre $a_i < a_{i+1}$ ed infine $a_i \to b$ per $i \to \infty$. È chiaro allora che la proposizione

$$P_{[a,b)}^{(f)} = \text{“}\textit{Il valore di } f \textit{ valutato sullo stato del sistema cade nell'insieme di Borel } [a, b)\text{”} ,$$

può esserre decomposta in una *disgiunzione infinita* di proposizioni ciascuna della forma

$$P_{[a_i,a_{i+1})}^{(f)} = \text{“}\textit{Il valore di } f \textit{ valutato sullo stato del sistema cade nell'insieme di Borel } [a_i, a_{i+1})\text{”}.$$

$$P_{[a,b)}^{(f)} = \vee_{i=1}^{+\infty} P_{[a_i,a_{i+1})}^{(f)} ,$$

che corrisponde alla decomposizione dell'insieme $P_{[a,b)}^{(f)}$ nell'unione disgiunta

$$P_{[a,b)}^{(f)} = \cup_{i=1}^{+\infty} P_{[a_i,a_{i+1})}^{(f)} .$$

Quindi vediamo che ha senso fisico assumere l'esistenza di almeno alcune proposizioni sul sistema costruite con infiniti connettivi logici ed infinite sotto proposizioni.

Per dualità rispetto alla negazione, che trasforma la disgiunzione $\vee$ in congiunzione $\wedge$, se assumiamo che l'insieme delle proposizioni ammissibili sia chiuso rispetto alla negazione $\neg$, dobbiamo assumere come fisicamente sensate anche proposizioni infinite costruite con il connettivo $\wedge$.
Il fatto di poter rappresentare le proposizioni come insiemi di una σ-algebra e quindi poterne calcolare la probabilità che siano vere su uno stato del sistema tramite la misura corrispondente, suggerisce di assumere che abbiano senso fisico proposizioni costruite con un'infinità numerabile connettivi logici di tipo $\vee$ oppure un'infinità numerabile di connettivi logici di tipo $\wedge$, perchè gli insiemi corrispondenti sono ancora elementi della σ-algebra che è chiusa rispetto all'unione ed all'intersezione numerabile di insiemi.
Per avere una struttura "isomorfa" a quella di σ-algebra costituita da proposizioni insieme ai connettivi connettivi logici $\vee$, $\wedge$, $\neg$, dobbiamo includere tra le proposizioni fisiche ancora due proposizioni che corrispondono agli insiemi $\emptyset$ e $\mathcal{F}_t$ rispettivamente. Tali proposizioni evidentemente sono quella sempre falsa (ovvero con probabilità 0 di essere vera per qualsivoglia stato) che indicheremo con $\mathbf{0}$ e quella sempre vera (ovvero con probabilità 1 di essere vera per qualsivoglia stato) che indicheremo con $\mathbf{1}$.

Possiamo dare due formulazioni all'assunzione che l'insieme delle proposizioni sul sistema contenga proposizioni "infinite" nel senso precisato sopra, una *più forte* ed una *più debole*. La formulazione più debole si può enunciare come:
W. *Fissata una grandezza $f : \mathcal{F}_t \to \mathbb{R}$ misurabile sul sistema, le proposizioni $P_A^{(f)}$ associate a f, al variare del borelliani A, insieme ai connettivi logici $\vee$, $\wedge$, $\neg$, costituiscano una struttura "isomorfa" ad una σ-algebra.*
La formulazione più forte si può enunciare come:
S. *L'insieme delle proposizioni elementari P relative al sistema fisico considerato, insieme ai connettivi logici $\vee$, $\wedge$, $\neg$, costituiscano una struttura "isomorfa" ad una σ-algebra.*

In entrambi i casi, il valore di verità oppure, a seconda del tipo di stato, la probabilità che una proposizione P sia vera sullo stato del sistema al tempo t si calcola comunque come $\nu_t(P)$ dove ora $P \subset \mathcal{F}_t$ è l'insieme associato alla proposizione. Tutto ciò è indipendente dal fatto che una proposizione sia "normale" o costituita con infiniti connettivi logici.

Nota. Assumendo la versione **S**, ci si può chiedere se la σ-algebra associata all'insieme di tutte le proposizioni sul sistema corrisponda alla σ-algebra di Borel di $\mathcal{F}_t$ o sia più piccola. È chiaro che se si assume che ogni funzione reale misurabile limitata definita su $\mathcal{F}_t$ sia una grandezza fisica allora la risposta è positiva perchè tra queste funzioni ci sono le funzioni caratteristiche di ogni insieme misurabile di $\mathcal{F}_t$. Se f è la funzione caratteristica dell'insieme di Borel E, per esempio, la proposizione $P_{[-1,1]}^{(f)}$ corrisponde all'insieme E.

6.2.5 Il reticolo distributivo, limitato, ortocomplementato e σ-completo delle proposizioni elementari.

In definitiva, sia nella versione debole che in quella forte dell'assunzione sulle proposizioni relative ad un istema fisico ad un istante di tempo, possiamo (anzi dobbiamo per dare seno alle proposizioni "infinite") identificare tali proposizioni con insiemi e dire che le proposizioni *sono insiemi*. Tuttavia, per il passaggio al caso quantistico in cui non esiste lo spazio delle fasi, è importante farsi la domanda che segue. Esistono strutture matematica "isomorfe" ad una σ-algebra che non sono σ-algebre di insiemi? La risposta è positiva e si ricava dalla teoria dei reticoli.

Definizione 6.1. *Un insieme parzialmente ordinato* $(X, \geq)$ *è detto* **reticolo** *quando soddisfa i due seguenti requisiti:*
(a) *per ogni coppia* $x, y \in X$ *esiste* $\sup\{a, b\}$*;*
(b) *per ogni coppia* $x, y \in X$ *esiste* $\inf\{a, b\}$*.*

Note.
(1) Nella teoria dei reticoli $\sup\{a, b\}$ si indica con $a \vee b$ e $\inf\{a, b\}$ si indica con con $a \wedge b$.
(2) Si osservi che per ogni reticolo X, dalla definizione di inf e sup, seguono le *proprietà associative*: per ogni terna $a, b, c \in X$,

$$\begin{aligned} (a \wedge b) \wedge c &= a \wedge (b \wedge c)\,, \\ (a \vee b) \vee c &= a \vee (b \vee c)\,, \end{aligned}$$

per cui, per esempio, si può scrivere $a \vee b \vee c \vee d$ e $a \wedge b \wedge c \wedge d$ senza ambiguità.
(3) Valgono anche le *proprietà di simmetria*:

$$\begin{aligned} a \vee b &= b \vee a\,, \\ a \wedge b &= b \wedge a\,, \end{aligned}$$

per ogni coppia $a, b \in X$. Di conseguenza, per esempio $a \vee b \vee c = c \vee a \vee b \vee a$.

Esistono diversi tipi di reticoli, la seguente definizione ne classifica alcuni.

Definizione 6.2. *Un reticolo* $(X, \geq)$ *si dice:*
(a) reticolo distributivo *se valgono le leggi distributive tra* $\vee$ *e* $\wedge$*: per ogni terna* $a, b, c \in X$

$$\begin{aligned} a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee b)\,, \\ a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge b)\,; \end{aligned}$$

(b) reticolo limitato *se* X *ammette minimo, indicato con* $\mathbf{0}$ *e massimo, indicato con* $\mathbf{1}$*;*
(c) reticolo ortocomplementato *se è dotato di un'applicazione* $X \ni a \mapsto \neg a$*, dove* $\neg a$ *è detto* **ortocomplemento** *di* a*, che soddisfa:*

(i) $a \vee \neg a = \mathbf{1}$ *per ogni* $a \in X$,
(ii) $a \wedge \neg a = \mathbf{0}$ *per ogni* $a \in X$,
(iii) $\neg(\neg a) = a$ *per ogni* $a \in X$,
(iv) se $a \geq b$ *allora* $\neg b \geq \neg a$ *per ogni coppia* $a, b \in X$.
(d) reticolo σ-completo, *se ogni insieme numerabile* $\{a_n\}_{n\in\mathbb{N}} \subset X$ *ammette estremo superiore che viene denotato con* $\vee_{n\in\mathbb{N}} a_n$.
Un reticolo che soddisfi (a), (b) e (c) è detto **algebra di Boole**.

Note.
(1) Si verifica facilmente che per ogni algebra di Boole X valgono le *leggi di de Morgan*:

$$\begin{aligned} \neg(a \vee b) &= \neg a \wedge \neg b\,, \\ \neg(a \wedge b) &= \neg a \vee \neg b\,, \end{aligned}$$

per ogni $a, b \in X$.
(2) Dalla stessa definizione di inf e sup è facile verificare che un'algebra di Boole è σ-completa se e solo se ogni insieme numerabile $\{a_n\}_{n\in\mathbb{N}} \subset X$ ammette estremo inferiore. Tale estremo inferiore viene denotato con $\wedge_{n\in\mathbb{N}} a_n$.

In base alle definizioni date risulta ovvio che vale la seguente proposizione.

Proposizione 6.1. *Ogni σ-algebra su un insieme X è un'algebra di Boole (ossia un reticolo distributivo, limitato, ortocomplementato), σ-completa in cui:*
(i) la relazione d'ordine è l'inclusione insiemistica (e quindi $\vee$ corrisponde a $\cup$ e $\wedge$ corrisponde a $\cap$),
(ii) il massimo ed il minimo dell'algebra di Boole sono X e $\emptyset$,
(iii) l'operazione di ortocomplemento corrisponde al complemento insiemistico rispetto allo spazio X.

Dobbiamo concludere la stessa cosa per l'insieme delle proposizioni relative ad un sistema fisico (sia nella versione forte **S** che nella versione debole **W**):
le proposizioni considerate in **S** *e in* **W** *costituiscono un reticolo distributivo, limitato, ortocomplementato e σ-completo in cui:*
(i) la relazione d'ordine è l'implicazione logica,
(ii) il massimo ed il minimo sono la proposizione sempre vera **1** *e quella sempre falsa* **0**,
(iii) l'operazione di ortocomplemento corrisponde alla negazione.

6.3 Le proposizioni relative a sistemi quantistici come insiemi di proiettori ortogonali.

Passiamo ora a considerare sistemi quantistici. Cercando di seguire un approccio per quanto possibile vicino al caso classico, vogliamo prima di tutto cercare di dare un modello matematico per la classe delle proposizioni relative allo stato quantistico ad un tempo t fissato, di un sistema quantistico. Non sappiamo ancora come descrivere lo stato quantistico, ma sappiamo qualcosa, **MQ1** e **MQ2**, riguardo alle grandezze quantistiche misurabili sul sistema quando ne è assegnato lo stato.

Nota. D'ora in poi, seguendo la tradizione, diremo **osservabili** le grandezze fisiche relative a sistemi quantistici e quindi soddifacenti **MQ1** e **MQ2**.

Per il momento ci concentreremo su **MQ2**. Sappiamo che esistono osservabili *incompatibili*. È immediato concludere che esistono *proposizioni incompatibili*: se A e B sono osservabili incompatibili

$$P_J^{(A)} = \text{“}\textit{Il valore di } A \textit{ valutato sullo stato del sistema cade nell'insieme di Borel } J \subset \mathbb{R}\text{”} ,$$

e

$$P_K^{(B)} = \text{“}\textit{Il valore di } B \textit{ valutato sullo stato del sistema cade nell'insieme di Borel } K \subset \mathbb{R}\text{”} ,$$

sono in generale incompatibili: i valori di verità di tali proposizioni si disturbano a vicenda esegundo la verifica di esse ripetutamente in successione con intervalli di tempo piccoli a piacere, in modo da non imputare il fenomeno all'evoluzione dello stato del sistema. Non ha molto senso fisico, in questa situazione, affermare che le due proposizioni considerate possano assumere sul sistema un valore di verità *contemporaneamente*. In altre parole non ha senso dare significato fisico a proposizioni del tipo $P_J^{(A)} \vee P_K^{(B)}$ oppure $P_J^{(A)} \wedge P_K^{(B)}$. Non possiamo allora assumere il modello dell'insieme delle proposizioni dato da qualche reticolo (o σ-algebra) di insiemi in cui la disgiunzione e la congiunzione sono sempre possibili. Dovremmo imporre degli ulteriori vincoli su tale modello, vietando certi sottoinsiemi. Una procedura alternativa, che si è rivelata molto più interessante è quella di modellizzare le proposizioni elementari tramite i proiettori ortogonali di uno spazio di Hilbert.

6.3.1 Reticoli di proiettori ortogonali su spazi di Hilbert.

L'insieme dei proiettori ortogonali su uno spazio di Hilbert godono di alcune interessantissime proprietà molto vicine a quelle dei reticoli booleani [5]. Non esamineremo la questione a fondo ma discuteremo solo alcune proprietà importanti.
Notiamo che e P e Q sono proiettori ortogonali, allora in generale PQ, $P + Q - PQ$ e $P + Q$ *non* sono più proiettori ortogonali. Tuttavia ci sono delle eccezioni notevoli.

Proposizione 6.2. *Sia $(H, (\ |\ ))$ uno spazio di Hilbert e $\mathfrak{P}(H)$ denoti la classe dei proiettori ortogonali di H. Valgono le proprietà che seguono per $P, Q \in \mathfrak{P}$.*
(a) *I seguenti fatti sono equivalenti:*
(i) $P \leq Q$;

(ii) $P(H)$ è sottospazio di $Q(H)$;
(iii) $PQ = QP = P$.
(b) *I seguenti fatti sono equivalenti:*
(i) $PQ = 0$;
(ii) $QP = 0$;
(iii) $P(H)$ e $Q(H)$ sono ortogonali.
Se vale (i) oppure (ii) oppure (iii), $P + Q$ è proiettore ortogonale e proietta su $P(H) \oplus Q(H)$.
(c) *Se $PQ = QP$ allora PQ è proiettore ortogonale e proietta su $P(H) \cap Q(H)$.*
(d) *Se $PQ = QP$ allora $P + Q - PQ$ è proiettore ortogonale e proietta sul sottospazio chiuso $\overline{< P(H), Q(H) >}$.*

Prova. (a) Prima di tutto notiamo che se P è un proiettore che proietta su M, allora $Pu = 0$ se e solo e $u \in M^{\perp}$ a causa della decomposizione diretta ortogonale $H = M \oplus M^{\perp}$ ((d) in teorema 3.1) e tenendo conto che il componente di u in M è proprio Pu.
(i) $\Rightarrow$ (ii). Se $P \leq Q$ allora $(u|Qu) \geq (u|Pu)$ ovvero essendo i proiettori idempotenti ed autoaggiunti: $(Qu|Qu) \geq (Pu|Pu)$ che si può anche scrivere come $||Qu|| \geq ||Pu||$. Quindi in particolare $Qu = 0$ implica $Pu = 0$ ovvero $Q(H)^{\perp} \subset P(H)^{\perp}$. Usando (e) di teorema 3.1 e notando che $Q(H)$ e $P(H)$ sono chiusi, troviamo infine che $P(H) \subset Q(H)$.
(ii) $\Rightarrow$ (iii). Si prova immediatamente scegliendo una base hilbertiana N di H costruita come l'unione di tre basi hilbertiane rispettivamente nei tre spazi adu a due ortogonali $S_1 = P(H)$, $S_2 = P(H)^{\perp} \cap Q(H)$, $S_3 = Q(H)^{\perp}$ e notando che il proiettore P si scrive $P = \sum_{u \in S_1} (u|\ )u$ mentre $Q = \sum_{u \in S_1} (u|\ )u + \sum_{u \in S_2} (u|\ )u$ ((d) di proposizione 3.9)
(iii) $\Rightarrow$ (i). Si prova immediatamentte dalle espressioni scritte sopra per P e Q.
(b) Supponiamo $PQ = 0$, prendendo l'aggiunto si ha che $QP = 0$, quindi si verifica immediatamente che $P(H)$ e $Q(H)$ sono ortogonali essendo $PQ = QP = 0$. Se $P(H)$ e $Q(H)$ sono ortogonali, prendiamo in ciascuno di tali spazi una base hilbertiana N e N' ripettivamente, scrivendo i proiettori P e Q come indicato in (d) di proposizione 3.9: $P = \sum_{u \in N} (u|\ )u$, $Q = \sum_{u \in N'} (u|\ )u$, si ricava immediatamente che $PQ = QP = 0$. Usando le espressioni di P e Q scritte sopra, tenendo conto che $N \cup N'$ è una bae hilbertiana per $P(H) \oplus Q(H)$ ed usando ancora (d) di proposizione 3.9 si ricava che $P + Q$ è il proiettore ortogonale che proietta su $P(H) \oplus Q(H)$.
(c) Il fatto che PQ sia proiettore ortogonale (cioè autoaggiunto ed idempotente) se $PQ = QP$ dove P e Q sono proiettori ortogonali è di immediata verifica. Se $u \in H$, $PQu \in P(H)$, ma anche $PQu = QPu \in Q(H)$ e quindi $PQu \in P(H) \cap Q(H)$. Abbiamo provato che $PQ(H) \subset P(H) \cap Q(H)$, per concludere basta provare che $P(H) \cap Q(H) \subset PQ(H)$. Se $u \in P(H) \cap Q(H)$ allora $Pu = u$ e $Qu = u$ e quindi vale anche $Pu = PQu = u$ ossia $u \in PQ(H)$. Questo significa che se $P(H) \cap Q(H) \subset PQ(H)$.
(d) Il fatto che $R := P + Q - PQ$ sia proiettore ortogonale è di immediata verifica. Consideriamo

lo spazio $\overline{< P(H), Q(H) >}$. Possiamo costruire una base hilbertiana per tale spazio come segue. Per prima cosa costruiamo una base hilbertiana N per il sottspazio chiuso $P(H) \cap Q(H)$. Quindi aggiungiamo a questa una seconda base hilbertiana per lo spazio che "rimane in $P(H)$ una volta tolto $P(H) \cap Q(H)$", cioè costruiamo una base hilbertiana N' per lo spazio chiuso ortogonale in $P(H)$ a $P(H) \cap Q(H)$: $P(H) \cap (P(H) \cap Q(H))^{\perp}$. Infine costruiamo, con lo stesso criterio, una terza base hilbertiana N'' per $Q(H) \cap (P(H) \cap Q(H))^{\perp}$. È facile convincersi che le tre basi sono a due a due ortogonali e insieme costituiscono una base per $\overline{< P(H), Q(H) >}$. Tutto ciò mostra che vale la decomposizione diretta ed ortogonale

$$\overline{< P(H), Q(H) >} = (P(H) \cap Q(H)) \oplus (P(H) \cap (P(H) \cap Q(H))^{\perp}) \oplus (Q(H) \cap (P(H) \cap Q(H))^{\perp}) \, .$$

Nelle nostre ipotesi, il proiettore sul primo pazio della decomposizione è PQ per (c). Il proiettore su $(P(H) \cap Q(H))^{\perp})$ è quindi $I - PQ$. Quindi, applicando nuovamente (c), il proiettore ortogonale sul secondo addendo della decomposizione è $P(I - PQ) = P - PQ$. Similmente, il proiettore sul terzo addendo della decompoizione è $Q(I - PQ) = Q - PQ$. Applicando (b), il proiettore sulla somma diretta ortogonale dei tre spazi detti, cioè su $\overline{< P(H), Q(H) >}$ sarà:

$$PQ + (P - PQ) + (Q - PQ) = P + Q - PQ \, .$$

Questo conclude la dimostrazione. □

In base a quanto dimostrato, se $P, Q \in \mathfrak{P}(H)$ commutano, allora definiti:
(i) $P \wedge Q := PQ$,
(ii) $P \vee Q := P + Q - PQ$,
(iii) $\neg P : I - P$,
i secondi membri delle definizioni date sopra sono ancora proiettori ortogonali. Inoltre questi proiettori soddisfano proprietà analoghe a quelle delle proposizioni costruite con connettivi logici. Per esempio $\neg(P \vee Q) = \neg P \wedge \neg Q$ e via di seguito.
Se però P e Q non commutano, PQ e $P + Q - PQ$ non sono nemmeno proiettori in generale. Tutto ciò sembra utile nel cercare un modello per le proposizioni dei sistemi quantistici tenendo conto di **MQ2**. L'idea è che:
le proposizioni dei sistemi quantistici sono in corrispondenza biunivoca con proiettori ortogonali in uno spazio di Hilbert in modo tale che:
(i) l'implicazione logica tra proposizioni $P \Rightarrow Q$ corrisponda alla relazione $P \leq Q$ tra i corrispondenti proiettori;
(ii) due proposizioni sono compatibili se e solo se i rispettivi proiettori commutano.

Se l'idea è sensata, le similitudini strutturali tra i costrutti con i proiettori ortogonali e quelli dell'algebra di Boole σ-finita delle proposizioni deve valere anche quando i considerino più di due proposizioni. La proposizione seguente mostra che è davvero così.

Proposizione 6.3. *Sia H uno spazio di Hilbert e sia $\mathfrak{P}_0(H) \subset \mathfrak{P}(H)$ un insieme di proiettori ortogonali commutanti che sia massimale rispetto alla condizione di commutatività (cioè non*

esiste in $\mathfrak{P}(H)$ un altro proiettore ortogonale che commuti con gli elementi di $\mathfrak{P}_0(H)$).
(a) *$\mathfrak{P}_0$ è un'algebra di Boole (ossia un reticolo distributivo, limitato, ortocomplementato), σ-completa in cui:*

(i) la relazione d'ordine è la relazione d'ordine $\geq$ tra proiettori,

(ii) il massimo ed il minimo dell'algebra di Boole sono I (operatore identità) e 0 (operatore nullo) rispettivamente,

(iii) l'operazione di ortocomplemento del proiettore ortogonale P corrisponde a

$$\neg P = I - P \,,$$

(proiettore che proietta su $P(H)^{\perp}$).
(b) *Risulta che, se $P, Q \in \mathfrak{P}_0(H)$:*

$$\begin{aligned} P \wedge Q &= PQ, \\ P \vee Q &= P + Q - PQ \,. \end{aligned}$$

Prova. Per prima cosa mostriamo che $\mathfrak{P}_0(H)$ è un'algebra di Boole. Notiamo che $\geq$ è una relazione d'ordine parziale in $\mathfrak{P}_0(H)$ per (f) di proposizione 3.8 ricordando che i proiettori ortogonali sono operatori positivi. È anche chiaro che per ogni proiettore ortogonale $P \geq 0$ e $I \geq P$ (perchè $I - P$ è proiettore ortogonale). Dato che 0 e I commutano con tutti gli elementi di $\mathfrak{P}_0(H)$ appartemgono a $\mathfrak{P}_0(H)$. Bisogna infine solo provare che per ogni coppia di proiettori P, Q nell'insieme detto esistono l'estremo superiore e quello inferiore per l'insieme $\{P, Q\}$ nello spazio $\mathfrak{P}(H)$ e si calcolano come detto in (b). È chiaro allora che $P \vee Q$ e $P \wedge Q$ appartengono a $\mathfrak{P}_0(H)$ perchè per costruzione commutano con tutti gli elementi di $\mathfrak{P}_0(H)$. Inoltre le proprietà di distributività e quelle dell'ortocomplementazione richieste nella definizione 6.2 seguono immediatamente da (b) e dalle proprietà di commutatività dei proiettori e dal fatto che soddifano la proprietà di idempotenza $PP = P$.

Capitolo 7

Appendice A: Relazioni d'ordine e Insiemi Parzialmente Ordinati.

Se X è un insieme arbitrario, una relazione $\geq$ su X viene detta **relazione d'ordine parziale** quando è *transitiva* (per $x, y, z \in X$, se $x \geq y$ e $y \geq z$ allora $x \geq z$) e *antisimmetrica* (per $x, y \in X$, se $x \geq y$ e $y \geq x$ allora $x = y$). In tal caso $(X, \geq)$ si dice **insieme parzialmente ordinato**.
Al posto di $a \geq b$ si può scrivere equivalentemente $b \leq a$.
La relazione d'ordine parziale $\geq$ si dice **relazione ordine** se vale ulteriormente la proprietà: per ogni coppia $x, y \in X$, $x \geq y$ oppure $y \geq x$.
Se $(X, \geq)$ è un insieme parzialmente ordinato:

(i) $Y \subset X$ si dice **limitato superiormente** (**inferiormente**) se ammette un **maggiorante** (**minorante**), cioè un elemento $x \in X$ tale che $x \geq y$ ($y \geq x$) per ogni $y \in Y$;

(ii) elemento $x_0 \in X$ tale che *non* esiste in X alcun $x \neq x_0$ con $x \geq x_0$, si dice **elemento massimale** di X. (Si noti che se x_0 è massimale è in generale falso che ogni $x \in X$ soddisfi $x_0 \geq x$ cioè un elemento massimale di X non è in generale un maggiorante di X).

Se $(X, \geq)$ è un insieme parzialmente ordinato, $Y \subset X$ si dice **ordinato** se $\geq$ ristretta a $Y \times Y$ è una relazione d'ordine.
Ricordiamo il lemma di Zorn equivalente all'Assioma della Scelta (o di Zermelo):

Lemma di Zorn. *Se in un insieme parzialmente ordinato $(X, \geq)$ ogni sottoinsieme ordinato è limitato superiormente, allora X ammette almeno un elemento massimale.*

Tra le nozioni utili nella teoria degli insiemi parzialmente ordinati ci sono anche quelle di estremo superiore ed estremo inferiore. Se $(X, \geq)$ è un insieme parzialmente ordinato:

(i) a è detto **estremo superiore** dell'insieme $A \subset X$ e si scrive $a = \sup A$, se a è un maggiorante di A e se ogni altro maggiorante di A, a', soddisfa $a \leq a'$;

(ii) a è detto **estremo inferiore** dell'insieme $A \subset X$ e si scrive $a = \inf A$, se a è un mino-

rante di A e se ogni altro minorante di A, a', soddisfa $a' \leq a$;

È immediato verificare che ogni sottoinsieme $A \subset X$ ammette al più un unico estremo superiore ed un unico estremo inferiore.

Capitolo 8

Appendice B: Dimostrazione di alcuni teoremi.

Capitolo 9

Appendice C: Integrale di funzioni a valori spazi di Banach.

Assumeremo nota la teoria della misura (positiva) per funzioni a valori complessi sviluppata come in [1, 13]. Vogliamo qui mostrare come si possa estendere [18] la teoria dell'integrazione al caso di funzioni a valori in spazi di Banach, $f : X \to B$, dove (X, Σ, μ) è uno spazio con misura positiva σ-additiva arbitrario. È chiaro che il problema maggiore risiede nel fatto che, in generale, in uno spazio di Banach B non esiste un ordinamento come in $\mathbb{R}$, per cui non possiamo definire l'integrale di una funzione $f : X \to B$ prendendo l'estremo superiore degli integrali delle funzioni semplici *maggiorate* da f. Bisogna elaborare un'altra procedura che però si dimostra essere equivalente alla procedura standard quando $dimB < \infty$ (in particolare $B = C$) e la misura è completa (cioè ogni sottoinsieme di un insieme di misura nulla appartiene alla σ-algebra Σ) e σ-finita (cioè lo spazio X è unione numerabile di elementi della σ-algebra Σ ciascuno con misura finita).

Nel seguito di questa appendice $(B, |\ |)$ è sempre uno spazio di Banach sul campo $\mathbb{C}$ (anche se quanto diremo funzionerà anche se il campo è $\mathbb{R}$) e (X, Σ, μ) è uno spazio con misura positiva.

Diremo che $s : X \to B$ è una **funzione a gradino** se e solo se è una funzione *semplice* (ossia l'immagine è costituita da un numero finito di valori le cui controimmagini sono insiemi misurabili) e le controimmagini dei valori assunti da s sono insiemi di misura *finita*. Indichiamo con $St(X, \mu)_B$ lo spazio vettoriale sul campo $\mathbb{K}$ delle funzioni a gradino.

Diremo che $f : X \to B$ è μ-misurabile se solo se esiste una successione $\{s_n\}_{n\in\mathbb{N}} \subset St(\mu, X)_B$ tale che $s_n \to f$, quasi ovunque rispetto a μ, per $n \to +\infty$. Indicheremo con $M(\mu, X)_B$ lo spazio vettoriale su $\mathbb{K}$ delle funzioni su X a valori in B μ-misurabili.

Nota. Si può provare che se $dimB < \infty$ (in particolare $B = \mathbb{C}$) e se μ è completa e σ-finita, allora $f : X \to B$ è **μ-misurabile** se e solo se è misurabile (e quindi la nozione non dipende

dalla scelta di μ).

Definiamo l'**integrale su** $Y \subset \Sigma$ della della funzione a gradino $s = \sum_{i=1}^{n} a_i \chi_{A_i}$ (dove χ_E è la funzione caratteristica dell'insieme E e $A_i := s^{-1}(a_i)$) come

$$\int_Y s \, d\mu := \sum_{i=1}^{n} a_i \mu(A_i \cap Y) \, .$$

Ovviamente risulta

$$\int_Y s \, d\mu = \int_X \chi_Y \cdot s \, d\mu \, .$$

Indicheremo con $\mathcal{L}^1(X, \mu)_B$ il sottospazio vettoriale di $M(X, \mu)$ costituito dalle $f : X \to B$ per le quali esiste una successone $\{s_n\} \subset St(X, \mu)_B$ che tende a f quasi ovunque rispetto a μ e che soddisfi:

$$\lim_{m \to +\infty} \int_X |f - s_m| \, d\mu = 0 \, , \tag{9.1}$$

Le funzioni di $\mathcal{L}^1(X, \mu)_B$ si dicono **funzioni integrabili secondo Bochner**.
Si prova che in tal caso esiste il limite

$$\lim_{n \to +\infty} \int_X s_n \, d\mu =: \int_X f \, d\mu \, ,$$

che viene detto essere l'**integrale di** f **rispetto a** μ. La definizione è ben posta in quanto, se $\{s_n\}$ e $\{s'_n\}$ sono successioni in $St(X, \mu)_B$ che tendono alla stessa $f \in \mathcal{L}^1(X, \mu)_B$, allora:

$$\lim_{n \to +\infty} \int_X s_n \, d\mu = \lim_{n \to +\infty} \int_X s'_n \, d\mu \, .$$

L'integrale di f su $Y \in \Sigma$ viene definito al solito come:

$$\int_Y f d\mu := \int_X \chi_Y \cdot f \, d\mu \, .$$

Si può definire una seminorma $|| \ ||_1$ su $\mathcal{L}^1(X, \mu)_B$ ponendo:

$$||f||_1 := \int_X |f| \, d\mu \, .$$

Questa seminorma ha la proprietà che se $f \in \mathcal{L}^1(X, \mu)_B$ allora:

$$\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu \, .$$

Si prova che la seminorma $|| \ ||_1$ è tale che $||f||_1 = 0$ se e solo se $f = 0$ quasi ovunque rispetto a μ. Definendo quindi $L^1(X, \mu)_B$ esattamente come nel caso in cui $B = \mathbb{C}$ si può provare che tale

spazio risulta essere spazio di Banach.

La teoria definita consente di generalizzare il teorema della convergenza dominata di Lebesgue come segue.

Teorema della convergenza dominata generalizzato. *Sia (X, Σ, μ) spazio con misura positiva e B spazio di Banach su $\mathbb{C}$.*
Se $\{f_n\}_{n\in\mathbb{N}}$ è una successione di funzioni in $\mathcal{L}^1(X,\mu)_B$ tali che esiste $g \in \mathcal{L}^1(X,\mu)_{\mathbb{C}}$ non negativa con

$$|f_n(x)| \leq g(x) \quad \textit{quasi ovunque rispetto a } \mu \textit{ su } X$$

e tale che

$$f_n(x) \to f(x) \quad \textit{quasi ovunque rispetto a } \mu \textit{ su } X \textit{ se } n \to +\infty\,,$$

allora:
(a) $f \in \mathcal{L}^1(X,\mu)_B$*;*
(b) $||f_n - f||_1 \to 0$ *per* $n \to +\infty$*;*
(c) *vale che:*

$$\int_X f\, d\mu = \lim_{n\to+\infty} \int_X f_n\, d\mu\,.$$

Questo teorema ha una serie notevole di conseguenze:
(1) se $g \in \mathcal{L}^1(X,\mu)_{\mathbb{C}}$, $f \in \mathcal{L}^1(X,\mu)_B$ e $|f(x)| \leq g(x)$ quasi ovunque rispetto a μ su X allora $f \in \mathcal{L}^1(X,\mu)_B$;
(2) $f \in \mathcal{L}^1(X,\mu)_B$ se e solo se $|f| \in \mathcal{L}^1(X,\mu)_{\mathbb{C}}$;
(3) se $\{f_n\}_{n\in\mathbb{N}}$ è una successione di funzioni in $\mathcal{L}^1(X,\mu)_B$ tali che $f_n(x) \to f(x)$ quasi ovunque rispetto a μ su X se $n \to +\infty$ ed esite $c < \infty$ con $||f||_1 \leq c$ per ogni $n \in \mathbb{N}$, allora $f \in \mathcal{L}^1(X,\mu)_B$ e $||f||_1 \leq c$;
(4) se B_1 è un secondo spazio di Banach e $T \in \mathfrak{B}(B_1, B)$, allora se $f \in \mathcal{L}^1(X,\mu)_{B_1}$ vale $T \circ f \in \mathcal{L}^1(X,\mu)_B$. Ulteriormente, l'applicazione

$$\mathcal{T} : \mathcal{L}^1(X,\mu)_{B_1} \ni f \mapsto T \circ f \in \mathcal{L}^1(X,\mu)_B$$

è lineare e continua e soddisfa.
(i) $||\mathcal{T}|| \leq ||T||$,
(ii) $T(\int_X f\, d\mu) = \int_X T \circ f\, d\mu\,.$

I teoremi di passaggio del simbolo di limite e di derivata sotto il segno di integrale ed il teorema di Fubini-Tonelli non cambiano forma rispetto alla formulazione standard della teoria della misura per funzioni a valori in $\mathbb{C}$.

Nota. Si può verificare che la definizione data di integrale coincide con quella ordinaria quando $B = \mathbb{C}$. Più precisamente $f \in \mathcal{L}^1(X,\mu)_{\mathbb{C}}$ se e solo se f è integrabile nel senso della teoria della misura ordinaria. In tal caso i due tipi di integrale forniscono lo stesso risultato e lo spazio

$\mathcal{L}^1(X, \mu)$ definito secondo la teoria ordinaria coincide con $\mathcal{L}^1(X, \mu)_{\mathbb{C}}$ definito sopra.

Bibliografia

[1] W. Rudin, *Analisi reale e complessa*, Boringhieri, Torino 1982

[2] W. Rudin, *Fuctional Analysis*, Second Edition, McGraw-Hill, Boston, 1991.

[3] R.F. Streater, A.S. Wightman, *PCT,Spin and Statistics, and All That*, Revised edition, Princeton University Press, Princeton, 2000.

[4] S.Weinberg, *La Teoria Quantistica dei Campi* Vol 1,2,3, Zanichelli, Bologna, 1999.

[5] G. Boniolo (a cura di), *Filosofia della fisica*, Bruno Mondadori, Milano, 1997.

[6] B. D'Espagnat, *I Fondamenti Concettuali della Meccanica Quantistica*, Bibliopolis, Napoli, 1980.

[7] O.Bratteli, D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol 1 and 2, Second Edition Springer-Verlag, Berlin, 2002.

[8] V.D. Braginsky and F.YA. Khalili, *Quantum measurements*, (first paperback edition) Cambridge University Press, Cambridge, 1995.

[9] R. Haag, *Local Quantum Physics*, Second Revised and Enlarged Edition, Springer-Verlag, Berlin, 1996.

[10] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin, 1932 (ristampe 1981,1996).
Ristampa in traduzione italiana:
J. von Neumann, *I fondamenti matematici della meccanica quantistica*, a cura di Giovanni Boniolo, Il Poligrafo, Padova, 1998.

[11] P.A.M. Dirac, *The principles of Quantum Mechanics* Oxford University Press, Oxford, 1930.
Ristampa in traduzione italiana:
P.A.M. Dirac, *I Principi della Meccanica Quantistica* Boringhieri, Torino, Oxford, 1976.

[12] P. Caldirola, R. Cirelli, G.M. Prosperi, *Introduzione alla Fisica Teorica*, UTET, Torino, 1982.

[13] P.R. Halmos, *Measure Theory*, Van Nostrand R. ,1969.

[14] G. Martucci, *Spazi di Hilbert con elementi di meccanica quantistica*. Pitagora Editrice Bologna, Bologna, 1982.

[15] L.V. Kantorovic, G.P. Akilov, *Analisi Funzionale*, Editori Riuniti, Roma, 1980.

[16] E. Prugovecki, *Quantum Mechanics in Hilbert Space*, Second Edition, Academic Press, New York, 1981.

[17] A.A. Kirillov, A.D. Gvisiani, *Teoremi e problemi dell'analisi funzionale*, MIR, Mosca, 1983.

[18] J.Diestel, J.J. Uhl, Jr., *Vector Measures*, American Mathematical Society, Providence, USA, 1977.

Notes on Multi-Linear Algebra and Tensor Calculus

(For the course Geometrical Methods in Mathematical Physics)

by **Valter Moretti**
www.science.unitn.it/∼moretti/homeE.html

Department of Mathematics,
Faculty of Science,
University of Trento
Italy

Academic year 2005-2006

Contents

1 Multi-linear Mappings and Tensors. 3
1.1 Dual space and conjugate space. 3
1.2 Multi linearity: tensor product, tensors, universality theorem. 7

2 Tensor algebra. Abstract Index Notation. 18
2.1 Tensor algebra generated by a vector space. 18
2.2 The abstract index notation and rules to handle tensors. 19
2.3 Physical invariance of the form of laws and tensors. 24

3 Some Applications. 26
3.1 Tensor products of group representations. 26
3.2 A quantum physical example. 29
3.3 Permutation group and symmetry of tensors. 30

4 Scalar Products and Metric Tools. 35
4.1 Scalar products. 35
4.2 Natural isomorphism between V and V^* and metric tensor. 38
4.3 Raising and lowering of indices of tensors. 43

5 Pseudo tensors, Ricci's pseudotensor and tensor densities. 46
5.1 Orientation and pseudo tensors. 46
5.2 Ricci's pseudo tensor. 47
5.3 Tensor densities. 52

6 Appendix: Square Roots of operators and Polar Decomposition Theorem. 55

Acknowledgments.

The author is grateful to R. Aramini who read these notes carefully correcting several misprints and to D. Ansari for helpful suggestions to improve the text.

1 Multi-linear Mappings and Tensors.

Within this section we introduce basic concepts concerning multi-linear algebra and tensors. The theory of vector spaces and linear mappings is assumed to be well known.

1.1 Dual space and conjugate space.

As a first step we introduce the dual space and the conjugate space of a given vector space.

Def.1.1. (Dual Space, Conjugate Dual Space and Conjugate space.) *Let V be a vector space on the field either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *The* **dual space** *of V, V^*, is the vector space of linear functionals on V, i.e., the linear mappings $f : V \to \mathbb{K}$.*
(b) *If $\mathbb{K} = \mathbb{C}$, the* **conjugate dual space** *of V, $\overline{V^*}$, is the vector space of anti-linear functionals on V, i.e., the antilinear mappings $g : V \to \mathbb{C}$. Finally the* **conjugate space** *of V, $\overline{V}$ is the space $(\overline{V^*})^*$*

Comments.
(1) If V and V' are vector spaces on $\mathbb{C}$, a mapping $f : V \to V'$ is called *anti linear* or *conjugate linear* if it satisfies

$$f(\alpha u + \beta v) = \overline{\alpha} f(u) + \overline{\beta} f(v)$$

for all $u, v \in V$ and $\alpha, \beta \in \mathbb{C}$, $\overline{\lambda}$ denoting the complex conjugate of $\lambda \in \mathbb{C}$. If $V' = \mathbb{C}$ the given definition reduces to the definition of *anti-linear functional.*
(2) V^*, $\overline{V^*}$ and $\overline{V}$ turn out to be vector spaces on the field $\mathbb{K}$ when the composition rule of vectors and the product with elements of the field are defined in the usual way. For instace, if $f, g \in V^*$ or $\overline{V^*}$, and $\alpha \in \mathbb{K}$ then $f + g$ and αf are functions such that:

$$(f + g)(u) := f(u) + g(u)$$

and

$$(\alpha f)(u) := \alpha f(u)$$

for all of $u \in V$.

Def.1.2. (Dual Basis, Conjugate Dual Basis and Conjugate Basis.) *Let V be a finite-dimensional vector space on either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $\{e_i\}_{i \in I}$ be a vector basis of V. The set $\{e^{*j}\}_{j \in I} \subset V^*$ whose elements are defined by*

$$e^{*j}(e_i) := \delta_i^j$$

for all $i, j \in I$ is called the **dual basis** *of $\{e_i\}_{i \in I}$.*
Similarly, if $\mathbb{K} = \mathbb{C}$, the set of elements $\{\overline{e^{*j}}\}_{j \in I} \subset \overline{V^*}$ defined by:

$$\overline{e^{*j}}(e_i) := \delta_i^j$$

for all $i, j \in I$ is called the **conjugate dual basis** *of $\{e_i\}_{i\in I}$.*
Finally the set of elements $\{\overline{e}_j\}_{j\in I} \subset \overline{V}$ defined by:

$$\overline{e}_p(\overline{e^{*q}}) := \delta^q_q$$

for all $p, q \in I$ is called the **conjugate basis** *of $\{e_i\}_{i\in I}$.*

Comments.
(1) We have explicitly assumed that V is finite dimensional. Anyway, each vector space admits a vector basis (i.e. a possibly infinite set of vectors such that each vector of V can be obtained as a finite linear combination of those vectors), therefore one could discuss on the validity of the given definitions in the general case of a non-finite dimensional vector space. We shall not follow that way because we are interested on algebraic features only and the infinite-dimensional case should be approached by convenient topological tools which are quite far form the goals of these introductory notes.
(2) If the vector space V with field $\mathbb{K}$ admits the vector basis $\{e_i\}_{i\in I}$, each linear or anti-linear mapping $f : V \to \mathbb{K}$ is completely defined by giving the values $f(e_i)$ for all of $i \in I$. This is because, if $v \in V$ then $v = \sum_{i\in I_v} c^i e_i$ for some numbers $c^i \in \mathbb{K}$, $I_v \subset I$ being *finite*. Then the linearity of f yelds $f(v) = \sum_{i\in I_v} c^i f(e_i)$. (This fact holds true no matter if $dimV < +\infty$). Using the result above, one realizes that the given definitions of elements e^{*j} and $\overline{e^*}^i$ completely determine these functionals and thus the given definitions are well-posed.
(3) One may wonder whether or not the dual basis and the conjugate basis are *proper* vector bases of respective vector spaces, the following theorem gives a positive answer.
(4) In *spinor* theory, if V is a complex two-dimensional space: The "upper-index spinor space", $\overline{V}$ is the space of "upper-index pointed spinors" or "upper-index primed spinors". $\overline{V^*}$ is the space of "lower-index pointed spinors" or "lower-index primed spinors". Finally V^* is the space of "upper-index spinors".

Theorem 1.1. *Concerning* **Def.1.2** *the dual basis, the conjugate dual basis and the conjugate basis of a base $\{e_i\}_{i\in I} \subset V$, with $dimV < \infty$, are vector bases for V^*, $\overline{V^*}$ and $\overline{V}$ respectively. As a consequence $dimV = dimV^* = dim\overline{V^*} = dim\overline{V}$.*

Proof. Consider the dual basis in V^*, $\{e^{*j}\}_{j\in I}$. We have to show that the functionals $e^{*j} : V \to \mathbb{K}$ are generators of V^* and are linearly independent.
(Generators.) Let us show that, if $f : V \mapsto \mathbb{K}$ is linear, then there are numbers $c_j \in \mathbb{K}$ such that $f = \sum_{j\in I} c_j e^{*j}$.
To this end define $f_j := f(e_j)$, $j \in I$, then we argue that $f = f'$ where $f' := \sum_{j\in I} f_j e^{*j}$. Indeed, any $v \in V$ may be decomposed as $v = \sum_{i\in I} v^i e_i$ and, by linearity, we have:

$$f'(v) = \sum_{j\in I} f_j e^{*j}(\sum_{i\in I} v^i e_i) = \sum_{i,j\in I} f_j v^i e^{*j}(e_i) = \sum_{i,j\in I} f_j v^i \delta^j_i = \sum_{j\in I} v^j f_j = \sum_{j\in I} v^j f(e_j) = f(v)\,.$$

(Notice that above we have used the fact that one can extract the summation symbol from the argument of each e^{*j}, this is because the sum on the index i is *finite* by hypotheses it being

$dim V < +\infty$.) Since $f'(v) = f(v)$ holds for all of $v \in V$, we conclude that $f' = f$.
(Linear independence.) We have to show that if $\sum_{k\in I} c_k e^{*k} = 0$ then $c_k = 0$ for all $k \in I$.
To achieve that goal notice that $\sum_{k\in I} c_k e^{*k} = 0$ means $\sum_{k\in I} c_k e^{*k}(v) = 0$ for all $v \in V$. Therefore, putting $v = e_i$ and using the definition of the dual basis, $\sum_{k\in I} c_k e^{*k}(e_i) = 0$ turns out to be equivalent to $c_k \delta_i^k = 0$, namely, $c_i = 0$. This result can be produced for each $i \in I$ and thus $c_i = 0$ for all $i \in I$. The proof for the conjugate dual basis is very similar and is left to the reader. The proof for the conjugate basis uses the fact that $\overline{V}$ is the dual space of $\overline{V^*}$ and thus the first part of the proof applies. The last statement of the thesis holds because, in the four considered cases, the set of the indices I is the same. □

Notation 1.1. *From now on we take advantage of the following convention. Whenever an index appears twice, once as an upper index and once as a lower index, in whatever expression, the summation over the values of that index is understood in the notation. E.g.,*

$$t_{ijkl} f^{rsil}$$

means

$$\sum_{i,l} t_{ijkl} f^{rsil} \,.$$

The range of the various indices should be evident from the context or otherwise specified.

We are naturally lead to consider the following issue. It could seem that the definition of dual space, V^* (of a vector space V) may be implemented on V^* itself producing the double dual space $(V^*)^*$ and so on, obtaining for instance $((V^*)^*)^*$ and, by that way, an infinite sequence of dual vector spaces. The theorem below shows that this is not the case because $(V^*)^*$ turns out to be *naturally isomorphic* to the initial space V and thus the apparently infinite sequence of dual spaces ends on the second step. We remind the reader that a vector space isomorphism $F : V \to V'$ is a linear map which is also one-to-one, i.e., injective and surjective. An isomorphism is called *natural* when it is built up using the definition of the involved algebraic structures only and it does not depend on "arbitrary choices". A more precise definition of *natural isomorphism* may be given by introducing the *theory of mathematical categories*.

Theorem 1.2. *If V is a finite-dimensional vector space on the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, there is a natural isomorphism $F : V \to (V^*)^*$ given by*

$$(F(v))\,(u) := u(v)\,,$$

for all $u \in V^*$ and $v \in V$.

Proof. Notice that $F(v) \in (V^*)^*$ because it is a linear functional on V^*:

$$(F(v))(\alpha u + \beta u') := (\alpha u + \beta u')(v) = \alpha u(v) + \beta u'(v) =: \alpha(F(v))(u) + \beta(F(v))(u')\,.$$

Let us prove that $v \mapsto F(v)$ with $(F(v))(u) := u(v)$ is linear, injective and surjective.
(Linearity.) We have to show that, for all $\alpha, \beta \in \mathbb{K}$, $v, v' \in V$,

$$F(\alpha v + \beta v') = \alpha F(v) + \beta F(v') \,.$$

This is equivalent to, by the definition of F given above,

$$u(\alpha v + \beta v') = \alpha u(v) + \beta u(v') \,,$$

for all $u \in V^*$. This is obvious because, u is a linear functional on V.
(Injectivity.) We have to show that $F(v) = F(v')$ implies $v = v'$. $(F(v))(u) = (F(v'))(u)$ can be re-written as $u(v) = u(v')$ or $u(v - v') = 0$. In our hypotheses, this holds true for all $u \in V^*$. Then, define $e_1 := v - v'$ and notice that, if $v - v' \neq 0$, one can complete e_1 with other vectors to get a vector basis of V, $\{e_i\}_{i \in I}$. Since u is arbitrary, choosing $u = e^{*1}$ we should have $u(e_1) = 1$ by definition of dual basis but this contradicts $u(e_1) = 0$, i.e., $u(v - v') = 0$. By consequence $v - v' = 0$ and $v = v'$.
(Surjectivity.) We have to show that if $f \in (V^*)^*$, there is $v_f \in V$ such that $F(v_f) = f$.
Fix a basis $\{e_i\}_{i \in I}$ in V and the dual one in V^*. $v_f := f(e^{*i})e_i$ fulfills the requirement $F(v_f) = f$.
□

Def.1.3. (Pairing.) *Let V be a finite-dimensional vector space on $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ with dual space V^*. The bi-linear map $\langle\, .,\, .\rangle : V \times V^* \to \mathbb{K}$ such that*

$$\langle u, v \rangle := v(u)$$

for all $u \in V$, $v \in V^*$, is called **pairing**.

Comment. Because of the theorem proven above, we may indifferently think $\langle u, v \rangle$ as representing either the action of $u \in (V^*)^*$ on $v \in V^*$ or the action of $v \in V^*$ on $u \in V$.

Notation 1.2. From now on

$$V \simeq W$$

indicates that the vector spaces V and W are isomorphic under some natural isomorphism. If the field of V and W is $\mathbb{C}$,

$$V \overline{\simeq} W$$

indicates that there is a natural anti-isomorphism, i.e. there is an injective, surjective, anti-linear mapping $G : V \to W$ built up using the definition of V and $\overline{V}$ only.

Exercises 1.1.
1.1.1. Show that if $v \in V$ then $v = \langle v, e^{*j} \rangle e_j$, where $\{e_j\}_{j \in I}$ is any basis of the finite dimensional vector space V.
(Hint. Decompose $v = c^i e_i$, compute $\langle v, e^{*k} \rangle$ taking the linearity of the the left entrance into

account. Alternatively, show that $v - \langle v, e^{*j}\rangle e_j = 0$ proving that $f(v - \langle v, e^{*j}\rangle e_j) = 0$ for every $f \in V^*$.)

1.1.2. Show that $V^* \overline{\simeq} \overline{V^*}$ if $dimV < \infty$ (and the field of V is $\mathbb{C}$). Similarly, show that $V \overline{\simeq} \overline{V}$ under the same hypotheses.

(Hint. The anti-isomorphism $G : V^* \to \overline{V^*}$, is defined by $(G(v))(u) := \overline{\langle v, u\rangle}$ and the anti-isomorphism $F : V \to \overline{V}$, is defined by $(F(v))(u) := \langle v, G^{-1}(u)\rangle$.)

1.1.3. Show that if the finite-dimensional vector spaces V and V' are isomorphic or anti-isomorphic, then V^* and V'^* are so.

(Hint. If the initial (anti-) isomorphism is $F : V \to V'$ consider $G : V'^* \to V^*$ defined by $\langle F(u), v'\rangle = \langle u, G(v')\rangle$.)

1.1.4. Show that $\overline{V}^* \simeq \overline{V^*}$ if V is a finite-dimensional vector space on $\mathbb{C}$.

(Hint. Use Theorem 1.2.)

1.1.5. Write explicitly the (anti-)isomorphisms in **1.1.2**, **1.1.3**, **1.1.4**, with respect to bases $\{e_i\}_{i\in I} \subset V$, $\{\overline{e}_i\}_{i\in I} \subset \overline{V}$ and the corresponding ones in V^*, $\overline{V^*}$, $\overline{V}$, $\overline{V^*}^*$.

1.2 Multi linearity: tensor product, tensors, universality theorem.

Let us consider $n \geq 1$ vector spaces $V_1, \ldots, V_n$ on the common field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and another vector space W on $\mathbb{K}$. In the following $\mathcal{L}(V_1, \ldots, V_n|W)$ denotes the vector space (on $\mathbb{K}$) of *multi-linear maps* from $V_1 \times \ldots \times V_n$ to W. We recall the reader that a mapping $f : V_1 \times \ldots \times V_n \to W$ is said to be *multi linear* if, arbitrarily fixing $n-1$ vectors, $v_i \in V_i$ for $i \neq k$, each mapping $v_k \mapsto f(v_1, \ldots, v_k, \ldots, v_n)$ is linear for every $k = 1, \ldots, n$. We leave to the reader the trivial proof of the fact that $\mathcal{L}(V_1, \ldots, V_n|W)$ is a *vector space* on $\mathbb{K}$, with respect to the usual sum of pair of functions and product of an element of $\mathbb{K}$ and a function. If $W = \mathbb{K}$ we use the shorter notation $\mathcal{L}(V_1, \ldots, V_n) := \mathcal{L}(V_1, \ldots, V_n|\mathbb{K})$, $\mathbb{K}$ being the common field of $V_1, \ldots, V_n$.

Exercises 1.2.

1.2.1. Suppose that $\{e_{k,i_k}\}_{i_k\in I_k}$ are bases of the vector spaces V_k, $k = 1, \ldots, n$ on the same field $\mathbb{K}$. Suppose also that $\{e_i\}_{i\in I}$ is a basis of the vector space W on $\mathbb{K}$. Show that each $f \in \mathcal{L}(V_1, \ldots, V_n|W)$ satisfies

$$f(v_1, \ldots, v_n) = v_1^{i_1} \cdots v_n^{i_n} \langle f(e_{1,i_1}, \ldots, e_{n,i_n}), e^{*i}\rangle e_i \,,$$

for all $(v_1, \ldots, v_n) \in V_1 \times \cdots \times V_n$.

We have a first fundamental and remarkable theorem. To introduce it we employ three steps.
(1) Take $f \in \mathcal{L}(V_1, ..., V_n|W)$. This means that f associates every string $(v_1, ..., v_n) \in V_1 \times \cdots \times V_n$ with a corresponding element $f(v_1, \ldots, v_n) \in W$, and the map

$$(v_1, \ldots, v_n) \mapsto f(v_1, \ldots, v_n)$$

is multilinear.
(2) Since, for *fixed* $(v_1, \ldots, v_n)$, the vector $f(v_1, \ldots, v_n)$ is an element of W, the action of $w^* \in W^*$

on $f(v_1, \ldots, v_n)$ makes sense, producing $\langle f(v_1, \ldots, v_n), w^* \rangle \in \mathbb{K}$.
(3) Allowing $v_1, \ldots v_n$ and w^* to range freely in the corresponding spaces, the map Ψ_f with

$$\Psi_f : (v_1, \ldots, v_n, w^*) \mapsto \langle f(v_1, \ldots, v_n), w^* \rangle \,,$$

turns out to be multilinear by construction. Hence, by definition $\Psi_f \in \mathcal{L}(V_1, \ldots, V_n, W^*)$.
The theorem concerns the map F which associates f with Ψ_f.

Theorem 1.3. *If $V_1, \ldots, V_n$ are finite-dimensional vector space on the common field $\mathbb{K}$ ($= \mathbb{C}$ or $\mathbb{R}$), and W is another finite-dimensional vector spaces on $\mathbb{K}$, the vector spaces $\mathcal{L}(V_1, \ldots, V_n|W)$ and $\mathcal{L}(V_1, \ldots, V_n, W^*)$ are naturally isomorphic by means of the map $F : \mathcal{L}(V_1, \ldots, V_n|W) \to \mathcal{L}(V_1, \ldots, V_n, W^*)$ with $F : f \mapsto \Psi_f$ defined by*

$$\Psi_f(v_1, \ldots, v_n, w^*) := \langle f(v_1, \ldots, v_n), w^* \rangle \,,$$

for all $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$ and $w^* \in W^*$.

Proof. Let us consider the mapping F defined above. We have only to establish that F is linear, injective and surjective. This ends the proof.
(Linearity.) We have to prove that $\Psi_{\alpha f + \beta g} = \alpha \Psi_f + \beta \Psi_g$ for all $\alpha, \beta \in \mathbb{K}$ and $f, g \in \mathcal{L}(V_1, \ldots, V_n|W)$.
In fact, making use of the left-hand linearity of the pairing, one has

$$\Psi_{\alpha f + \beta g}(v_1, \ldots, v_n, w^*) = \langle (\alpha f + \beta g)(v_1, \ldots, v_n), w^* \rangle = \alpha \langle f(v_1, \ldots, v_n), w^* \rangle + \beta \langle g(v_1, \ldots, v_n), w^* \rangle$$

and this is nothing but:

$$\Psi_{\alpha f + \beta g}(v_1, \ldots, v_n, w^*) = (\alpha \Psi_f + \beta \Psi_g)(v_1, \ldots, v_n, w^*) \,.$$

Since $(v_1, \ldots, v_n, w^*)$ is arbitrary, the thesis is proven.
(Injectivity.) We have to show that if $\Psi_f = \Psi_{f'}$ then $f = f'$.
In fact, if $\Psi_f(v_1, \ldots, v_n, w^*) = \Psi_{f'}(v_1, \ldots, v_n, w^*)$ for all $(v_1, \ldots, v_n, w^*)$, using the definition of Ψ_g we have $\langle f(v_1, \ldots, v_n), w^* \rangle = \langle f'(v_1, \ldots, v_n), w^* \rangle$, or, equivalently

$$\langle f(v_1, \ldots, v_n) - f'(v_1, \ldots, v_n), w^* \rangle = 0 \,,$$

for all $(v_1, \ldots, v_n)$ and w^*. Then define $e_1 := f(v_1, \ldots, v_n) - f'(v_1, \ldots, v_n)$, if $e_1 \neq 0$ we can complete it to a basis of W. Fixing $w^* = e^{*1}$ we should have

$$\langle f(v_1, \ldots, v_n) - f'(v_1, \ldots, v_n), w^* \rangle = 1$$

which contradicts the statement above. Therefore $f(v_1, \ldots, v_n) - f'(v_1, \ldots, v_n) = 0$ for all $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$, in other words $f = f'$.
(Surjectivity.) We have to show that for each $\Phi \in \mathcal{L}(V_1, \ldots, V_n, W^*)$ there is a $f_\Phi \in \mathcal{L}(V_1, \ldots, V_n|W)$ with $\Psi_{f_\Phi} = \Phi$.

To this end, fix vector bases $\{e_{r,i_r}\}_{i_r\in I_r} \subset V_r$ and a (dual) basis $\{e^{*k}\}_{k\in I} \subset W^*$. Then, take $\Phi \in \mathcal{L}(V_1,\ldots,V_n,W^*)$ and define the mapping $f_\Phi \in \mathcal{L}(V_1,\ldots,V_n|W)$ given by

$$f_\Phi(v_1,\ldots,v_n) := v_1^{i_1}\cdots v_n^{i_n}\Phi(e_{1,i_1},\ldots,e_{n,i_n},e^{*k})e_k\ .$$

By construction that mapping is multilinear and, using multilinearity and Exercise **1.2.1**, we find

$$\Psi_{f_\Phi}(v_1,\ldots,v_n,w^*) = \langle v_1^{i_1}\cdots v_n^{i_n}\Phi(e_{1,i_1},\ldots,e_{n,i_n},e^{*k})e_k, w_h^* e^{*h}\rangle = \Phi(v_1,\ldots,v_n,w^*)\ ,$$

for all $(v_1,\ldots,v_n,w^*)$ and this is equivalent to $\Psi_{f_\Phi} = \Phi$. $\square$

Important Remark.
An overall difficult point in understanding and trusting in the statement of the theorem above relies upon the fact that the function f has n entries, whereas the function Ψ_f has $n+1$ entries but f is identified to Ψ_f by means of F. This fact may seem quite weird at first glance and the statement of the theorem may seem suspicious by consequence. The difficulty can be clarified from a practical point of view as follows.
If bases $\{e_{1,i_1}\}_{i_1\in I_1}$,...,$\{e_{n,i_n}\}_{i_n\in I_n}$, for V_1,...,V_n respectively, and $\{e_k\}_{k\in I}$ for W are fixed and $\Psi \in \mathcal{L}(V_1,\ldots,V_n,W^*)$, one has, for constant coefficients $P^k_{i_1\ldots i_n}$ depending on Ψ,

$$\Psi(v_1,\ldots,v_n,w^*) = P^k_{i_1\ldots i_n} v_1^{i_1}\ldots v_n^{i_n} w_k^*\ ,$$

where w_k^* are the components of w^* in the dual basis $\{e^{*k}\}_{k\in I}$ and $v_p^{i_p}$ the components of v_p in the basis $\{e_{p,i_p}\}_{i_p\in I_p}$. Now consider the other map $f \in \mathcal{L}(V_1,\ldots,V_n|W)$ whose argument is the string $(v_1,\ldots,v_n)$, *no further vectors being necessary*:

$$f : (v_1,\ldots,v_n) \mapsto P^k_{i_1\ldots i_n} v_1^{i_1}\ldots v_n^{i_n} e_k$$

The former mapping Ψ, which deals with $n+1$ arguments, is associated with f which deals with n arguments. The point is that we have taken advantage of the bases, $\{e_k\}_{k\in I_k}$ in particular. Within this context, Theorem 1.3 just proves that *the correspondence which associates Ψ to f is linear, bijective and independent from the chosen bases.*

Theorem 1.3 allow us to restrict our study to the spaces of multi-linear *functionals* $\mathcal{L}(V_1,\ldots,V_k)$, since the spaces of multi-linear *maps* are completely encompassed. Let us introduce the concept of *tensor product* of vector spaces. The following definitions can be extended to encompass the case of non-finite dimensional vector spaces by introducing suitable topological notions (e.g., Hilbert spaces). To introduce the tensor product we need a trivial but relevant preliminary definition.

Def.1.4. (Linear action of a vector space.) *Let U and V two finite-dimensional vector space on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We say that U* **linearly acts on V by means of F**, *if*

$F : U \to V^*$ is a natural isomorphism.

Examples.
(1) If V is a finite-dimensional vector space on $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, V^* linearly acts on V by $F = id_{V^*}$.
(2) If V is that as in **(1)**, V linearly acts on V^* by means of the natural isomorphism between V and $(V^*)^*$.

In the following if U linearly acts on V by means of F, and $u \in U$, $v \in V$ we write $u(v)$ instead of $(F(u))(v)$ whenever F is unambiguously determined. For instance, in the second example above, $v(u)$ means $\langle v, u\rangle$ when $v \in V$ and $u \in V^*$.

Def.1.5. (Tensor product.) *Let $U_1, \ldots, U_n$ be $n \geq 1$ finite-dimensional vector spaces on the common field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, suppose each U_i linearly acts on V_i (by means of some F_i), for $i = 1, \ldots, n$.*
(1) *if $(u_1, \ldots, u_n) \in U_1 \times \ldots \times U_n$, $u_1 \otimes \ldots \otimes u_n$ denotes the multi linear mapping in $\mathfrak{L}(V_1, \ldots, V_n)$ defined by*

$$(u_1 \otimes \ldots \otimes u_n)(v_1, \ldots, v_n) := u_1(v_1) \cdots u_n(v_n) \,,$$

for all $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$. $u_1 \otimes \ldots \otimes u_n$ is called **tensor product of vectors** *$u_1, \ldots, u_n$.*
(2) *The mapping $\otimes : U_1 \times \ldots \times U_n \to \mathfrak{L}(V_1, \ldots, V_n)$ given by: $\otimes : (u_1, \ldots, u_n) \mapsto u_1 \otimes \ldots \otimes u_n$, is called* **tensor product map**.
(3) *The vector subspace of $\mathfrak{L}(V_1, \ldots, V_n)$ generated by all of $u_1 \otimes \ldots \otimes u_n$ for all $(u_1, \ldots, u_n) \in U_1 \times \ldots \times U_n$ is called* **tensor product of spaces** *$U_1, \ldots, U_n$ and is indicated by $U_1 \otimes \ldots \otimes U_n$. The vectors in $U_1 \otimes \ldots \otimes U_n$ are called* **tensors**.

Remarks.
(1) $U_1 \otimes \ldots \otimes U_n$ is made of all the linear combinations of the form $\sum_{j=1}^N \alpha_j u_{1,j} \otimes \ldots \otimes u_{n,j}$, where $\alpha_j \in \mathbb{K}$, $u_{k,j} \in U_k$ and $N = 1, 2, \ldots$.
(2) It is trivially proven that the tensor product map:

$$(u_1, \ldots, u_n) \mapsto u_1 \otimes \ldots \otimes u_n \,,$$

is *multi linear*.
That is, for any $k \in \{1, 2, \ldots, n\}$, the following identity holds for all $u, v \in V_k$ and $\alpha, \beta \in \mathbb{K}$:

$$\begin{aligned} & u_1 \otimes \ldots \otimes u_{k-1} \otimes (\alpha u + \beta v) \otimes u_{k+1} \otimes \ldots \otimes u_k \\ & = \alpha(u_1 \otimes \ldots \otimes u_{k-1} \otimes u \otimes u_{k+1} \otimes \ldots \otimes u_k) \\ & + \beta(u_1 \otimes \ldots \otimes u_{k-1} \otimes v \otimes u_{k+1} \otimes \ldots \otimes u_k) \,. \end{aligned}$$

As a consequence, it holds

$$(\alpha u) \otimes v = u \otimes (\alpha v) = \alpha(u \otimes v) \,,$$

and similar identites hold considering whatever number of factor spaces in a tensor product of vector spaces.

(3) From the given definition, if $V_1, \ldots, V_n$ are given finite-dimensional vector spaces on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, it is clear what we mean by $V_1 \otimes \ldots \otimes V_n$ or $V_1^* \otimes \ldots \otimes V_n^*$ but also, for instance, $V_1^* \otimes V_2 \otimes V_3$ or $V_1 \otimes V_2^* \otimes V_3^*$. One simply has to take the natural linear action of V on V^* and V^* on V into account.

The natural question which arises is if $U_1 \otimes \ldots \otimes U_n$ is a *proper* subspace of $\mathcal{L}(V_1, \ldots, V_n)$ or, conversely, recovers the whole space $\mathcal{L}(V_1, \ldots, V_n)$. The following theorem gives an answer to that question.

Theorem 1.4. *Referring to* **Def.1.4**, *it holds:*

$$U_1 \otimes \ldots \otimes U_n = \mathcal{L}(V_1, \ldots, V_n) \, .$$

Proof. It is sufficient to show that if $f \in \mathcal{L}(V_1, \ldots, V_n)$ then $f \in U_1 \otimes \ldots \otimes U_n$. To this end fix bases $\{e_{k,i}\}_{i \in I_k} \subset V_k$ for $k = 1, \ldots, n$. f above is completely determined by coefficients $f_{i_1,\ldots,i_n} := f(e_{1,i_1}, \ldots, e_{n,i_n})$ since, every $v_k \in V_k$ can be decomposed as $v_k = v_k^{i_k} e_{ki_k}$ and thus, by multi linearity:

$$f(v_1, \ldots, v_n) = v_1^{i_1} \cdots v_n^{i_n} f(e_{1,i_1}, \ldots, e_{n,i_n}) \, .$$

Thenconsider the tensor $t_f \in U_1 \otimes \ldots \otimes U_n$ defined by:

$$t_f := f_{i_1 \ldots i_n} e_1^{*i_1} \otimes \ldots \otimes e_n^{*i_n} \, ,$$

where we have identified each U_k with the corresponding space V_k^* which linearly acts on V_k. Then, by **Def.1.5**, one can directly prove that, by multi linearity

$$t_f(v_1, \ldots, v_n) = v_1^{i_1} \ldots v_n^{i_n} f_{i_1 \ldots i_n} = f(v_1, \ldots, v_n) \, ,$$

for all of $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$. This is nothing but $t_f = f$. $\square$

Another relevant result is stated by the theorem below.

Theorem 1.5. *Referring to* **Def.1.4**, *the following statements hold.*
(a) *The dimension of* $U_1 \otimes \ldots \otimes U_n$ *is:*

$$dim(U_1 \otimes \ldots \otimes U_n) = \prod_{k=1}^{n} dim U_k = \prod_{k=1}^{n} dim V_k \, .$$

(b) *If* $\{e_{k,i_k}\}_{i_k \in I_k}$ *is a basis of* U_k, $k = 1, \ldots, n$, *then :* $\{e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}\}_{(i_1,\ldots,i_n) \in I_1 \times \ldots \times I_n}$ *is a vector basis of* $U_1 \otimes \ldots \otimes U_n$.
(c) *If* $t \in U_1 \otimes \ldots \otimes U_n$, *the components of* t *with respect to a basis* $\{e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}\}_{(i_1,\ldots,i_n) \in I_1 \times \ldots \times I_n}$ *are given by (identifying each* V_i *with the corresponding* U_i^**):*

$$t^{i_1 \ldots i_n} = t(e_1^{*i_1}, \ldots, e_n^{*i_n})$$

and thus it holds

$$t = t^{i_1 \ldots i_n} e_{1,i_1} \otimes \ldots \otimes e_{n,i_n} \, .$$

Proof. Notice that **(b)** trivially implies **(a)** because the elements $e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}$ are exactly $\prod_{k=1}^n dim U_k$. (Also $\prod_{k=1}^n dim U_k = \prod_{k=1}^n dim V_k$ because each U_k is isomorphic to the corresponding V_k^* and $dim V_k = dim V_k^*$ by definition of dual basis.) So it is sufficient to show that the second statement holds true. To this end, since elements $e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}$ are generators of $U_1 \otimes \ldots \otimes U_n$, it is sufficient to show that they are linearly independent. Consider the generic vanishing linear combination

$$C^{i_1 \ldots i_n} e_{1,i_1} \otimes \ldots \otimes e_{n,i_n} = 0 \, ,$$

We want to show that all of the coefficients $C^{i_1 \ldots i_n}$ vanish. We may identify each V_k with the corresponding U_k^* by means of the natural isomorphism of U and $(U^*)^*$ and thus we may consider each dual-basis vector $e_r^{*j_r} \in U_j^*$ as an element of V_j. Then the action of that linear combination of multi-linear functionals on the generic element $(e_1^{*j_1}, \ldots, e_n^{*j_n}) \in V_1 \times \cdots \times V_n$ produces the result

$$C^{j_1 \ldots j_n} = 0 \, .$$

Since we can arbitrarily fix the indices $j_1, \ldots j_n$ this proves the thesis.
Concerning **(c)**, by the uniqueness of the components of a vector with respect to a basis, it is sufficient to show that, defining

$$t' := t^{i_1 \ldots i_n} e_{1,i_1} \otimes \ldots \otimes e_{n,i_n} \, ,$$

where

$$t^{i_1 \ldots i_n} := t(e_1^{*i_1}, \ldots, e_n^{*i_n}) \, ,$$

it holds

$$t'(v_1, \ldots, v_n) = t(v_1, \ldots, v_n) \, ,$$

for all $(v_1, \ldots, v_n) \in V_1^* \times \ldots \times V_n^*$. By multi linearity,

$$t(v_1, \ldots, v_n) = v_{1\,i_1} \cdots v_{n\,i_n} t(e_1^{*i_1}, \ldots, e_n^{*i_n}) \, ,$$

By multi linearity, the left-hand side is nothing but $t'(v_1, \ldots, v_n)$. This concludes the proof. □

There are two important theorems which, *together with the identification of* V *and* $(V^*)^*$, imply that all of the spaces which one may build up coherently using the symbols $\otimes$, V_k, * and () are naturally isomorphic to spaces which are of the form $V_{i_1}^{(*)} \otimes \ldots \otimes V_{i_n}^{(*)}$. The rule to produce spaces naturally isomorphic to a given initial space is that one has to (1) ignore parentheses, (2) assume that * is *distributive with respect to* $\otimes$ and (3) assume that * is *involutive* (i.e. $(X^*)^* \simeq X$). For instance, one has:

$$((V_1^* \otimes V_2) \otimes (V_3 \otimes V_4^*))^* \simeq V_1 \otimes V_2^* \otimes V_3^* \otimes V_4 \, .$$

Let us state the theorems corresponding to the rules (2) and (1) respectively (the rule (3) being nothing but **Theorem 1.2**).

Theorem 1.6. ** is* **distributive** *with respect to $\otimes$ by means of* **natural isomorphisms** *$F : V_1^* \otimes \ldots \otimes V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$. In other words, if $V_1 \ldots V_n$ are finite dimensional vector spaces on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, it holds:*

$$(V_1 \otimes \ldots \otimes V_n)^* \simeq V_1^* \otimes \ldots \otimes V_n^* \,.$$

under F and furthermore

$$\langle u_1 \otimes \ldots \otimes u_n, F(v_1^* \otimes \ldots \otimes v_n^*)\rangle = \langle u_1, v_1^*\rangle \cdots \langle u_n, v_n^*\rangle \,,$$

for every choice of $v_i^* \in V_i^*$, $u_i \in V_i$ and $i = 1, 2, \ldots, n$.

Theorem 1.7. *The tensor product is* **associative** *by means of* **natural isomorphisms**. *In other words considering a space which is made of tensor products of tensor products of finite-dimensional vector spaces on the same field $\mathbb{R}$ or $\mathbb{C}$, one may omit parenthesis everywhere obtaining a space which is naturally isomorphic to the initial one. So, for instance*

$$V_1 \otimes V_2 \otimes V_3 \simeq V_1 \otimes (V_2 \otimes V_3) \,.$$

where the natural isomprphism $F_1 : V_1 \otimes V_2 \otimes V_3 \to V_1 \otimes (V_2 \otimes V_3)$ satisfies

$$F_1 : v_1 \otimes v_2 \otimes v_3 \to v_1 \otimes (v_2 \otimes v_3)$$

for every choice of $v_i \in V_i$, $i = 1, 2, 3$.

Sketch of proof. The natural isomorphisms required in the two theorems above can be built up as follows. In the former case, consider a linear mapping $F : V_n^* \otimes \ldots \otimes V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$ which satisfies

$$F(v_1^* \otimes \ldots \otimes v_n^*) \in (V_1 \otimes \ldots \otimes V_n)^* \,,$$

where, for all $u_1 \otimes \ldots \otimes u_n \in V_1 \otimes \ldots \otimes V_n$,

$$F(v_1^* \otimes \ldots \otimes v_n^*)(u_1 \otimes \ldots \otimes u_n) := \langle u_1, v_1^*\rangle \ldots \langle u_n, v_n^*\rangle \,. \tag{1}$$

In the latter case, consider the first proposed example. The required natural isomorphism can be build up as a linear mapping $F_1 : V_1 \otimes V_2 \otimes V_3 \to (V_1 \otimes V_2) \otimes V_3$ such that

$$F_1 : v_1 \otimes v_2 \otimes v_3 \mapsto (v_1 \otimes v_2) \otimes v_3 \,,$$

for all $v_1 \otimes v_2 \otimes v_3 \in V_1 \otimes V_2 \otimes V_3$.
Using the linearity and the involved definitions the reader can simply prove that the mappings F, F_1 are surjective. Moreover, by the previously proven theorems

$$dim(V_1 \otimes \cdots \otimes V_n)^* = dim(V_1 \otimes \cdots \otimes V_n) = dim(V_1^* \otimes \cdots \otimes V_n^*)$$

and

$$dim((V_1 \otimes V_2) \otimes V_3) = dim(V_1 \otimes V_2) \cdot dim(V_3) = dim(V_1 \otimes V_2 \otimes V_3)\,.$$

As a consequence, F and F_1 must be also injective and thus they are isomorphisms. $\Box$

Very important remark. The central point is that the mappings F and F_1 above have been given by specifying their action on tensor products of elements (e.g, $F(v_1 \otimes \ldots \otimes v_n)$) and not on *linear combinations* of these tensor products of elements. Remind that, for instance $V_1^* \otimes \ldots \otimes V_n^*$ is not the set of products $u_1^* \otimes \ldots \otimes u_n^*$ but it is the set of *linear combinations* of those products. Hence, in order to completely define F and F_1, one must require that F and F_1 admit *uniquely determined linear extensions* on their initial domains in order to encompass the whole tensor spaces generated by linear combinations of simple tensor products. In other words one has to complete the given definition, in the former case, by adding the further requirement

$$F(\alpha\, u_1^* \otimes \ldots \otimes u_n^* + \beta\, v_1^* \otimes \ldots \otimes v_n^*) = \alpha F(u_1^* \otimes \ldots \otimes u_n^*) + \beta F(v_1^* \otimes \ldots \otimes v_n^*)\,,$$

and similarly in the latter case. Despite it could seem a trivial task it is not the case. Indeed it is worthwhile noticing that each product $v_1^* \otimes \ldots \otimes v_n^*$ can be re-written using linear combinations, e.g., concerning F:

$$v_1^* \otimes \ldots \otimes v_n^* = [(v_1^* + u_1^*) \otimes \ldots \otimes v_n^*] \;-\; [u_1^* \otimes \ldots \otimes v_n^*]\,.$$

Now consider the identities, which has to hold as a consequence of the assumed linearity of F, such as:

$$F(v_1^* \otimes \ldots \otimes v_n^*) = F((v_1^* + u_1^*) \otimes \ldots \otimes v_n^*) - F(u_1^* \otimes \ldots \otimes v_n^*)$$

Above $F(v_1^* \otimes \ldots \otimes v_n^*)$, $F(u_1^* \otimes \ldots \otimes v_n^*)$, $F((v_1^* + u_1^*) \otimes \ldots \otimes v_n^*)$ are *independently* defined as we said at the beginning and there is no reason, in principle, for the validity of the constraint:

$$F(v_1^* \otimes \ldots \otimes v_n^*) = F((v_1^* + u_1^*) \otimes \ldots \otimes v_n^*) - F(u_1^* \otimes \ldots \otimes v_n^*)\,.$$

Similar problems arise concerning F_1.
The general problem which arises by the two considered cases can be stated as follows. Suppose we are given a tensor product of vector spaces $V_1 \otimes \ldots \otimes V_n$ and we are interested in the possible *linear extensions* of a mapping f on $V_1 \otimes \ldots \otimes V_n$, with values in some vector space W, when f is initially defined on simple products $v_1 \otimes \ldots \otimes v_n \in V_1 \otimes \ldots \otimes V_n$ only.
Is there any general prescription on the specification of values $f(v_1 \otimes \ldots \otimes v_n)$ which assures that f can be extended, uniquely*, to a linear mapping from $V_1 \otimes \ldots \otimes V_n$ to W?*
An answer is given by the following very important **universality theorem.**

Theorem 1.8. (Universality Theorem.) *Given $n \geq 1$ finite-dimensional vector spaces $U_1, \ldots, U_n$ on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, the following statements hold.*
(a) *For each finite-dimensional vector space W and each multi-linear mapping $f : U_1 \times \ldots \times U_n \to$*

W, there is a unique linear mapping $f^{\otimes} : U_1 \otimes \ldots \otimes U_n \to W$ such that the diagram below **commutes** *(in other words, $f^{\otimes} \circ \otimes = f$).*

$$\begin{tikzcd} U_1 \times \ldots \times U_n \arrow[r, "\otimes"] \arrow[dr, "f"] & U_1 \otimes \ldots \otimes U_n \arrow[d, "f^{\otimes}"] \\ & W \end{tikzcd}$$

(b) *For $U_1, \ldots, U_n$ fixed as above, suppose there is another pair (T, U_T) such that, for each W and each f with $f : U_1 \times \ldots \times U_n \to W$ multi linear, there is a unique linear map $f^T : U_T \to W$ such that the diagram below commute,*

$$\begin{tikzcd} U_1 \times \ldots \times U_n \arrow[r, "T"] \arrow[dr, "f"] & U_T \arrow[d, "f^T"] \\ & W \end{tikzcd}$$

then there is a natural vector space isomorphism $\phi : U_1 \otimes \ldots \otimes U_n \to U_T$ with $f^T \circ \phi = f^{\otimes}$. In other words, given $U_1, \ldots, U_n$ as in the hypotheses, the pair $(\otimes,\, U_1 \otimes \ldots \otimes U_n)$ is determined up to vector spaces isomorphisms by the diagram property above.

Remark. The property of the pair $(\otimes,\, U_1 \otimes \ldots \otimes U_n)$ stated in the second part of the theorem is called **universality property**.

Before proving the theorem, let us explain how it gives a precise answer to the stated question. The theorem says that a *linear extension* of any function f with values in W, initially defined on simple tensor products only, $f(v_1 \otimes \ldots \otimes v_n)$ with $v_1 \otimes \ldots \otimes v_n \in U_1 \otimes \ldots \otimes U_n$, does *exist* on the whole domain space $U_1 \otimes \ldots \otimes U_n$ and it is *uniquely* determined provided $f(v_1 \otimes \ldots \otimes v_n) = g(v_1, \ldots, v_n)$ where $g : U_1 \times \ldots \times U_n \to W$ is some *multi-linear function.*
Concerning the mappings F and F_1 introduced above, we may profitably use the universality theorem to show that they are well-defined on the whole domain made of linear combinations of simple tensor products. In fact, consider F for example. We can define the *multi-linear* mapping $G : V_1^* \times \ldots \times V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$ such that $G(v_1^*, \ldots, v_n^*)$ it the unique multi-linear function with:

$$\left(G(v_1^*, \ldots, v_n^*)\right)(u_1, \ldots, u_n) := \langle u_1, v_1^* \rangle \cdots \langle u_n, v_n^* \rangle \,, \quad \text{for all } (u_1, \ldots, u_n) \in U_1 \times \cdots \times U_n \tag{2}$$

(The reader should check on the well-posedness and cd multi linearity of G.) Then the universality theorem with $U_k = V_k^*$ and $W = (V_1 \otimes \ldots \otimes V_n)^*$ assures the existence of a linear mapping $F : V_1^* \otimes \ldots \otimes V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$ with the required properties because $F := G^{\otimes}$ is such that

$$F(v_1^* \otimes \ldots \otimes v_n^*) = (G^{\otimes} \circ \otimes)(v_1^*, \ldots, v_n^*) = G(v_1^*, \ldots, v_n^*) \,,$$

si that (1) is fulfilled due to (2). A similar multi-linear mapping G_1 can be found for F_1:

$$G_1 : (v_1, v_2, v_3) \mapsto (v_1 \otimes v_2) \otimes v_3 \, .$$

$F_1 := G_1^{\otimes}$ can be used in the proof of Theorem 1.7.

Proof of **Theorem 1.8**. **(a)** Fix bases $\{e_{k,i_k}\}_{i_k \in I_k} \subset U_k$, $k = 1, \ldots, n$ and $\{e_j\}_{j \in I} \subset W$. Because of the multi-linearity property, the specification of coefficients $f^j_{i_1 \ldots i_n} \in \mathbb{K}$, $i_k \in I_k, j \in I$, $k = 1, \ldots, n$ uniquely determines a multi-linear mapping $f : U_1 \times \ldots \times U_n \to W$ by defining $f^j_{i_1 \ldots i_n} := \langle f(e_{1,i_1}, \ldots, e_{n,i_n}), e^{*j} \rangle$. Conversely, a multi-linear mapping $f : U_1 \times \ldots \times U_n \to W$ uniquely determines coefficients $f^j_{i_1 \ldots i_n} \in \mathbb{K}$ by the same rule. Similarly, because of the linearity property, a linear mapping $g : U_1 \otimes \ldots \otimes U_n \to W$ is completely and uniquely determined by coefficients $g^j_{i_1 \ldots i_n} := \langle g(e_{i_1} \otimes \ldots \otimes e_{i_n}), e^j \rangle$. Therefore, if $f : U_1 \times \ldots \times U_n \to W$ is given and it is determined by coefficients $f^j_{i_1 \ldots i_n}$ as above, defining $f^{\otimes} : U_1 \otimes \ldots \otimes U_n \to W$ by the coefficients $(f^{\otimes})^j_{i_1 \ldots i_n} := f^j_{i_1 \ldots i_n}$, we have

$$\langle f^{\otimes}(v_1 \otimes \ldots \otimes v_n), e^{*j} \rangle = v^{i_1} \ldots v^{i_n} f^j_{i_1 \ldots i_n} = \langle f(v_1, \ldots, v_n), e^{*j} \rangle,$$

for all $(v_1, \ldots, v_n) \in U_1 \times \ldots \times U_n$ and $j \in I$. This means that $f^{\otimes}(v_1 \otimes \ldots \otimes v_n) = f(v_1, \ldots, v_n)$ for all $(v_1, \ldots, v_n) \in U_1 \times \ldots \times U_n$. In other words $f^{\otimes} \circ \otimes = f$. The uniqueness of the mapping $f^{\otimes}$ is obvious: suppose there is another mapping $g^{\otimes}$ with $g^{\otimes} \circ \otimes = f$ then $(f^{\otimes} - g^{\otimes}) \circ \otimes = 0$. This means in particular that

$$\langle (f^{\otimes} - g^{\otimes})(e_{i_1} \otimes \ldots \otimes e_{i_n}), e^{*j} \rangle = 0 \, ,$$

for all $i_k \in I_k$, $k = 1, \ldots, n$ and $j \in I$. Since the coefficients above completely determine a map, the considered mapping must be the null mapping and thus: $f^{\otimes} = g^{\otimes}$.

(b) Suppose that there is a pair (T, U_T) as specified in the hypotheses. Then, using the part **(a)** with $f = T$ and $W = U_T$ we have the diagram below ,where we have a former diagram commutation relation:

$$T^{\otimes} \circ \otimes = T.$$

$$\begin{tikzcd}
U_1 \times \ldots \times U_n \arrow[r, "\otimes"] \arrow[dr, "T"] & U_1 \otimes \ldots \otimes U_n \arrow[d, "T^{\otimes}"] \\
& U_T
\end{tikzcd}$$

On the other hand, using the analogous property of the pair (T, U_T) with $f = \otimes$ and $W = U_1 \otimes \ldots \otimes U_n$ we also have the commutative diagram

$$\begin{tikzcd}
U_1 \times \ldots \times U_n \arrow[r, "T"] \arrow[dr, "\otimes"] & U_T \arrow[d, "\otimes^T"] \\
& U_1 \otimes \ldots \otimes U_n
\end{tikzcd}$$

which involves a second diagram commutation relation:

$$\otimes^T \circ T = \otimes \, .$$

The two obtained relations imply

$$(T^{\otimes} \circ \otimes^T) \circ T = T$$

and

$$(\otimes^T \circ T^{\otimes}) \circ \otimes = \otimes \, ,$$

In other words, if $RanT \subset U_T$ is the range of the mapping T and $Ran\otimes \subset U_1 \otimes \ldots \otimes U_n$ is the analogue for $\otimes$:

$$(T^{\otimes} \circ \otimes^T)|_{RanT} = Id_{RanT}$$

and

$$(\otimes^T \circ T^{\otimes})|_{Ran\otimes} = Id_{Ran\otimes}$$

Then consider the latter. Notice that $Ran\otimes$ is not a subspace, but the subspace spanned by $Ran\otimes$ coincides with $U_1 \otimes \ldots \otimes U_n$ by definition of tensorial product. Since $\otimes^T \circ T^{\otimes}$ is linear (because composition of linear maps) and defined on the whole space $U_1 \otimes \ldots \otimes U_n$, we conclude that

$$\otimes^T \circ T^{\otimes} = Id_{U_1 \otimes \ldots \otimes U_n} \, .$$

Concerning the former identity we could conclude analogously, i.e.,

$$T^{\otimes} \circ \otimes^T = Id_{U_T} \, ,$$

provided the subspace spanned by $RanT$, $Span(RanT)$, coincides with the whole space U_T. Anyway this holds true by the uniqueness property of f^T for a fixed f (see hypotheses about the pair (T, U_T)). (If $Span(RanT) \neq U_T$, then $U_T = Span(RanT) \oplus S$, $S \neq \{0\}$ being some proper subspace of U_T with $S \cap Span(RanT) = \{0\}$. Then, a mapping f^T could not be uniquely determined by the requirement $f^T \circ T = f$ because such a relation would be preserved under modifications of f^T on the subspace S).
We conclude that the linear mapping $\phi := T^{\otimes} : U_1 \otimes \ldots \otimes U_n \to U_T$ is a natural isomorphism with inverse $\otimes^T$.
As far as the property $f^T \circ \phi = f^{\otimes}$ is concerned we may deal with as follows. Since $\phi : U_1 \otimes \ldots \otimes U_n \to U_T$ is an isomorphism, $f^T \circ T = f$ implies $(f^T \circ \phi) \circ (\phi^{-1} \circ T) = f$, but $\phi^{-1} \circ T = \otimes^T \circ T = \otimes$, therefore $(f^T \circ \phi) \circ \otimes = f$. By the uniqueness of the mapping $f^{\otimes}$ which satisfies the same identity, it has to hold $f^T \circ \phi = f^{\otimes}$. $\square$

Exercises 1.3.
1.3.1. Consider a finite-dimensional vector space V and its dual V^*. Show by the universality theorem that there is a natural isomorphism such that

$$V \otimes V^* \simeq V^* \otimes V \, .$$

(Hint. Consider the bilinear mapping $f : V \times V^* \to V^* \otimes V$ with $f : (v_1, v_2^*) \mapsto v_2^* \otimes v_1$. Show that $f^{\otimes}$ is injective and thus surjective because $dim(V^* \otimes V) = dim(V \otimes V^*)$.)

2 Tensor algebra. Abstract Index Notation.

2.1 Tensor algebra generated by a vector space.

Let V be a finite-dimensional vector space on the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Equipped with this algebraic structure, in principle, we may build up several different tensor spaces. Notice that we have to consider also $\mathbb{K}$, $\mathbb{K}^*$ and V^* as admissible tensor factors. (In the following we shall not interested in conjugate spaces.) Obviously, we are interested in tensor products which are not identifiable by some natural isomorphism.
First consider the dual $\mathbb{K}^*$ when we consider $\mathbb{K}$ as a vector space on the field $\mathbb{K}$ itself. $\mathbb{K}^*$ is made of linear functionals from $\mathbb{K}$ to $\mathbb{K}$. Each $c^* \in \mathbb{K}^*$ has the form $c^*(k) := c \cdot k$, for all $k \in \mathbb{K}$, where $c \in \mathbb{K}$ is a fixed field element which completely determines c^*. The mapping $c \mapsto c^*$ is a (natural) vector space isomorphism. Therefore

$$\mathbb{K} \simeq \mathbb{K}^* \,.$$

Then we pass to consider products $\mathbb{K} \otimes \ldots \otimes \mathbb{K} \simeq \mathbb{K}^* \otimes \ldots \otimes \mathbb{K}^* = \mathcal{L}(\mathbb{K}, \ldots, \mathbb{K})$. Each multi-linear mapping $f \in \mathcal{L}(\mathbb{K}, \ldots, \mathbb{K})$ is completely determined by the number $f(1, \ldots, 1)$, since $f(k_1, \ldots, k_n) = k_1 \cdots k_n f(1, \ldots, 1)$. One can trivially show that the mapping $f \mapsto f(1, \ldots, 1)$ is a (natural) vector space isomorphism between $\mathbb{K} \otimes \ldots \otimes \mathbb{K}$ and $\mathbb{K}$ itself. Therefore

$$\mathbb{K} \otimes \ldots \otimes \mathbb{K} \simeq \mathbb{K} \,.$$

Also notice that the found isomorphism trivially satisfies $c_1 \otimes \ldots \otimes c_n \mapsto c_1 \cdots c_n$, and thus the tensor product mapping reduces to the ordinary product of the field.
We pass to consider the product $\mathbb{K} \otimes V = \mathcal{L}(\mathbb{K}^*, V^*) \simeq \mathcal{L}(\mathbb{K}, V^*)$. Each multi-linear functional f in $\mathcal{L}(\mathbb{K}, V^*)$ is completely determined by the element of V, $f(1, \cdot) : V^* \to \mathbb{K}$, which maps each $v^* \in V^*$ in $f(1, v^*)$. Once again, it is a trivial task to show that $f \mapsto f(1, \cdot)$ is a (natural) vector space isomorphism between $\mathbb{K} \otimes V$ and V itself.

$$\mathbb{K} \otimes V \simeq V \,.$$

Notice that the found isomorphism satisfies $k \otimes v \mapsto kv$ and thus the tensor product mapping reduces to the ordinary product of a field element and a vector.
Concluding, only the spaces $\mathbb{K}$, V, V^* and whatever products of V and V^* may be significantly different. This result leads us to the following definition.

Def.2.1. (Tensor Algebra generated by V.) *Let V be a finite-dimensional vector space with field $\mathbb{K}$.*
(1) *The* **tensor algebra** $\mathcal{A}_{\mathbb{K}}(V)$ *generated by V with field $\mathbb{K}$ is the class whose elements are the vector spaces: $\mathbb{K}$, V, V^* and all of tensor products of factors V and V^* in whatever order and number.*
(2) *The tensors of $\mathbb{K}$ are called* **scalars**, *the tensors of V are called* **contravariant vectors**, *the tensors of V^* are called* **covariant vectors**, *the tensors of spaces $V^{n\otimes} := V \otimes \ldots \otimes V$,*

where V appears $n \geq 1$ times, are called **contravariant tensors of order** n *or* **tensors of order** $(n, 0)$*, the tensors of spaces $V^{*n\otimes} := V^* \otimes \ldots \otimes V^*$, where V^* appears $n \geq 1$ times, are called* **covariant tensors of order** n *or* **tensors of order** $(0, n)$*, the remaining tensors which belong to spaces containing n factors V and m factors V^* are called* tensors of order (n, m) *not depending on the order of factors.*

Remark. Obviously, pairs of tensors spaces made of the same number of factors V and V^* in different order, are naturally isomorphic (see **exercise 1.3.1**). However, for practical reasons it is convenient to consider these spaces as different spaces and use the identifications when and if necessary.

2.2 The abstract index notation and rules to handle tensors.

Let us introduce the **abstract index notation**. Consider a finite-dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. After specification of a basis $\{e_i\}_{i\in I} \subset V$ and the corresponding basis in V^*, each tensor is completely determined by giving its components with respect to the induced basis in the corresponding tensor space in $\mathcal{A}_{\mathbb{K}}(V)$. We are interested in the transformation rule of these components under change of the base in V. Suppose to fix another basis $\{e'_j\}_{j\in I} \subset V$ with $e_i = A^j{}_i e'_j$. The coefficients $A^j{}_i$ determine a matrix $A := [A^j{}_i]$ in the matrix group $GL(dimV, \mathbb{K})$, i.e., the group (see the next section) of $dimV \times dimV$ matrices with coefficients in $\mathbb{K}$ and non-vanishing determinant. First consider a contravariant vector $t = t^i e_i$, passing to the other basis, we have $t = t^i e_i = t'^j e'_j$ and thus $t'^j e'_j = t^i A^j{}_i e'_j$. This is equivalent to $(t'^j - A^j{}_i t^i) e'_j = 0$ which implies

$$t'^j = A^j{}_i t^i \ ,$$

because of the linear independence of vectors e'_j. Similarly, if we specify a set of components in $\mathbb{K}$, $\{t^i\}_{i\in I}$ for each basis $\{e_i\}_{i\in I} \subset V$ and these components, changing the basis to $\{e'_j\}_{j\in I}$, transform as

$$t'^j = A^j{}_i t^i \ ,$$

where the coefficients $A^j{}_i$ are defined by

$$e_i = A^j{}_i e'_j \ ,$$

then a contravariant tensor t is defined. It is determined by $t := t^i e_i$ in each basis. The proof is self evident.

Concerning covariant vectors, a similar result holds. Indeed a covariant vector $u \in V^*$ is completely determined by the specification of a set of components $\{u_i\}_{i\in I}$ for each basis $\{e^{*i}\}_{i\in I} \subset V^*$ (dual basis of $\{e_i\}_{i\in I}$ above) when these components, changing the basis to $\{e'^{*j}\}_{j\in I}$ (dual base of $\{e'_j\}_{j\in I}$ above), transform as

$$u'_j = B_j{}^i u_i \ ,$$

where the coefficients ${B_l}^r$ are defined by

$$e^{*i} = {B_j}^i e'^{*j} \ .$$

What is the relation between the matrix $A = [{A^j}_i]$ and the matrix $B := [{B_k}^h]$? The answer is clear: it must be

$$\delta^i_j = \langle e_j, e^{*i} \rangle = {A^l}_j {B_k}^i \langle e'_l, e'^{*k} \rangle = {A^l}_j {B_k}^i \delta^k_l = {A^k}_j {B_k}^i \ .$$

In other words it has to hold $I = A^t B$, which is equivalent to

$$B = A^{-1t} \ .$$

(Notice that ^{t} and $^{-1}$ commute.)
Proper tensors have components which transform similarly to the vector case. For instance, consider $t \in V \otimes V^*$, fix a basis $\{e_i\}_{i \in I} \subset V$, the dual one $\{e^{*i}\}_{i \in I} \subset V^*$ and consider that induced in $V \otimes V^*$, $\{e_j \otimes e^{*j}\}_{(i,j) \in I \times I}$. Then $t = {t^i}_j e_i \otimes e^{*j}$. By bi linearity of the tensor product map, if we pass to consider another basis $\{e'_i\}_{i \in I} \subset V$ and those associated in the relevant spaces as above, concerning the components ${t'^k}_l$ of t in the new tensor space basis, one trivially gets

$${t'^k}_l = {A^k}_i {B_l}^j {t^i}_j \ ,$$

where the matrices $A = [{A^j}_i]$ and $B := [{B_k}^h]$ are those considered above. It is clear that the specification of a tensor of $V \otimes V^*$ is completely equivalent to the specification of a set of components for each basis of $V \otimes V^*$, $\{e_j \otimes e^{*j}\}_{(i,j) \in I \times I}$, provided these components transform as specified above under change of basis.
We can generalize the obtained results after a definition.

Def. 2.2 (Canonical bases.) *Let $\mathcal{A}_{\mathbb{K}}(V)$ be the tensor algebra generated by the finite-dimensional vector space V on the field $\mathbb{K}$ (= $\mathbb{C}$ or $\mathbb{R}$). If $B = \{e_i\}_{i \in I}$ is a basis in V with dual basis $B^* = \{e^{*i}\}_{i \in I} \subset V^*$, the* **canonical bases** *associated to the former are the bases in the tensor spaces of $\mathcal{A}_{\mathbb{K}}(V)$ obtained by tensor products of elements of B and B^*.*

Remark. Notice that also $\{e_i \otimes e'_j\}_{i,j \in I}$ is a basis of $V \otimes V$ if $\{e_i\}_{i \in I}$ and $\{e'_j\}_{j \in I}$ are bases of V. However, $\{e_i \otimes e'_j\}_{i,j \in I}$ is *not* canonical unless $e_i = e'_i$ for all $i \in I$.

Theorem 2.1. *Consider the tensor algebra $\mathcal{A}_{\mathbb{K}}(V)$ generated by a finite-dimension vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and take a tensor space $V^{n\otimes} \otimes V^{*m\otimes} \in \mathcal{A}_{\mathbb{K}}(V)$. The specification of a tensor $t \in V^{n\otimes} \otimes V^{*m\otimes}$ is completely equivalent to the specification of a set of components*

$$\{{t^{i_1 \ldots i_n}}_{j_1 \ldots j_m}\}_{i_1,\ldots,i_n,j_1,\ldots,j_m \in I}$$

with respect to each canonical basis of $V^{n\otimes} \otimes V^{*m\otimes}$,

$$\{e_{i_1} \otimes \ldots \otimes e_{i_n} \otimes e^{*j_1} \otimes \ldots \otimes e^{*j_n}\}_{i_1,\ldots i_n,j_1,\ldots j_m \in I}$$

which, under change of basis:

$$\{e'_{i_1} \otimes \ldots \otimes e'_{i_n} \otimes e'^{*j_1} \otimes \ldots \otimes e'^{*j_n}\}_{i_1,\ldots i_n, j_1,\ldots j_m \in I}$$

transform as:

$$t'^{i_1\ldots i_n}{}_{j_1\ldots j_m} = A^{i_1}{}_{k_1} \ldots A^{i_n}{}_{k_n} B_{j_1}{}^{l_1} \ldots B_{j_m}{}^{l_m} t^{k_1\ldots k_n}{}_{l_1\ldots l_m} \,,$$

where

$$e_i = A^j{}_i e'_j \,,$$

and the coefficients $B_j{}^l$ *are those of the matrix:*

$$B = A^{-1t} \,,$$

with $A := [A^j{}_i]$*. The associated tensor* t *is represented by*

$$t = t^{i_1\ldots i_n}{}_{j_1\ldots j_m} e_{i_1} \otimes \ldots \otimes e_{i_n} \otimes e^{*j_1} \otimes \ldots \otimes e^{*j_n}$$

for each considered canonical basis. Analogous results hold for tensor spaces whose factors V *and* V^* *take different positions.*

Notation 2.1. *In the* **abstract index notation** *a tensor is indicated by writing its generic component in a non-specified basis. E.g.* $t \in V^* \otimes V$ *is indicated by* $t_i{}^j$*.*

Somewhere in the following we adopt a cumulative index notation, i.e., letters $A, B, C, \ldots$ denote set of covariant, contravariant or mixed indices. For instance $t^{ijk}{}_{lm}$ can be written as t^A with ${}^A = {}^{ijk}{}_{lm}$. Similarly e_A denotes the element of a canonical basis $e_i \otimes e_j \otimes e_k \otimes e^{*l} \otimes e^{*m}$. Moreover, if ^{A} and ^{B} are cumulative indices, the indices of the cumulative index ^{AB} are those of ^{A} immediately followed by those of ^{B}, e.g, if ${}^A = {}^{ijk}{}_{lm}$, and ${}^B = {}^{pq}{}_u{}^n$, ${}^{AB} = {}^{ijk}{}_{lm}{}^{pq}{}_u{}^n$.

Let us specify the allowed mathematical rules to produces tensors from given tensors. To this end we shall use both the synthetic and the index notation.

Linear combinations of tensors of a fixed tensor space. Take a tensor space $S \in \mathcal{A}_{\mathbb{K}}(V)$. This is a vector space by definition, and thus picking out $s, t \in S$, and $\alpha, \beta \in \mathbb{K}$, linear combinations can be formed which still belong to S. in other words we may define the tensor of S

$$u := \alpha s + \beta t \quad \text{or, in the abstract index notation} \quad u^A = \alpha s^A + \beta t^A \,.$$

The definition of u above given by the abstract index notation means that the components of u are related to the components of s and t by a linear combination which has *the same form in whatever canonical basis of the space* S and the coefficients α,β *do not depend on the basis.*

Products of tensors of generally different tensor spaces. Take two tensor spaces $S, S' \in \mathcal{A}_{\mathbb{K}}(V)$ and pick out $t \in S$, $t' \in S'$. Then the tensor $t \otimes t' \in S \otimes S'$ is well defined. Using the

associativity of the tensor product by natural isomorphisms, we found a unique tensor space $S'' \in \mathcal{A}_{\mathbb{K}}(V)$ which is isomorphic to $S \otimes S'$ and thus we can identify $t \otimes t'$ with a tensor in S'' which we shall indicate by $t \otimes t'$ once again with a little misuse of notation. $t \otimes t'$ is called the **product** of tensors t and t'. Therefore, the product of tensors coincides with the usual tensor product of tensors up to a natural isomorphism. What about the abstract index notation? We leave to the reader the trivial proof of the following fact based on **Theorem 1.7**,

$$(t \otimes t')^{AB} = t^A t'^B \,.$$

Thus, for instance, if $S = V \otimes V^*$ and $S' = V^*$ the tensors t and t' are respectively indicated by ${t^i}_j$ and s_k and thus $(t \otimes s)^i{}_{jk} = {t^i}_j s_k$.

Contractions. Consider a tensor space of $\mathcal{A}_{\mathbb{K}}(V)$ of the form

$$U_1 \otimes \ldots \otimes U_k \otimes V \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes V^* \otimes U_{l+1} \otimes \ldots \otimes U_n$$

where U_i denotes either V or V^*. Everything we are going to say can be re-stated for the analogous space

$$U_1 \otimes \ldots \otimes U_k \otimes V^* \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes V \otimes U_{l+1} \otimes \ldots \otimes U_n \,.$$

Then consider the *multi-linear* mapping C with domain

$$U_1 \times \ldots \times U_k \times V \times U_{k+1} \times \ldots \times U_l \times V^* \times U_{l+1} \times \ldots \times U_n \,,$$

and values in

$$U_1 \otimes \ldots \otimes U_k \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes U_{l+1} \otimes \ldots \otimes U_n$$

defined by:

$$(u_1, \ldots, u_k, v, u_{k+1}, \ldots, u_l, v^*, v_{l+1}, \ldots, v_n) \mapsto \langle v, v^* \rangle u_1 \otimes \ldots \otimes u_k \otimes u_{k+1} \otimes \ldots \otimes u_l \otimes u_{l+1} \otimes \ldots \otimes u_n .$$

By the universality theorem there is a *linear* mapping $C^{\otimes}$, called **contraction** of V and V^*, defined on the whole tensor space

$$U_1 \otimes \ldots \otimes U_k \otimes V \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes V^* \otimes U_{l+1} \otimes \ldots \otimes U_n$$

taking values in

$$U_1 \otimes \ldots \otimes U_k \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes U_{l+1} \otimes \ldots \otimes U_n$$

such that, on simple products of vectors reduces to

$$u_1 \otimes \ldots \otimes u_k \otimes v \otimes u_{k+1} \otimes \ldots \otimes u_l \otimes v^* \otimes u_{l+1} \otimes \ldots \otimes u_n \mapsto \langle v, v^* \rangle u_1 \otimes \ldots \otimes u_k \otimes u_{k+1} \otimes \ldots \otimes u_l \otimes u_{l+1} \otimes \ldots \otimes u_n .$$

This linear mapping takes tensors in a tensor product space with $n+2$ factors and produces tensors in a space with n factors. The simplest case arises for $n = 0$. In that case $C : V \times V^* \to \mathbb{K}$

is nothing but the bilinear pairing $C : (v, v^*) \mapsto \langle v, v^* \rangle$ and $C^\otimes$ is the linear associated mapping by the universality theorem.
Finally, let us represent the contraction mapping within the abstract index picture. It is quite simple to show that, $C^\otimes$ takes a tensor $t^{AiB}{}_j{}^C$ where ^{A}, ^{B} and ^{C} are arbitrary cumulative indices, and produces the tensor $(C^\otimes t)^{ABC} := t^{AkB}{}_k{}^C$ where we remark the convention of *summation of the double repeated index* k. To show that the abstract-index representation of contractions is that above notice that the contractions are linear and thus

$$C^\otimes(t^{AiB}{}_j{}^C e_A \otimes e_i \otimes e_B \otimes e^{*j} \otimes e_C) = t^{AiB}{}_j{}^C C^\otimes(e_A \otimes e_i \otimes e_B \otimes e^{*j} \otimes e_C) = t^{AiB}{}_j{}^C \delta^j_i e_A \otimes e_B \otimes e_C \,,$$

and thus

$$C^\otimes(t^{AiB}{}_j{}^C e_A \otimes e_i \otimes e_B \otimes e^{*j} \otimes e_C) = t^{AkB}{}_k{}^C e_A \otimes e_B \otimes e_C \,.$$

This is nothing but:

$$(C^\otimes t)^{ABC} := t^{AkB}{}_k{}^C \,.$$

To conclude we pass to consider a final theorem which shows that there is a one-to-one correspondence between linear mappings on tensors and tensors them-selves.

Theorem 2.2 (Linear mappings and tensors.) *Let S, S' be a pair of tensor spaces of a tensor algebra $\mathcal{A}_{\mathbb{K}}(V)$. The vector space of linear mappings from S to S' is naturally isomorphic to $S^* \otimes S'$ (which it is naturally isomorphic to the corresponding tensor space of $\mathcal{A}_{\mathbb{K}}(V)$). The isomorphism $F : \mathcal{L}(S|S') \to S^* \otimes S'$ is given by $f \mapsto t(f)$, where the tensor $t_f \in S^* \otimes S'$ is that (uniquely) determined by the requirement*

$$C^\otimes(s \otimes t_f) = f(s) \,, \quad \textit{for all } s \in S$$

Moreover, fixing a basis $\{e_i\}_{i \in I}$ in V, let $\{e_A\}$ denote the canonical basis induced in S, $\{e'_B\}$ that induced in S' and $\{e'^{*C}\}$ that induced in S'^*. With those definitions

$$t(f)_A{}^C = \langle f(e_A), e'^{*C} \rangle$$

and, if $s = s^A E_A \in S$,

$$f(s)^C = s^A t(f)_A{}^C \,.$$

The isomorphism F is the composition of the isomorphisms in **Theorem 1.3** *and* **Theorem 1.2***: $\mathcal{L}(S|S') \simeq \mathcal{L}(S, S'^*) = S^* \otimes (S'^*)^* \simeq S^* \otimes S'$.*

Skech of proof. Defining t_f in components as said above one has a linear map $F : f \mapsto t_f$ satisfying $C^\otimes(s \otimes t_f) = f(s)$, for all $s \in S$. Since $S^* \otimes S'$ and $\mathcal{L}(S|S') \simeq \mathcal{L}(S, S'^*) \simeq S^* \otimes S'$ have the same dimension, injectivity of F would imply surjectivity. Injectivity can be straightforwardly proved using the definition of t_f itself. Similarly, using the rules to change bases as in **Theorem 2.1**, one finds the independence on the used canonical bases. The last statement can be checked

by direct inspection. □

Let us illustrate how one may use that theorem. For instance consider a linear mapping $f : V \to V$ where V is finite-dimensional with field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. Then f defines a tensor of $V^* \otimes V$. In fact, fixing a basis $\{e_i\}_{i\in I}$in V and considering those canonically associated, by the linearity of f and the pairing:

$$f(v) = \langle f(v), e^{*k}\rangle e_k = v^i \langle f(e_i), e^{*k}\rangle e_k \,.$$

We may define the tensor $t(f) \in V^* \otimes V$ such that

$$t(f) := \langle f(e_i), e^{*k}\rangle e^{*i} \otimes e_k \,.$$

Employing **Theorem 2.1** one can trivially show that the given definition is well posed. We leave to the reader the proof that the obtained maps which associates linear functions with tensors is the isomorphism of the theorem above.
The action of f on v can be represented in terms of $t(f)$ and v using the rules presented above. Indeed, by the abstract index notation, one has

$$(f(v))^k = v^i t(f)_i{}^k \,.$$

In other words the action of f on v reduces to (1) a product of the involved tensors:

$$v^i t(f)_j{}^k \,,$$

(2) followed by a convenient contraction:

$$(f(v))^k = v^i t(f)_i{}^k \,.$$

More complicate cases can be treated similarly. For example, linear mappings f from $V^* \otimes V$ to $V^* \otimes V \otimes V$ are determined by tensors $t(f)^i{}_{jk}{}^{lm}$ of $V \otimes V^* \otimes V^* \otimes V \otimes V$ and their action on tensors $u_p{}^q$ of $V^* \otimes V$ is

$$f(u)_k{}^{lm} = u_i{}^j t(f)^i{}_{jk}{}^{lm}$$

i.e., a product of tensors and two contractions. Obviously

$$t(f)^i{}_{jk}{}^{lm} = \big(f(e^{*i} \otimes e_j)\big)\,(e_k, e^{*l}, e^{*m})$$

where we have used the multi-linear action of elements of $V^* \otimes V \otimes V$ on elements of $V \times V^* \times V^*$.

2.3 Physical invariance of the form of laws and tensors.

A physically relevant result is that the rules given above to produce a new tensor from given tensors have the same form whatever basis one use to handle tensors. For that reason the abstract index notation makes sense. In physics the choice of a basis is associated with the choice of a reference frame. As is well known, various relativity principles (*Galileian Principle*, *Special*

Relativity Principle and *General Relativity Principle*) assume that
"*the law of Physics can be written in such a way that they preserve their form whatever reference frame is used*".
The allowed reference frames range in a class singled out by the considered relativity principle, for instance in Special Relativity the relevant class is that of inertial reference frames.
It is clear that the use of tensors and rules to compose and decompose tensors to represent physical laws is very helpful in implementing relativity principles. In fact the theories of Relativity can be profitably formulated in terms of tensors and operations of tensors just to assure the invariance of the physical laws under change of the reference frame. When physical laws are given in terms of tensorial relations one says that those laws are *covariant*. It is worthwhile stressing that covariance is *not* the only way to state physical laws which preserve their form under changes of reference frames. For instance the Hamiltonian formulation of mechanics, in relativistic contexts, is invariant under change of reference frame but it is not formuled in terms of tensor relations (in spacetime).

3 Some Applications.

In this section we present a few applications of the theory previously developed.

3.1 Tensor products of group representations.

As is known a *group* is an algebraic structure, $(G, \circ)$, where G is a set and $\circ : G \times G \to G$ is a mapping called the *composition rule* of the group. Moreover the following three conditions have to be satisfied.
(1) $\circ$ is *associative*, i.e.,

$$g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3 \ , \quad \text{for all} \quad g_1, g_2, g_3 \in G \ .$$

(2) There is a *group unit*, i.e., there is $e \in G$ such that

$$e \circ g = g \circ e = g \ , \quad \text{for all} \quad g \in G \ .$$

(3) Each element $g \in G$ admits an *inverse element*, i.e.,

$$\text{for each} \quad g \in G \quad \text{there is} \quad g^{-1} \in G \quad \text{with} \quad g \circ g^{-1} = g^{-1} \circ g = e \ .$$

We remind the reader that the unit element turns out to be unique and so does the inverse element for each element of the group (the reader might show those uniqueness properties as an exercise). A group $(G, \circ)$ is said to be *commutative* or *Abelian* if $g \circ g' = g' \circ g$ for each pair of elements, $g, g' \in G$; otherwise it is said to be *non-commutative* or *non-Abelian*. A subset $G' \subset G$ of a group is called *subgroup* if it is a group with respect to the restriction to $G' \times G'$ of the composition rule of G.

If $(G_1, \circ_1)$ and $(G_2, \circ_2)$ are groups, a (group) *homomorphism* from G_1 to G_2 is a mapping $h : G_1 \to G_2$ which *preserves the group structure*, i.e., the following requirement has to be fulfilled:

$$h(g \circ_1 g') = h(g) \circ_2 h(g') \quad \text{for all} \quad g, g' \in G_1 \ ,$$

As a consequence, they also hold with obvious notations:

$$h(e_1) = e_2 \ ,$$

and

$$h(g^{-1_1}) = (h(g))^{-1_2} \quad \text{for each} \quad g \in G_1 \ .$$

Indeed, if $g \in G_1$ one has $h(g) \circ e_2 = h(g) = h(g \circ e_1) = h(g) \circ h(e_1)$. Applying $h(g)^{-1}$ on the left, one finds $e_2 = h(e_1)$. On the other hand $h(g)^{-1} \circ h(g) = e_2 = h(e_1) = h(g^{-1} \circ g) = h(g^{-1}) \circ h(g)$ implies $h(g)^{-1} = h(g^{-1})$.
A group *isomorphism* is a *bijective* group homomorphism.
A simple example of group is $GL(n, \mathbb{K})$ which is the set of the $n \times n$ matrices A with components in the field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and $detA \neq 0$. The group composition rule is the usual product

of matrices. If $n > 0$ the group is non-commutative. An important subgroup of $GL(n, \mathbb{K})$ is $SL(n, \mathbb{K})$, i.e., the set of matrices B in $GL(n, \mathbb{K})$ with $detB = 1$. (The reader might show that $SL(n, \mathbb{K})$ is a subgroup of $GL(n, \mathbb{K})$.) However there are groups which are not defined as group of matrices, e.g., $(\mathbb{Z}, +)$. An non completely trivial example is given by the *group of permutations of n elements* which we shall consider in the next subsection.

Exercises 3.1.
3.1.1. Prove the uniqueness of the unit element and the inverse element in any group.
3.1.2. Show that in any group G the unique element e such that $e^2 = e$ is the unit element.
3.1.3. Show that if G' is a subgroup of G the unit element of G' must coincide with the unit element of G, and, if $g \in G'$, the inverse element g^{-1} in G' coincides with the inverse element in G.
3.1.4. Show that if $h : G_1 \to G_2$ is a group homomorphism, then $h(e_1) = e_2$ and $h(g^{-1_1}) = (h(g))^{-1_2}$ for each $g \in G_1$.
(Hint. Use **3.1.2.** to achieve the former statement.)
3.1.5. Show that if $h : G_1 \to G_2$ is a group homomorphism, then $h(G_1)$ is a subgroup of G_2.

We are interested in the concept of (linear) *representation of a group on a vector space.* In order to state the corresponding definition, notice that, if V is a (not necessarily finite-dimensional) vector space, $\mathcal{L}(V|V)$ contains an important group. This is $GL(V)$ which is the set of both injective and surjective elements of $\mathcal{L}(V|V)$ equipped with the usual composition rule of maps. We can give the following definition.

Def.3.1. (Linear group on a vector space.) *If V is a vector space, $GL(V)$ denotes the group of linear mappings $f : V \to V$ such that f is injective and surjective, with group composition rule given by the usual mappings composition. $GL(V)$ is called the* **linear group on** V.

Remarks.
(a) If $V := \mathbb{K}^n$ then $GL(V) = GL(n, \mathbb{K})$.
(b) If $V \neq V'$ it is not possible to define the analogue of $GL(V)$ considering some subset of $\mathcal{L}(V|V')$. (The reader should explain the reason.)

Def.3.2. (Linear group representation on a vector space.) *Let $(G, \circ)$ be a group and V a vector space. A (linear group)* **representation of** G **on** V *is a homomorphism $\rho : G \to GL(V)$. Moreover a representation $\rho : G \to GL(V)$ is called:*
(1) faithful *if is injective,*
(2) free *if, for any $v \in V$, the subgroup of G made of the elements h_v such that $\rho(h_v)v = v$ contains only the unit element of G,*
(3) transitive *if for each pair $v, v' \in V$ there is $g \in G$ with $v' = \rho(g)v$.*
(4) irreducible *if there is no* proper *vector subspace $S \subset V$ which is* **invariant** *under the action of $\rho(G)$, i.e., which satisfies $\rho(g)S \subset S$ for all $g \in G$.*

Equipped with the above-given definitions we are able to study the simplest interplay of tensors and group representations. We want to show that the notion of tensor product allows the definitions of *tensor products of representations.* That mathematical object is of fundamental importance in applications to Quantum Mechanics, in particular as far as systems with many components are concerned.
Consider a group G (from now on we omit to specify the symbol of the composition rule whenever it does not produces misunderstandings) and several representations of the group $\rho_i : G \to GL(V_i)$, where $V_1, \ldots, V_n$ are finite-dimensional vector spaces on the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. For each $g \in G$, we may define a multi-linear mapping $[\rho_i(g), \ldots, \rho_n(g)] \in \mathcal{L}(V_1, \ldots, V_n | V_1 \otimes \ldots \otimes V_n)$ given by, for all $(\rho_i(g), \ldots, \rho_n(g)) \in V_1 \times \ldots \times V_n$,

$$[\rho_i(g), \ldots, \rho_n(g)] : (v_1, \ldots, v_n) \mapsto (\rho_1(g)v_1) \otimes \ldots \otimes (\rho_n(g)v_n) \,.$$

That mapping is multi linear because of the multi linearity of the tensor-product mapping and the linearity of the operators $\rho_k(g)$. Using the universality theorem, we *uniquely* find a linear mapping which we indicate by $\rho_1(g) \otimes \ldots \otimes \rho_n(g) : V_1 \otimes \ldots \otimes V_n \to V_1 \otimes \ldots \otimes V_n$ such that:

$$\rho_1(g) \otimes \ldots \otimes \rho_n(g)(v_1 \otimes \ldots \otimes v_n) = (\rho_1(g)v_1) \otimes \ldots \otimes (\rho_n(g)v_n) \,.$$

Def.3.3. (Tensor product of representations.) *Let $V_1, \ldots V_n$ be finite-dimensional vector spaces on the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and suppose there are n representations $\rho_i : G \to GL(V_k)$ of the same group G on the given vector spaces. The set of linear maps*

$$\{\rho_1(g) \otimes \ldots \otimes \rho_n(g) : V_1 \otimes \ldots \otimes V_n \to V_1 \otimes \ldots \otimes V_n \mid g \in G\} \,,$$

defined above is called **tensor product of representations** $\rho_1, \ldots, \rho_n$.

The relevance of the definition above is evident because of the following theorem,

Theorem 3.1. *Referring to the definition above, the elements of a tensor product of representations $\rho_1, \ldots \rho_n$ of the group G on spaces $V_1, \ldots, V_n$ define a* **linear group representation of G on the tensor-product space** $V_1 \otimes \ldots \otimes V_n$.

Proof. We have to show that the mapping

$$g \mapsto \rho_1(g) \otimes \ldots \otimes \rho_n(g) \,,$$

is a group homomorphism from G to $GL(V_1 \otimes \ldots \otimes V_n)$. Taking account of the fact that each ρ_i is a group homomorphism, if $g, g' \in G$, one has

$$\rho_1(g') \otimes \ldots \otimes \rho_n(g')(\rho_1(g) \otimes \ldots \otimes \rho_n(g)(v_1 \otimes \ldots \otimes v_n)) = \rho_1(g') \otimes \ldots \otimes \rho_n(g')((\rho_1(g)v_1) \otimes \ldots \otimes (\rho_n(g)v_n))$$

and this is

$$(\rho_1(g' \circ g)v_1) \otimes \ldots \otimes (\rho_n(g' \circ g)v_n) \,.$$

The obtained result holds true also using a canonical basis for $V_1 \otimes \ldots \otimes V_n$ made of usual elements $e_{1,i_1} \otimes \ldots e_{n,i_n}$ in place of $v_1 \otimes \ldots \otimes v_n$. By linearity, it means that

$$(\rho_1(g') \otimes \ldots \otimes \rho_n(g'))(\rho_1(g) \otimes \ldots \otimes \rho_n(g)) = \rho_1(g' \circ g) \otimes \ldots \otimes \rho_n(g' \circ g) \, .$$

To conclude notice that $\rho_1(g) \otimes \ldots \otimes \rho_n(g) \in GL(V_1 \otimes \ldots \otimes V_n)$ because $\rho_1(g) \otimes \ldots \otimes \rho_n(g)$ is (1) linear and furthermore it is (2) bijective. The latter can be proved as follows: $\rho_1(g^{-1}) \otimes \ldots \otimes \rho_n(g^{-1}) \circ (\rho_1(g) \otimes \ldots \otimes \rho_n(g)) = (\rho_1(g) \otimes \ldots \otimes \rho_n(g)) \circ \rho_1(g^{-1}) \otimes \ldots \otimes \rho_n(g^{-1}) = \rho_1(e) \otimes \ldots \otimes \rho_n(e) = I$. The last identity follows by linearity form $(\rho_1(e) \otimes \ldots \otimes \rho_n(e))(v_1 \otimes \cdots \otimes v_n) = v_1 \otimes \cdots \otimes v_n$. $\square$

More generally, if $A_k : V_k \to U_k$ are n linear mappings (operators), and all involved vector spaces are finite dimensional and with the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, it is defined the *tensor product of operators.*

Def.3.4 (Tensor Product of Operators.) *If $A_k : V_k \to U_k$, $k = 1, \ldots, n$ are n linear mappings (operators), and all the vector spaces U_i, V_j are finite dimensional with the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, the* **tensor product of** *$A_1, \ldots, A_n$ is the linear mapping*

$$A_1 \otimes \ldots \otimes A_n : V_1 \otimes \ldots \otimes V_n \to U_1 \otimes \ldots \otimes U_n$$

uniquely determined by the universality theorem and the requirement:

$$(A_1 \otimes \ldots \otimes A_n) \circ \otimes = A_1 \times \ldots \times A_n \, ,$$

where

$$A_1 \times \ldots \times A_n : V_1 \times \ldots \times V_n \to U_1 \otimes \ldots \otimes U_n \, ,$$

is the multi-linear mapping such that, for all $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$:

$$A_1 \times \ldots \times A_n : (v_1, \ldots, v_n) \mapsto (A_1 v_1) \otimes \ldots \otimes (A_n v_n) \, .$$

Remark. Employing the given definition and the same proof used for the relevant part of the proof of **Theorem 3.1**, it is simply shown that if $A_i : V_i \to U_i$ and $B_i : U_i \to W_i$, $i = 1, \ldots, n$ are $2n$ linear maps and all involved spaces V_i, U_j, W_k are finite dimensional with the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, then

$$B_1 \otimes \ldots \otimes B_n \;\; A_1 \otimes \ldots \otimes A_n = (B_1 A_1) \otimes \ldots \otimes (B_n A_n) \, .$$

3.2 A quantum physical example.

Physicsts are involved with Hilbert spaces whenever they handle quantum mechanics. A Hilbert space is noting but a complex vector space equipped with a Hermitean scalar product (see the next chapter) such that it is complete with respect to the norm topology induced by that scalar product. As far as we are concerned we need only the structure of vector space. Physically

speaking, the vectors of the Hilbert space represent the states of the considered physical system (actually things are more complicated but we do not matter). To consider the simplest case we assume that the vector space which describes the states of the system is finite-dimensional (that is the case for the spin part of a quantum particle). Moreover, physics implies that the space $\mathcal{H}$ of the states of a composit system S made of two systems S_1 and S_2 associated with Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, is the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. There should be several remarks in the infinite dimensional case since our definition of tensor product works for finite-dimensional spaces only, however suitable and well-behaved generalizations are, in fact, possible. Let the system S_1 be described by a state $\psi \in \mathcal{H}_1$ and suppose to transform the system by the action of an element R of some physical group of transformations $\mathcal{G}$ (e.g. $SO(3)$). The transformed state ψ' is given by $U_R^{(1)}\psi$ where $\mathcal{G} \ni R \mapsto U_R^{(1)}$ is a representation of $\mathcal{G}$ in terms of linear transformations $U_R^{(1)} : \mathcal{H}_1 \to \mathcal{H}_1$. Actually, physics and the celebrated *Wigner's theorem* in particular, requires that every $U_R^{(1)}$ be a *unitary* (or *anti unitary*) transformation but this is not relevant for our case. The natural question concerning the representation of the action of $\mathcal{G}$ on the composit system S is:
"If we know the representations $\mathcal{G} \ni R \mapsto U_R^{(1)}$ and $\mathcal{G} \ni R \mapsto U_R^{(2)}$, what about the representation of the action of $\mathcal{G}$ on S in terms of linear transformations in the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$?"
The answer given by physics, at least when the systems S_1 and S_2 do not interact, is that $U_g := U_g^{(2)} \otimes U_g^{(1)}$.

3.3 Permutation group and symmetry of tensors.

We remind the definition of the *group of permutations of n objects* and give some know results of basic group theory whose proofs may be found in any group-theory textbook.

Def.3.5. (Group of permutations.) *Consider the set $I_n := \{1, \ldots, n\}$, the* **group of permutations of n objects**, *$\mathcal{P}_n$ is the set of the bijective mappings $\sigma : I_n \to I_n$ equipped with the composition rule given by the usual composition rule of functions. Moreover,*
(a) *the elements of $\mathcal{P}_n$ are called* **permutations** *(of n objects);*
(b) *a permutation of $\mathcal{P}_n$ with $n \geq 2$ is said to be a* **transposition** *if differs from the identity mapping and reduces to the identity mapping when restricted to some subset of I_n containing $n-2$ elements.*

Comments.
(1) $\mathcal{P}_n$ contains $n!$ elements.
(2) Each permutation $\sigma \in \mathcal{P}_n$ can be represented by a corresponding string $(\sigma(1), \ldots, \sigma(n))$.
(3) If, for instance $n = 5$, with the notation above $(1,2,3,5,4)$, $(5,2,3,4,1)$, $(1,2,4,3,5)$ are transpositions, $(2,3,4,5,1)$, $(5,4,3,2,1)$ are not.
(3) It is possible to show that each permutation $\sigma \in \mathcal{P}_n$ can be decomposed as a product of transpositions $\sigma = \tau_1 \circ \ldots \circ \tau_k$. In general there are several different transposition-product decompositions for each permutation, however it is possible to show that if $\sigma = \tau_1 \circ \ldots \circ \tau_k =$

$\tau'_1 \circ \ldots \circ \tau'_r$, where τ_i and τ'_j are transpositions, then $r+k$ is even. Equivalently, r is even or odd if and only if k is so. This defines the **parity**, $\epsilon_\sigma \in \{-1,+1\}$, of a permutation σ, where, $\epsilon_\sigma = +1$ if σ can be decomposed as a product of an *even* number of transpositions and $\epsilon_\sigma = -1$ if σ can be decomposed as a product of an *odd* number of transpositions.
(4) If $A = [A_{ij}]$ is a real or complex $n \times n$ matrix, it is possible to show (by induction) that:

$$detA = \sum_{\sigma \in \mathcal{P}_n} \epsilon_\sigma A_{1\sigma(1)} \cdots A_{n\sigma(n)} \, .$$

Alternatively, the identity above may be used to define the determinant of a matrix.

We pass to consider the action of $\mathcal{P}_n$ on tensors. Fix $n > 1$, consider the tensor algebra $\mathcal{A}_{\mathbb{K}}(K)$ and single out the tensor space $V^{n\otimes} := V \otimes \ldots \otimes V$ where the factor V appears n times Then consider the following action of $\mathcal{P}_n$ on $V^{n\times} := V \times \ldots \times V$ where the factor V appears n times. For each $\sigma \in \mathcal{P}_n$ consider the mapping

$$\hat{\sigma} : V^{n\times} \to V^{n\otimes} \quad \text{such that} \quad \hat{\sigma}(v_1, \ldots, v_n) \mapsto v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(n)} \, .$$

It is quite straightforward to show that $\hat{\sigma}$ is multi linear. Therefore, let

$$\sigma^\otimes : V^{n\otimes} \to V^{n\otimes} \, ,$$

be the linear mapping uniquely determined by $\hat{\sigma}$ by means of the universality theorem. By definition, it is completely determined by linearity and the requirement

$$\sigma^\otimes : v_1 \otimes \ldots \otimes v_n \mapsto v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(n)} \, .$$

Theorem 3.2. *The above-defined linear mapping*

$$\sigma \mapsto \sigma^\otimes$$

with $\sigma \in \mathcal{P}_n$, is a group representation of $\mathcal{P}_n$ on $V^{n\otimes}$.

Proof. First we show that if $\sigma, \sigma' \in \mathcal{P}_n$ then

$$\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = (\sigma \circ \sigma')^\otimes(v_1 \otimes \ldots \otimes v_n) \, ,$$

This follows from the definition: $\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = \sigma^\otimes(v_{\sigma'^{-1}(1)} \otimes \ldots \otimes v_{\sigma'^{-1}(n)})$. Re-defining $u_i := v_{\sigma'^{-1}(i)}$ so that $u_{\sigma^{-1}(j)} := v_{\sigma'^{-1}(\sigma^{-1}(j))}$, one finds the identity $\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = u_{\sigma^{-1}(1)} \otimes \ldots \otimes u_{\sigma^{-1}(n)} = v_{\sigma'^{-1} \circ \sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma'^{-1} \circ \sigma^{-1}(n)} = v_{(\sigma \circ \sigma')^{-1}(1)} \otimes \ldots \otimes v_{(\sigma \circ \sigma')^{-1}(n)} = (\sigma \circ \sigma')^\otimes(v_1 \otimes \ldots \otimes v_n)$. In other words

$$\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = (\sigma \circ \sigma')^\otimes(v_1 \otimes \ldots \otimes v_n) \, .$$

In particular, that identity holds alsofor a canonical basis of elements $e_{i_1} \otimes \ldots \otimes e_{i_n}$

$$\sigma^\otimes(\sigma'^\otimes(e_{i_1} \otimes \ldots \otimes e_{i_n})) = (\sigma \circ \sigma')^\otimes(e_{i_1} \otimes \ldots \otimes e_{i_n}) \, .$$

By linearity such an identity will hold true for all arguments in $V^{n\otimes}$ and thus we have

$$\sigma^{\otimes}\sigma'^{\otimes} = (\sigma \circ \sigma')^{\otimes} \, .$$

The mappings $\sigma^{\otimes}$ are linear by constructions and are bijective because

$$\sigma^{\otimes}\sigma'^{\otimes} = (\sigma \circ \sigma')^{\otimes}$$

implies

$$\sigma^{\otimes}\sigma^{-1\otimes} = \sigma^{-1\otimes}\sigma^{\otimes} = e^{\otimes} = I \, .$$

The identity $e^{\otimes} = I$ can be proven by noticing that $e^{\otimes} - I$ is linear and vanishes when evaluated on any canonical base of $V^{n\otimes}$. We have shown that $\sigma^{\otimes} \in GL(V^{n\otimes})$ for all $\sigma \in \mathcal{P}_n$ and the mapping $\sigma \mapsto \sigma^{\otimes}$ is a homomorphism. This concludes the proof. □

Let us pass to consider the abstract index notation and give a representation of the action of $\sigma^{\otimes}$ within that picture.

Theorem 3.3. *If t is a tensor in $V^{n\otimes} \in \mathcal{A}_{\mathbb{K}}(V)$ with $n \geq 2$ and $\sigma \in \mathcal{P}_n$, then the components of t with respect to any canonical basis of $V^{n\otimes}$ satisfy*

$$(\sigma^{\otimes}t)^{i_1 \ldots i_n} = t^{i_{\sigma(1)} \ldots i_{\sigma(n)}} \, .$$

Proof.

$$\sigma^{\otimes}t = \sigma^{\otimes}(t^{j_1 \ldots j_n} e_{j_1} \otimes \ldots \otimes e_{j_n}) = t^{j_1 \ldots j_n} e_{j_{\sigma^{-1}(1)}} \otimes \ldots \otimes e_{j_{\sigma^{-1}(n)}} \, .$$

Since $\sigma : I_n \to I_n$ is bijective, if we define $i_k := j_{\sigma^{-1}(k)}$, it holds $j_k = i_{\sigma(k)}$. Using this identity above we find

$$\sigma^{\otimes}t = t^{i_{\sigma(1)} \ldots i_{\sigma(n)}} e_{i_1} \otimes \ldots \otimes e_{i_n} \, .$$

That is nothing but the thesis. □

To conclude we introduce the concept of symmetric or anti-symmetric tensor.

Def. 3.6 (Symmetric and anti-symmetric tensors.) *Let V be a finite-dimensional vector space with field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. Consider the space $V^{n\otimes} \in \mathcal{A}_{\mathbb{K}}(V)$ of tensors of order $(n, 0)$ with $n \geq 2$.*
(a) *$t \in V^{n\otimes}$ is said to be* **symmetric** *if*

$$\sigma^{\otimes}t = t \, ,$$

for all of $\sigma \in \mathcal{P}_n$, or equivalently, using the abstract index notation,

$$t^{j_1 \ldots j_n} = t^{j_{\sigma(1)} \ldots j_{\sigma(n)}} \, ,$$

for all of $\sigma \in \mathcal{P}_n$.
(b) $t \in V^{n\otimes}$ *is said to be* **anti symmetric** *if*

$$\sigma^{\otimes} t = \epsilon_\sigma \, t \, ,$$

for all of $\sigma \in \mathcal{P}_n$*, or equivalently using the abstract index notation:*

$$t^{j_1 \ldots j_n} = \epsilon_\sigma \, t^{j_{\sigma(1)} \ldots j_{\sigma(n)}} \, ,$$

for all of $\sigma \in \mathcal{P}_n$.

Remark. Concerning the definitition in **(b)** notice that $\epsilon_\sigma = \epsilon_{\sigma^{-1}}$.

Examples 3.1.
3.1.1. Suppose $n = 2$, then a symmetric tensor $s \in V \otimes V$ satisfies $s^{ij} = s^{ji}$ and an antisymmetric tensor $a \in V \otimes V$ satisfies $a^{ij} = -a^{ji}$.
3.1.2. Suppose $n = 3$, then it is trivially shown that $\sigma \in \mathcal{P}_3$ has parity 1 if and only if σ is a **cyclic permutation**, i.e., $(\sigma(1), \sigma(2), \sigma(3)) = (1, 2, 3)$ or $(\sigma(1), \sigma(2), \sigma(3)) = (2, 3, 1)$ or $(\sigma(1), \sigma(2), \sigma(3)) = (3, 1, 2)$.
Now consider the vector space V with $dim V = 3$. It turns out that a tensor $e \in V \otimes V \otimes V$ is anti symmetric if and only if

$$e^{ijk} = 0 \, ,$$

if (i, j, k) is *not* a permutation of $(1, 2, 3)$ and, otherwise,

$$e^{ijk} = \pm e^{123} \, ,$$

where the sign $+$ takes place if the permutation $(\sigma(1), \sigma(2), \sigma(3)) = (i, j, k)$ is cyclic and the sign $-$ takes place otherwise. That relation between parity of a permutation and cyclicity does *not* hold true for $n > 3$.

Remarks.
(1) Consider a generic tensor space $S \in \mathcal{A}_{\mathbb{K}}(V)$ which contains $n \geq 2$ spaces V as factors. We may suppose for sake of simplicity $S = S_1 \otimes V^{n\otimes} \otimes S_2$ where $S_1 = U_1 \otimes \ldots \otimes U_{n-1}$, $S_2 = U_{n+1} \otimes \ldots \otimes U_m$ and $U_i = V$ or $U_i = V^*$. Anyway all what we are going to say holds true also if the considered n spaces V do not define a unique block $V^{n\otimes}$. We may define the action of $\sigma \in \mathcal{P}_n$ on the whole space S starting by a multi linear mapping

$$\hat{\sigma} : U_1 \times \ldots \times U_k \times V^{n\times} \times U_{k+1} \times \ldots \times U_m \to U_1 \otimes \ldots \otimes U_k \otimes V^{n\otimes} \otimes U_{k+1} \otimes \ldots \otimes U_m \, ,$$

such that reduces to the tensor-product mapping on $U_1 \times \ldots \times U_k$ and $U_{k+1} \times \ldots \times U_m$:

$$\sigma : (u_1, \ldots, u_k, v_1, \ldots v_n, u_k, \ldots, u_m) \mapsto u_1 \otimes \ldots \otimes u_k \otimes v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)} \otimes u_{k+1} \otimes \ldots \otimes u_m \, .$$

Using the universality theorem as above, we build up a representation of $\mathcal{P}_n$ on S, $\sigma \mapsto \sigma^{\otimes}$ which "acts on $V^{n\otimes}$ only". Using the abstract index notation the action of $\sigma^{\otimes}$ is well represented:

$$\sigma^{\otimes} : t^{A\ i_1 \dots i_n\ B} \mapsto t^{A\ i_{\sigma(1)} \dots i_{\sigma(n)}\ B}\ .$$

This allows one to define and study the symmetry of a tensor referring to a few indices singled out among the complete set of indices of the tensors. E.g., a tensor $t^{ij}{}_k{}^r$ may be symmetric or anti symmetric, for instance, with respect to the indices i and j or j and r or ijr.
(2) Everything we have obtained and defined may be similarly re-obtained and re-defined considering spaces of tensors of order $(0, n)$ with $n \geq 2$, i.e. covariant tensors. In that case the *antisymmetric tensors* of order $(0, n)$ are called n**-forms**.
(3) Notice that no discussion on the symmetry of indices of different kind (one covariant and the other contravariant) is possible.

Exercises 3.2.
3.2.1. Let t be a tensor in $V^{n\otimes}$ (or $V^{*n\otimes}$). Show that t is symmetric or anti symmetric if there is a canonical basis where the components have symmetric or anti symmetric indices, i.e., $t^{i_1 \dots i_n} = t^{i_{\sigma(1)} \dots i_{\sigma(n)}}$ or respectively $t^{i_1 \dots i_n} = \epsilon_\sigma\, t^{i_{\sigma(1)} \dots i_{\sigma(n)}}$ for all $\sigma \in \mathcal{P}_n$.
Note. The result implies that, to show that a tensor is symmetric or anti symmetric, it is sufficient to verify the simmetry or anti symmetry of its components within a *single* canonical basis.
3.2.2. Show that the sets of symmetric tensors of order $(n, 0)$ and $(0, n)$ are vector subspaces of $V^{n\otimes}$ and $V^{*n\otimes}$ respectively.
3.2.3. Show that the subspace of anti-symmetric tensors of order $(0, n)$ (the space of n-forms) in $V^{*n\otimes}$ has dimension $\binom{dimV}{n}$ if $n \leq dimV$. What about $n > dimV$?
3.2.4. Consider a tensor $t^{i_1 \dots i_n}$, show that the tensor is symmetric if and only if it is symmetric with respect to each arbitrary chosen pair of indices, i.e.

$$t^{\dots i_k \dots i_p \dots} = t^{\dots i_p \dots i_k \dots}\ ,$$

for all $p, k \in \{1, \dots n\}$, $p \neq k$.
3.2.5. Consider a tensor $t^{i_1 \dots i_n}$, show that the tensor is anti symmetric if and only if it is anti symmetric with respect to each arbitrarily chosen pair of indices, i.e.

$$t^{\dots i_k \dots i_p \dots} = -t^{\dots i_p \dots i_k \dots}\ ,$$

for all $p, k \in \{1, \dots n\}$, $p \neq k$.
3.2.6. Show that $V \otimes V = A \oplus S$ where $\oplus$ denotes the direct sum and A and S are respectively the space of anti-symmetric and symmetric tensors in $V \otimes V$. Does such a direct decomposition hold if considering $V^{n\otimes}$ with $n > 2$?

4 Scalar Products and Metric Tools.

This section concerns the introduction of the notion of scalar product and several applications on tensors.

4.1 Scalar products.

First of all we give the definition of a *pseudo scalar product* and *semi scalar products* which differ from the notion of *scalar product* for the positivity and the non-degenerateness requirement respectively. In fact, a pseudo scalar product is a generalization of the usual definition of scalar product which has many applications in mathematical physics, relativistic theories in particular. Semi scalar products are used in several applications of quantum field theory (for instance in the celebrated *GNS theorem*).

Def.4.1. (Pseudo Scalar Product.) *Let V be a vector space on the field either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *A* **pseudo scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{K}$ which is:*
(i) **bi linear**, *i.e., for all $u \in V$ both $(u|\cdot) : v \mapsto (u|v)$ and $(\cdot|u) : v \mapsto (v|u)$ are linear functionals on V;*
(ii) **symmetric**, *i.e., $(u|v) = (v|u)$ for all $u, v \in V$;*
(iii) **non-degenerate**, *i.e., $(u|v) = 0$ for all $v \in V$ implies $u = 0$.*
(b) *If $\mathbb{K} = \mathbb{C}$, a* **Hermitian pseudo scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{K}$ which is:*
(i) **sesquilinear**, *i.e., for all $u \in V$, $(u|\cdot)$ and $(\cdot|u)$ are a linear functional and an anti-linear functional on V respectively;*
(ii) **Hermitian**, *i.e., $(u|v) = \overline{(v|u)}$ for all $u, v \in V$;*
(iii) **non-degenerate**.

Def.4.2 (Semi Scalar Product.) *Let V be a vector space on the field either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *If $\mathbb{K} = \mathbb{R}$, a* **semi scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{R}$ which satisfies* (ai),(aii) *above and is*
(iv) **semi-defined positive**, *i.e., $(u|u) \geq 0$ for all $u \in V$.*
(b) *If $\mathbb{K} = \mathbb{C}$, a* **Hermitian semi scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{K}$ which satisfies* (bi),(bii) *above and is*
(iv) **semi-defined positive.**

Finally we give the definition of scalar product.

Def.4.3. (Scalar Product.) *Let V be a vector space on the field $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{C}$) endowed with a pseudo scalar product (resp. Hermitian pseudo scalar product) $(\,|\,)$. $(\,|\,)$ is called* **scalar product** *(***resp. Hermitian scalar product***) if $(\,|\,)$ is also a semi scalar product, i.e., if it is* **semi-defined positive**.

Comments.
(1) Notice that all given definitions do not require that V is finite dimensional.
(2) If $\mathbb{K} = \mathbb{C}$ and $(|)$ is not Hermitian, in general, any requirement on positivity of $(u|u)$ does not make sense because $(u|u)$ may *not* be real. If instead Hermiticity holds, we have $(u|u) = \overline{(u|u)}$ which assures that $(u|u) \in \mathbb{R}$ and thus positivity may be investigated.
(3) Actually, a (Hermitian) scalar product is *positive defined*, i.e,

$$(u|u) > 0 \quad \text{if} \quad u \in V \setminus \{0\} ,$$

because of Cauchy-Schwarz' inequality

$$|(u|v)|^2 \leq (u|u)(v|v) ,$$

which we shall prove below for semi scalar products.
(4) A **semi norm** on a vector space V with field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, is a mapping $|| \ || : V \to \mathbb{K}$ such that:
(i) $||v|| \in \mathbb{R}$ and in particular $||v|| \geq 0$ for all $v \in V$;
(ii) $||\alpha v|| = |\alpha| ||v||$ for all $\alpha \in \mathbb{K}$ and $v \in V$;
(iii) $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in V$.
A semi norm $|| \ || : V \to \mathbb{K}$ is a **norm** if
(iv) $||v|| = 0$ implies $v = 0$.
Notice that for semi norms it holds: $||0|| = 0$ because of (ii) above. With the given definitions, it quite simple to show (the reader might try to give a proof) that if V with field $\mathbb{K} = \mathbb{R}$ $(\mathbb{C})$ is equipped by a (Hermitian) semi scalar product then $||v|| := \sqrt{(v|v)}$ for all $v \in V$ defines a semi norm. Furthermore if $(\, | \,)$ is a scalar product, then the associated semi norm is a norm.
(5) If a vector space V is equipped with a norm $|| \ ||$ it becomes a *metric space* by defining the *distance* $d(u, v) := ||u - v||$ for all $u, v \in V$. A **Banach space** $(V, || \ ||)$ is a vector space equipped with a norm such that the associates metric space is *complete*, i.e, all Cauchy's sequences converge. A **Hilbert space** is a Banach space with norm given by a (Hermitean if the field is $\mathbb{C}$) scalar product as said above. Hilbert spaces are the central mathematical objects used in quantum mechanics.

Exercices 4.1.
4.1.1. Show that if $(\, | \,)$ is a (Hermitian) semi scalar product on V with field $\mathbb{R}$ $(\mathbb{C})$ then the mapping on V, $v \mapsto ||v|| := \sqrt{(v|v)}$, satisfies $||u + v|| \leq ||u|| + ||v||$ as a consequence of Cauchy-Schwarz' inequality

$$|(u|v)|^2 \leq (u|u)(v|v) ,$$

which holds true by all (Hermitian) semi scalar product.
(Hint. Compute $||u + v||^2 = (u + v|u + v)$ using bi linearity or sesquilinearity property of $(\, | \,)$, then use Cauchy-Schwarz' inequality.)

Theorem 4.1. (Cauchy-Schwarz' inequality.) *Let V be a vector space with field $\mathbb{R}$ ($\mathbb{C}$) equipped with a (Hermitian) semi scalar product $(\,|\,)$. Then, for all $u, v \in V$, Cauchy-Schwarz' inequality holds:*

$$|(u|v)|^2 \leq (u|u)(v|v)\,.$$

Proof. Consider the complex case with a Hermitian semi scalar product. Take $u, v \in V$. For all $z \in \mathbb{C}$ it must hold $(zu + v|zu + v) \geq 0$ by definition of Hermitean semi scalar product. Using sesquilinearity and Hermiticity :

$$0 \leq \overline{z}z(u|u) + (v|v) + \overline{z}(u|v) + z(v|u) = |z|^2(u|u) + (v|v) + \overline{z}(u|v) + z\overline{(u|v)}\,,$$

which can be re-written as

$$|z|^2(u|u) + (v|v) + 2Re\{\overline{z}(u|v)\} \geq 0\,.$$

Then we pass to the polar representation of z, $z = re^{i\,\alpha}$ with $r, \alpha \in \mathbb{R}$ arbitrarily and independently fixed. Decompose also $(u|v)$, $(u|v) = |(u|v)|e^{i\ arg(u|v)}$. Inserting above we get:

$$F(r, \alpha) := r^2(u|u) + 2r|(u|v)|Re[e^{i(arg(u|v)\,-\,\alpha)}] + (v|v) \geq 0\,,$$

for all $r \in \mathbb{R}$ when $\alpha \in R$ is fixed arbitrarily. Since the right-hand side above is a second-order polynomial in r, the inequlity implies that, for all $\alpha \in \mathbb{R}$,

$$\left\{2|(u|v)|Re[e^{i(arg(u|v)\,-\,\alpha)}]\right\}^2 - 4(v|v)(u|u) \leq 0\,,$$

which is equivalent to

$$|(u|v)|^2 \cos(arg(u|v)\ -\ \alpha) - (u|u)(v|v) \leq 0\,,$$

for all $\alpha \in \mathbb{R}$. Choosing $\alpha = arg(u|v)$, we get Cauchy-Schwarz' inequality:

$$|(u|v)|^2 \leq (u|u)(v|v)\,.$$

The real case can be treated similarly, replacing $z \in \mathbb{C}$ with $x \in \mathbb{R}$ and the proof is essentially the same. $\square$.

Corollary. *A bilinear symmetric (resp. sesquilinear Hermitian) mapping $(\,|\,) : V \times V \to \mathbb{R}$ (resp. $\mathbb{C}$) is a scalar product (resp. Hermitian scalar product) if and only if it is* **positive defined***, that is $(u|u) > 0$ for all $u \in V \setminus \{\mathbf{0}\}$.*

Proof. Assume that $(\,|\,)$ is a (Hermitian) scalar product. Hence $(u|u) \geq 0$ by definition and $(\,|\,)$ is non-degenerate. Moreover it holds $|(u|v)|^2 \leq (u|u)(v|v)$. As a consequence, if $(u|u) = 0$ then $(u|v) = 0$ for all $v \in V$ and thus $u = 0$ because $(\,|\,)$ is non-degenerate. We have proven

that $(u|u) > 0$ if $u \neq 0$. That is, a scalar product is a positive-defined bilinear symmetric (resp. sesquilinear Hermitian) mapping $(\ |\ ) : V \times V \to \mathbb{R}$ (resp. $\mathbb{C}$). Now assume that $(\ |\ )$ is positive-defined bilinear symmetric (resp. sesquilinear Hermitian). By definition it is a semi scalar product since positive definiteness implies positive semi-definiteness. Let us prove that $(|)$ is non-degenerate and this concludes the proof. If $(u|v) = 0$ for all $v \in V$ then, choosing $v = u$, the positive definiteness implies $u = \mathbf{0}$. □

4.2 Natural isomorphism between V and V^* and metric tensor.

Let us show that if V is a finite-dimensional vector space endowed with a pseudo scalar product, V is isomorphic to V^*. That isomorphism is *natural* because it is built up using the structure of vector space with scalar product only, specifying nothing further.

Theorem 4.2. (Natural (anti)isomorphism between V and V^*.) *Let V be a finite-dimensional vector space with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *If V is endowed with a pseudo scalar product $(\ |\ )$ (also if $\mathbb{K} = \mathbb{C}$),*
(i) *the mapping defined on V, $h : u \mapsto (u|\cdot)$, where $(u|\cdot)$ is the linear functional, $(u|\cdot) : v \mapsto (u|v)$, is an isomorphism;*
(ii) *$(h(u)|h(v))^* := (u|v)$ defines a pseudo scalar product on V^*.*
(b) *If $\mathbb{K} = \mathbb{C}$ and V is endowed with a Hermitean pseudo scalar product $(\ |\ )$,*
(i) *the mapping defined on V, $h : u \mapsto (u|\cdot)$, where $(u|\cdot)$ is the linear functional, $(u|\cdot) : v \mapsto (u|v)$, is an anti isomorphism;*
(ii) *$(h(u)|h(v))^* := \overline{(u|v)}$ $(= (v|u))$ defines a Hermitian pseudo scalar product on V^*.*

Proof. First consider (i) in the cases (a) and (b). It is obvious that $(u|\cdot) \in V^*$ in both cases. Moreover the linearity or antilinearity of the mapping $u \mapsto (u|\cdot)$ is a trivial consequence of the definition of pseudo scalar product and Hermitian pseudo scalar product respectively.
Then remind the well-known theorem, $dim(Kerf) + dimf(V) = dimV$, which holds true for linear and anti-linear mappings from some finite-dimensional vector space V to some vector space V'. Since $dimV = dimV^*$, it is sufficient to show that $h : V \to V^*$ defined by $u \mapsto (u|\cdot)$ has trivial kernel, i.e., is injective: this also assures the surjectivity of the map. Therefore, we have to show that $(u|\cdot) = (u'|\cdot)$ implies $u = u'$. This is equivalent to show that $(u - u'|v) = 0$ for all $v \in V$ implies $u - u' = 0$. This is nothing but the non-degenerateness property, which holds by definition of (Hermitean) scalar product.
Statements (ii) cases are obvious in both by definition of (Hermitian) pseudo scalar products using the fact that h is a (anti) isomorphism. □

Remarks.
(1) Notice that, with the definitions given above it holds also $(u|v)^* = (h^{-1}u|h^{-1}v)$ and, for the Hermitian case, $(u|v)^* = \overline{(h^{-1}u|h^{-1}v)}$ for all $u, v \in V^*$. This means that h and h^{-1} *(anti)preserve* the scalar products.
(2) The theorem above holds also considering a Hilbert space and its topological dual space

(i.e., the subspace of the dual space consisting of continuous linear functionals on the Hilbert space). That is the mathematical content of celebrated Riesz' representation theorem.

Exercise 4.2.
4.2.1. Show that, for all $u, v \in V$:

$$(u|v) = \langle u, h(v)\rangle = (h(v)|h(u))^* \ ,$$

no matter if $(\ |\ )$ is Hermitian or not.

From now on we specialize to the pseudo-scalar-product case dropping the Hermitian case. Suppose $(|)$ is a pseudo scalar product on a finite-dimensional vector space V with field $\mathbb{K}$ either $\mathbb{R}$ or $\mathbb{C}$. The mapping $(u, v) \mapsto (u|v)$ is bi linear on $V \times V$ and thus is a tensor $g \in V^* \otimes V^*$. Fixing a canonical basis in $V^* \otimes V^*$ induced by a basis $\{e_i\}_{i\in I} \subset V$, we can write:

$$\mathbf{g} = g_{ij} e^{*i} \otimes e^{*j} \ ,$$

where, by **Theorem 1.5**,

$$g_{ij} = (e_i|e_j) \ .$$

Def.4.4. (Metric Tensor.) *A pseudo scalar product* $(|) = \mathbf{g} \in V^* \otimes V^*$ *on a finite-dimensional vector space* V *with field* $\mathbb{K}$ *either* $\mathbb{R}$ *or* $\mathbb{C}$ *is called* **pseudo-metric tensor**. *If* $\mathbb{K} = \mathbb{R}$, *a pseudo-metric tensor is called* **metric tensor** *if it defines a scalar product.*

Remarks.
(1) By **Theorem 2.2**, the isomorphism $h : V \to V^*$ is represented by a tensor of $V^* \otimes V^*$ which acts on elements of V by means of a product of tensors and a contraction. The introduction of the pseudo-metric tensor allows us to represent the isomorphism $h : V \to V^*$ by means of the abstract index notation determining the tensor representing h. Indeed, since $h : u \mapsto (u|\ )$ and $(u|v) = (e_i|e_j)u^i v^j = g_{ij} v^i u^j$ we trivially have:

$$(hu)_i = g_{ij} u^j \ .$$

Hence h is represented by g itself.
(2) Notice that pseudo-metric tensors are **symmetric** because of the symmetry of pseudo scalar products:

$$\mathbf{g}(u, v) = (u|v) = (v|u) = \mathbf{g}(v, u) \ .$$

Components of pseudo-metric tensors with respect to canonical basis enjoy some simple but important properties which are listed below.

Theorem 4.3. (Properties of the metric tensor.) *Referring to* **Def.4.4**, *the components of any pseudo-metric tensor* $\mathbf{g}$, $g_{ij} := \mathbf{g}(e_i, e_j)$ *with respect to the canonical basis induced in*

$V^* \otimes V^*$ by any basis $\{e_i\}_{i=1,\ldots,n} \subset V$, enjoy the following properties:
(1) *define a symmetric matrix $[g_{ij}]$, i.e,*

$$g_{ij} = g_{ji}\,;$$

(2) *$[g_{ij}]$ is non singular, i.e., it satisfies:*

$$det[g_{ij}] \neq 0\,;$$

(3) *if $\mathbb{K} = \mathbb{R}$ and g is a scalar product, the matrix $[g_{ij}]$ is positive defined.*

Proof. **(1)** It is obvious: $g_{ij} = (e_i|e_j) = (e_j|e_i) = g_{ji}$.
(2) Suppose $det[g_{ij}] = 0$ and define $n = dimV$. The linear mapping $\mathbb{K}^n \to \mathbb{K}^n$ determined by the matrix $g := [g_{ij}]$ has a non-trivial kernel. In other words, there are n reals u^j, $j = 1, \ldots, n$ defining a $\mathbb{K}^n$ vector $[u] := (u^1, \ldots, u^n)^t$ with $g[u] = 0$ and $[u] \neq 0$. In particular $[v]^t g[u] = 0$ for whatever choice of $[v] \in \mathbb{K}^n$. Defining $u := u^j e_j$, the obtained result implies that there is $u \in V \setminus \{0\}$ with $(u|v) = (v|u) = 0$ for all $v \in V$. This is impossible because $(\,|\,)$ is non degenerate by hypothesis.
(3) The statement, $(u|u) > 0$ if $u \in V \setminus \{\mathbf{0}\}$, reads, in the considered canonical basis $[u]^t g[u] > 0$ for $[u] \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. That is one of the equivalent definitions of a positive defined matrix g. □

The following theorem shows that a (pseudo) scalar product can be given by the assignment of a convenient tensor which satisfies some properties when represented in some canonical bases. The important point is that there is no need to check on these properties for *all* canonical bases, verification for a single canonical basis is sufficient.

Theorem 4.4 (Assignment of a (pseudo) scalar product.) *Let V be a finite-dimensional vector space with field $\mathbb{K}$ either $\mathbb{R}$ or $\mathbb{C}$. Suppose $\mathbf{g} \in V^* \otimes V^*$ is a tensor such that there is a canonical basis of $V^* \otimes V^*$ where the components g_{ij} of $\mathbf{g}$ define a symmetric matrix $g := [g_{ij}]$ with non-vanishing determinant. Then $\mathbf{g}$ is a pseudo-metric tensor, i.e. a pseudo scalar product. Furthermore, if $\mathbb{K} = \mathbb{R}$ and $[g_{ij}]$ is positive defined, the pseudo scalar product is a scalar product.*

Proof. If $\mathbf{g}$ is represented by a symmetric matrix of components in a canonical basis then it holds in all remaining bases and the tensor is symmetric (see **Exercise 3.2.1**). This implies that $(u|v) := \mathbf{g}(u, v)$ is a bi-linear symmetric functional. Suppose $(\,|\,)$ is degenerate, then there is $u \in V$ such that $u \neq \mathbf{0}$ and $(u|v) = 0$ for all $v \in V$. Using notations of the proof of the item (2) of **Theorem 4.3**, we have in components of the considered canonical bases, $[u]^t g[v] = 0$ for all $[v] = (v^1, \ldots, v^n)^t \in \mathbb{K}^n$ where $n = dimV$. Choosing $[v] = g[u]$, it also holds $[u]^t g g[u] = 0$. Since $g = g^t$, this is equivalent to $(g[u])^t g[u] = 0$ which implies $g[u] = 0$. Since $[u] \neq \mathbf{0}$, g cannot be injective and $detg = 0$. This is not possible by hypotheses, thus $(\,|\,)$ is non-degenerate. We conclude that $(u|v) := g(u, v)$ define a pseudo scalar product.
Finally, if $\mathbb{K} = \mathbb{R}$ and g is also positive defined, $(\,|\,)$ itself turns out to be positive defined, i.e., it is a scalar product since $(u|u) = [u]^t g[u] > 0$ if $[u] \neq \mathbf{0}$ (which is equivalent to $u \neq \mathbf{0}$). □

Let us introduce the concept of *signature* of a pseudo-metric tensor in a vector space with field $\mathbb{R}$ by reminding Sylvester's theorem whose proof can be found in any linear algebra textbook. The definition is interlaced with the definition of an *orthonormal basis*.

Theorem 4.5 (Sylvester's theorem.) *Let A be a real symmetric $n \times n$ matrix.*
(a) *There is a non-singular (i.e., with non vanishing determinant) real $n \times n$ matrix D such that:*

$$DAD^t = diag(0, \ldots, 0, -1, \ldots, -1, +1, \ldots, +1) \ ,$$

where the reals $0, -1, +1$ appear $v \geq 0$ times, $m \geq 0$ times and $p \geq 0$ times respectively with $v + m + p = n$.
(b) *the triple (v, m, p) does not depend on D. In other words, if, for some non-singular real $n \times n$ matrix matrix $E \neq D$, EAE^t is diagonal and the diagonal contains reals $0, -1, +1$ only (in whatever order), then $0, -1, +1$ respectively appear v times, m times and p times.*

If $\mathbf{g} \in V^* \otimes V^*$ is a pseudo-metric tensor on the finite-dimensional vector space v with field $\mathbb{R}$, the transformation rule of the components of g with respect to canonical bases (see **Theorem 2.1**) induced by bases $\{e_i\}_{i \in I}$, $\{e'_j\}_{j \in I}$ of V are

$$g'_{pq} = B_p{}^i B_q{}^j g_{ij} \ .$$

Defining $g' := [g_{pq}]$, $g := [g_{ij}]$, $B := [B_h{}^k]$, they can be re-written as

$$g' = BgB^t \ .$$

We remind (see **Theorem 2.1**) that the non-singular matrices B are defined by $B = A^{-1t}$, where $A = [A^i{}_j]$ and $e_m = A^l{}_m e'_l$. Notice that the specification of B is completely equivalent to the specification of A because $A = B^{-1t}$.
Hence, since g is real and symmetric by **Theorem 4.3**, Sylvester's theorem implies that, starting from any basis $\{e_i\}_{i \in I} \subset V$ one can find another basis $\{e'_j\}_{j \in I}$, which induces a canonical basis in $V^* \otimes V^*$ where the pseudo-metric tensor is represented by a diagonal matrix. It is sufficient to pick out a transformation matrix B as specified in **(a)** of **Theorem 4.5**. In particular, one can find B such that each element on the diagonal of g' turns out to be either -1 or $+1$ only. The value 0 is not allowed because it would imply that the matrix has vanishing determinant and this is not possible because of **Theorem 4.3**. Moreover the pair (m, p), where (m, p) are defined in **Theorem 4.5**, does not depend on the basis $\{e'_j\}_{j \in I}$. In other words, it is an *intrisic* property of the pseudo-metric tensor: that is the *signature* of the pseudo-metric tensor.

Def.4.5 (Pseudo Orthonormal Bases and Signature). *Let $\mathbf{g} \in V^* \otimes V^*$ be a pseudo-metric tensor on the finite-dimensional vector space V with field $\mathbb{R}$.*
(a) *A basis $\{e_i\}_{i \in I} \subset V$ is called* **pseudo orthonormal** *with respect to $\mathbf{g}$ if the components of $\mathbf{g}$ with respect to the canonical basis induced in $V^* \otimes V^*$ form a diagonal matrix with eigenvalues in $\{-1, +1\}$. In other words, $\{e_i\}_{i \in I}$ is pseudo orthonormal if*

$$(e_i, e_j) = \pm \delta_{ij} \ .$$

If the pseudo-metric tensor is a metric tensor the pseudo-orthonormal bases are called orthonormal bases.
(b) *The pair* (m, p)*, where* m *is the number of eigenvalues* -1 *and* p *is the number of eigenvalues* $+1$ *of a matrix representing the components of* $\mathbf{g}$ *in an orthonormal basis is called* **signature** *of* $\mathbf{g}$.
(c) $\mathbf{g}$ *and its signature are said* **elliptic** *or* **Euclidean** *or* **Riemannian** *if* $m = 0$*,* **hyperbolic** *if* $m > 0$ *and* $p \neq 0$*,* **Lorentzian** *or* **normally hyperbolic** *if* $m = 1$ *and* $p \neq 0$.
(d) *If* $\mathbf{g}$ *is hyperbolic, an orthonormal basis* $\{e_i\}_{i\in I}$ *is said to be* **canonical** *if the matrix of the components of* $\mathbf{g}$ *takes the form:*

$$diag(-1, ..., -1, +1, ..., +1) \ .$$

Exercises. 4.2.
4.2.1. Show that a *pseudo-metric* tensor $\mathbf{g}$ is a *metric* tensor if and only if its signature is elliptic.

Remark. If $\{e_i\}_{i\in I}$ is an orthonormal basis with respect to a hyperbolic pseudo-metric tensor g, one can trivially re-order the vectors of the basis giving rise to a canonical orthonormal basis.

Comment. Let us consider a pseudo-metric tensor $\mathbf{g}$ in V with field $\mathbb{R}$. Let (m, p) be the signature of $\mathbf{g}$ and let $\mathcal{N}_g$ be the class of all of the canonical pseudo-orthonormal bases in V with respect to $\mathbf{g}$. In the following we shall indicate by η the matrix $diag(-1, ..., -1, +1, ..., +1)$ which represents the components of $\mathbf{g}$ with respect to each basis of $\mathcal{N}_{\mathbf{g}}$. If A is a matrix corresponding to a change of basis in $\mathcal{N}_{\mathbf{g}}$, and $B := A^{-1t}$ is the associated matrix concerning change of basis in V^*, it has to hold

$$\eta = B\eta B^t \ .$$

Conversely, each real $n \times n$ matrix B which satisfies the identity above determines $A = B^{-1t}$ which represents a change of basis in $\mathcal{N}_{\mathbf{g}}$. In particular each B which satisfies the identity above must be *non singular* because A is such. Furthermore, taking the determinant of both sides in the identity above and taking $det\eta = (-1)^m$ into account, we have that $(detB)^2 = 1$ and thus $detB = \pm 1$. Where we have also used $detB = detB^t$.
Noticing that $\eta = \eta^{-1}$ and $A = B^{-1t}$, the identity above can be equivalently re-written in terms of the matrices A:

$$\eta = A\eta A^t \ .$$

The equation above completely determines the set $O(m, p) \subset GL(n, \mathbb{R})$ $(n = m + p)$ of all real non-singular $n \times n$ matrices which correspond to changes of bases in $\mathcal{N}_{\mathbf{g}}$.
It is possible to show that $O(m, p)$ is a subgroup of $GL(n, \mathbb{R})$ called the **pseudo orthogonal** group of order (m, p). Notice that, if $m = 0$, $O(0, p) = O(n)$ reduces to the usual orthogonal group of order n. $O(1, 3)$ is the celebrated **Lorentz group** which is the central mathematical object in relativistic theories.

Exercises. 4.3.
4.3.1. Show that if $A \in O(m,p)$ then A^{-1} exists and

$$A^{-1} = \eta A^t \eta \, .$$

4.3.2. Show that $O(m,p)$ is a group with respect to the usual multiplication of matrices.
Note. This implies that $O(m,p)$ is a subgroup of $GL(n,\mathbb{R})$ with $n = p + m$.
(Hint. You have to prove that, (1) the identity matrix I belongs to $O(m,p)$, (2) if A and A' belong to $O(m,p)$, AA' belongs to $O(m,p)$, (3) if A belongs to $O(m,p)$, then A^{-1} exists and belongs to $O(m,p)$.)
4.3.3. Show that $SO(m,p) := \{A \in O(m,p) \mid detA = 1\}$ is not the empty set and is a subgroup of $O(m,p)$.
Note. $SO(m,p)$ is called the **special pseudo orthogonal group** of order (m,p).
4.3.4. Consider the special Lorentz group $SO(1,3)$ and show that the set

$$S\uparrow O(m,p) := \{A \in SO(1,3) \ | \ A^1{}_1 > 0\}$$

is a not empty subgroup.
Note. $S\uparrow O(m,p)$ is called the **special orthocronous Lorentz group**.

4.3 Raising and lowering of indices of tensors.

Consider a finite dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ endowed with a pseudo-metric tensor $\mathbf{g}$. As we said above, there is a natural isomorphism $h : V \to V^*$ defined by $h : u \mapsto (u|\cdot) = \mathbf{g}(u,\cdot)$. This isomorphism may be extended to the whole tensor algebra $\mathcal{A}_{\mathbb{K}}(V)$ using the universality theorem and **Def.3.4**.
Indeed, consider a space $S \in \mathcal{A}_{\mathbb{K}}(V)$ of the form $A \otimes V \otimes B$, where A and B are tensor spaces of the form $U_1 \otimes \ldots \otimes U_k$, and $U_{k+1} \otimes \ldots \otimes U_m$ respectively, U_i being either V or V^*. We may define the operators:

$$h^{\otimes} := I_1 \otimes \ldots \otimes I_k \otimes h \otimes I_{k+1} \otimes \ldots \otimes I_m : A \otimes V \otimes B \to A \otimes V^* \otimes B \, ,$$

and

$$(h^{-1})^{\otimes} := I_1 \otimes \ldots \otimes I_k \otimes h^{-1} \otimes I_{k+1} \otimes \ldots \otimes I_m : A \otimes V^* \otimes B \to A \otimes V \otimes B \, ,$$

where $I_j : U_j \to U_j$ is the identity operator. Using *Remark* after **Def.3.4**, one finds

$$(h^{-1})^{\otimes} h^{\otimes} = I_1 \otimes \ldots \otimes I_k \otimes (h^{-1}h) \otimes I_{k+1} \otimes \ldots \otimes I_m = Id_{A \otimes V \otimes B} \, ,$$

and

$$h^{\otimes} (h^{-1})^{\otimes} = I_1 \otimes \ldots \otimes I_k \otimes (hh^{-1}) \otimes I_{k+1} \otimes \ldots \otimes I_m = Id_{A \otimes V^* \otimes B} \, .$$

Therefore $h^{\otimes}$ is an isomorphism with inverse $(h^{-1})^{\otimes}$.

The action of $h^{\otimes}$ and $(h^{-1})^{\otimes}$ is that of **lowering** and **raising indices** respectively. In fact, in abstract index notation, one has:

$$h^{\otimes} : t^{AiB} \mapsto t^{A}{}_{j}{}^{B} := t^{AiB} g_{ij} \ ,$$

and

$$(h^{-1})^{\otimes} : u^{A}{}_{i}{}^{B} \mapsto u^{AjB} := t^{A}{}_{i}{}^{B} \tilde{g}^{ij} \ .$$

Above g_{ij} represents the pseudo-metric tensor as specified in *Remark* **1** after **Def.4.4**. What about the tensor $\tilde{g} \in V \otimes V$ representing h^{-1} via **Theorem 2.2**?

Theorem 4.6. *Let $h : V \to V^*$ be the isomorphism determined by a pseudo scalar product, i.e. a pseudo-metric tensor $\mathbf{g}$ on the finite-dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *The inverse mapping $h^{-1} : V^* \to V$ is represented via* **Theorem 2.2** *by a* **symmetric** *tensor $\tilde{\mathbf{g}} \in V \otimes V$ such that, if $\{e_i\}_{i \in I}$ is a basis of V, $\tilde{g}^{rs} := \tilde{g}(e^{*r}, e^{*s})$ and $g_{ij} := \mathbf{g}(e_i, e_j)$, then the matrix $[\tilde{g}^{ij}]$ is the* inverse matrix *of $[g_{ij}]$.*
(b) *The tensor $\tilde{\mathbf{g}}$ coincides with the pseudo-metric tensor* **with both indices raised**.

Proof. **(a)** By **Theorem 2.2**, h^{-1} determines a tensor $\tilde{\mathbf{g}} \in V \otimes V$ with $h^{-1}(u^*) = \tilde{\mathbf{g}}(u^*, \cdot)$. In components $(h^{-1}u^*)^i = u^*_k \tilde{g}^{ki}$. On the other hand it must be

$$h(h^{-1}u^*) = u^*$$

or,

$$u^*_k \tilde{g}^{ki} g_{ir} = u^*_r \ ,$$

for all $u^* \in V^*$. This is can be re-written

$$[u^*]^t (\tilde{g} g - I) = 0 \ ,$$

for all $\mathbb{K}^n$ vectors $[u^*] = (u^*_1, \ldots, u^*_n)$. Then the matrix $(\tilde{g} g - I)^t$ is the null matrix. This implies that

$$\tilde{g} g = I \ ,$$

which is the thesis. $\tilde{g}$ is symmetric because is the inverse of a symmetric matrix and thus also the tensor $\tilde{\mathbf{g}}$ is symmetric.
(b) Let g^{ij} be the pseudo-metric tensor with both indices raised, i.e.,

$$g^{ij} := g_{rk} \ \tilde{g}^{kj} \tilde{g}^{ri} \ .$$

By **(a)**, the right-hand side is equal to:

$$\delta^j_r \tilde{g}^{ri} = \tilde{g}^{ji} = \tilde{g}^{ij} \ .$$

That is the thesis. □

Remark. Another result which arises from the proof of the second part of the theorem is that

$$g_i^j = \delta_i^j \ .$$

Comments.
(1) When a vector space is endowed with a pseudo scalar product, tensors can be viewed as abstract objects which may be represented either as covariant or contravariant concrete tensors using the procedure of raising and lowering indices. For instance, a tensor t^{ij} of $V \otimes V$ may be viewed as a covariant tensor when "represented" in its covariant form $t_{pq} := g_{pi}\, g_{qj}\, t^{ij}$. Also, it can be viewed as a mixed tensor $t_p{}^j := g_{pi}\, t^{ij}$ or $t^i{}_q := g_{qj}\, t^{ij}$.
(2) Now consider a finite-dimensional vector space on $\mathbb{R}$, V, endowed with a metric tensor $\mathbf{g}$, i.e., with *elliptic signature*. In orthonormal bases the contravariant and covariant components numerically coincides because $g_{ij} = \delta_{ij} = g^{ij}$. This is the reason because, using the usual scalar product of vector spaces isomorphic to $\mathbb{R}^n$ and working in orthonormal bases, the difference between covariant and contravariant vectors does not arise.
Conversely, in relativistic theories where a Lorentzian scalar product is necessary, the difference between covariant and contravariant vectors turns out to be evident also in orthonormal bases, since the diagonal matrix $[g_{ij}]$ takes an eigenvalue -1.

5 Pseudo tensors, Ricci's pseudotensor and tensor densities.

This section is devoted to introduce very important tools either in theoretical/mathematical physics and in pure mathematics: pseudo tensors and tensor densities.

5.1 Orientation and pseudo tensors.

The first example of "pseudo" object we go to discuss is a *orientation* of a *real* vector space.

Consider a finite-dimensional vector space V with field $\mathbb{R}$. In the following $\mathcal{B}$ indicates the set of all the vector bases of V. Consider two bases $\{e_i\}_{i\in I}$ and $\{e'_j\}_{j\in I}$ in $\mathcal{B}$. Concerning the determinant of the transformation matrix $A := [A^r{}_s]$, with $e_i = A^j{}_i e'_j$, there are two possibilities only: $detA > 0$ or $detA < 0$. It is a trivial task to show that the relation in $\mathcal{B}$:

$$\{e_i\}_{i\in I} \sim \{e'_j\}_{j\in I} \quad \text{iff} \quad detA > 0$$

where A indicates the transformation matrix as above, is an *equivalence relation*. Since there are the only two possibilities above, the partition of $\mathcal{B}$ induced by $\sim$ is made of two *equivalence classes* $\mathcal{B}_1$ and $\mathcal{B}_2$. Hence if a basis belongs to $\mathcal{B}_1$ or $\mathcal{B}_2$ any other basis belongs to the same set if and only if the transformation matrix has positive determinant.

Def.5.1. (Orientation of a vector space.) *Consider a finite-dimensional vector space V with field $\mathbb{R}$, an* **orientation** *of V is a bijective mapping $\mathcal{O} : \{\mathcal{B}_1, \mathcal{B}_2\} \to \{-1, +1\}$. If V has an orientation $\mathcal{O}$, is said to be* **oriented** *and a basis $\{e_i\}_{i\in I} \in \mathcal{B}_k$ is said to be* **positive oriented** *if $\mathcal{O}(\mathcal{B}_k) = +1$ or* **negative oriented** *if $\mathcal{O}(\mathcal{B}_k) = -1$.*

Comment. The usual physical vector space can be orietend "by hand" using the natural basis given by our own right hand. When we use the right hand to give an orientation we determine $\mathcal{O}^{-1}(+1)$ by the exhibition of a basis contained therein.

The given definition can be, in some sense, generalized with the introduction of the concept of *pseudo tensor*.

Def.5.2. (Pseudotensors.) *Let V be a finite-dimensional vector space with field $\mathbb{R}$. Let S be a tensor space of $\mathcal{A}_{\mathbb{R}}(V)$. A* **pseudo tensor of** *S is a bijective mapping $t_s : \{\mathcal{B}_1, \mathcal{B}_2\} \to \{s, -s\}$, where $s \in S$. Moreover:*
(a) *the various tensorial properties enjoyed by both s and $-s$ are attributed to t_s. (So, for instance, if s, and thus $-s$, is symmetric, t_s is said to be symmetric);*
(b) *if $\{e_i\}_{i\in I} \in \mathcal{B}_i$, the* **components of** *t_s with respect to the canonical bases induced by that basis are the components of $t_s(\mathcal{B}_i)$.*

Remarks.
(1) The given definition encompasses the definition of *pseudo scalar*.
(2) It is obvious that the assignment of a pseudo tensor t_s of, for instance, $V^{n\otimes} \otimes V^{*m\otimes}$, is equivalent to the assignment of components

$$t^{i_1 \ldots i_n}{}_{j_1 \ldots j_m}$$

for each canonical basis

$$\{e_{i_1} \otimes \ldots e_{i_n} \otimes e^{*j_1} \otimes \cdots \otimes e^{*j_m}\}_{i_1,\ldots,i_n,j_1,\ldots,j_m \in I}$$

such that the transformation rules passing to the basis

$$\{e'_{r_1} \otimes \ldots e'_{r_n} \otimes e'^{*l_1} \otimes \cdots \otimes e'^{*l_m}\}_{r_1,\ldots,r_n,l_1,\ldots,l_m \in I} \ ,$$

are given by:

$$t'^{k_1 \ldots k_n}{}_{h_1 \ldots h_m} = \frac{detA}{|detA|} \ A^{k_1}_{\ i_1} \cdots A^{k_n}_{\ i_n} \ B_{h_1}^{\ j_1} \cdots B_{h_m}^{\ j_m} \ t^{i_1 \ldots i_n}{}_{j_1 \ldots j_n} \ ,$$

where $e_l = A^m_{\ l} e'_m$ and $B = A^{-1t}$ with $B := [B_k^{\ j}]$ and $A := [A^p{}_q]$.
In fact, $t^{i_1 \ldots i_n}{}_{j_1 \ldots j_m} = s^{i_1 \ldots i_n}{}_{j_1 \ldots j_m}$ if the considered base is in $t_s^{-1}(+1)$ or $t^{i_1 \ldots i_n}{}_{j_1 \ldots j_m} = (-s)^{i_1 \ldots i_n}{}_{j_1 \ldots j_m}$ if the considered base is in $t_s^{-1}(-1)$.

Example 5.1. Consider the **magnetic field** $B = B^i e_i$ where e_1, e_2, e_3 is a right-hand orthonormal basis of the space V_3 of the vectors with origin in a point of the physical space E_3. Actually, as every physicist knows, changing basis, the components of B changes as usual only if the new basis is a right-hand basis, otherwise a sign $-$ appears in front of each component. That is a physical requirement due to the Lorentz law. This means that the magnetic field has to be represented in terms of **pseudo vectors**.

5.2 Ricci's pseudo tensor.

A particular pseudo tensor is Ricci's one which is very important in physical applications. The definition of this pseudo tensor requires a preliminary discussion.
Consider a finite-dimensional vector space V with a pseudo-metric tensor $\mathbf{g}$. We know that, changing basis $\{e_i\}_{i \in I} \to \{e'_j\}_{j \in I}$, the components of the pseudo-metric tensor referred to the corresponding canonical bases, transform as:

$$g' = BgB^t \ ,$$

where $g = [g_{ij}]$, $g' = [g'_{pq}]$ and $B = A^{-1t}$, $A := [A^p{}_q]$, $e_l = A^m{}_l e'_m$. This implies that

$$detg' = (detB)^2 \ detg \ , \tag{3}$$

which is equivalent to

$$\sqrt{|detg'|} = |detA|^{-1} \ \sqrt{|detg|} \tag{4}$$

Now fix a basis $\{e_i\}_{i\in I}$ in V and consider the canonical basis induced in $V^{*n\otimes}$ where $n = dimV$, $\{e^{*i_1} \otimes \ldots \otimes e^{*i_n}\}_{i_1,\ldots,i_n\in I}$. Then consider components $\eta_{i_1\ldots i_n}$ referred to the considered basis, given by:

$$\eta_{i_1\ldots i_n} = \epsilon_{\sigma_{i_1\ldots i_n}} \ ,$$

if $(i_1, \ldots, i_n)$ is a permutated string of $(1, 2, \ldots, n)$, otherwise

$$\eta_{i_1\ldots i_n} = 0 \ ,$$

where $\epsilon_{\sigma_{i_1\ldots i_n}}$ is the *parity* of the *permutation* $\sigma_{i_1\ldots i_n} \in \mathcal{P}_n$ defined by:

$$(\sigma(1), \ldots, \sigma(n)) = (i_1, \ldots, i_n) \ .$$

Finally define the components:

$$\varepsilon_{i_1\ldots i_n} := \sqrt{|detg|} \ \eta_{i_1\ldots i_n} \ .$$

We want to show that, if we define analogous components in each canonical basis of $V^{*n\otimes}$, an anti-symmetric $(0, n)$-order pseudo tensor turns out to be defined by the whole assignment of components.

Taking *Remark* **(2)** above into account, it is sufficient to show that, under a change of basis,

$$\sqrt{|detg'|} \ \eta_{i_1\ldots i_n} = \frac{detA}{|detA|} \ B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \sqrt{|detg|} \ \eta_{j_1\ldots j_n} \ . \tag{5}$$

We start by noticing that:

$$B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \eta_{j_1\ldots j_n} = \sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1\ldots j_n}} \ ,$$

where $(j_1, \ldots, j_n)$ ranges over the set of all the permutated strings of $(1, 2, \ldots, n)$. We consider the various cases separately.

(1) Suppose that $i_p = i_q$ with $p \neq q$. To fix the extent assume $i_1 = i_2$. Therefore

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_1} B_{i_2}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2\ldots j_n}} = \sum_{(j_1,\ldots,j_n)} B_{i_2}^{\ j_1} B_{i_1}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2\ldots j_n}} \ .$$

The right-hand side can be re-written

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_2} B_{i_2}^{\ j_1} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2\ldots j_n}} \ ,$$

that is, interchanging the names of j_1 and j_2,

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_2} B_{i_2}^{\ j_1} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_2 j_1\ldots j_n}} \ ,$$

But

$$\epsilon_{\sigma_{j_2 j_1 \dots j_n}} = -\epsilon_{\sigma_{j_1 j_2 \dots j_n}} \ ,$$

so that we finally get

$$\sum_{(j_1,\dots,j_n)} B_{i_1}^{\ j_1} B_{i_2}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \dots j_n}} = - \sum_{(j_1,\dots,j_n)} B_{i_1}^{\ j_1} B_{i_2}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \dots j_n}} \ .$$

In other words, *if* $i_1 = i_2$,

$$B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \sqrt{|detg|} \ \eta_{j_1 \dots j_n} = 0$$

and consequently (5) holds:

$$\sqrt{|detg'|} \ \eta_{i_1 \dots i_n} = \frac{detA}{|detA|} \ B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \sqrt{|detg|} \ \eta_{j_1 \dots j_n} \ .$$

because the left-hand side vanishes too. The remaining subcases of $i_p = i_q$ with $p \neq q$, can be treated analogously.

(2) Next we pass to consider the case of $i_k = k$ for $k = 1, 2, \dots, n$. In that case (5), that we need to prove, reduces to

$$\sqrt{|detg'|} = \sqrt{|detg|} \frac{detA}{|detA|} \ B_1^{\ j_1} \cdots B_n^{\ j_n} \quad \eta_{j_1 \dots j_n} \ ,$$

that is

$$\frac{1}{detA} = B_1^{\ j_1} \cdots B_n^{\ j_n} \quad \eta_{j_1 \dots j_n} \ , \tag{6}$$

where we have used (4). (6) can equivalently be re-written

$$detB = B_1^{\ j_1} \cdots B_n^{\ j_n} \quad \eta_{j_1 \dots j_n} \ ,$$

which is trivially true by the properties of the determinant.

(3) To conclude, it remains to consider the case of a permutation $(i_1, i_2, \dots, i_n)$ of $(1, 2, \dots, n)$. In that case, taking (4) into account, the identity (5) which has to be proven, reduces to

$$detB \epsilon_{\sigma_{i_1 \dots i_n}} = B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \quad \eta_{j_1 \dots j_n} \ . \tag{7}$$

Let us prove (7). Suppose for instance that $i_1 = 2$ and $i_2 = 1$ while $i_k = k$ if $k > 2$. In that case (7) holds true because the left-hand side is noting but $-detB$ and the right-hand side can be re-written

$$\sum_{(j_1,\dots,j_n)} B_2^{\ j_1} B_1^{\ j_2} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \dots j_n}} = \sum_{(j_1,\dots,j_n)} B_1^{\ j_2} B_2^{\ j_1} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \dots j_n}} \ ,$$

that is, interchanging the names of j_1 and j_2,

$$\sum_{(j_1,\dots,j_n)} B_1^{\ j_1} B_2^{\ j_2} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_2 j_1 \dots j_n}}$$

which, in turn, equals just

$$-\sum_{(j_1,\ldots,j_n)} B_1{}^{j_1} B_2{}^{j_2} \cdots B_n{}^{j_n} \epsilon_{\sigma_{j_1 j_2 \cdots j_n}} = -detB \, .$$

If $(i_1, \ldots, i_n)$ is different from $(1, \ldots, 2)$ just for a transposition only, the same procedure can be adapted trivially. If $(i_1, \ldots, i_n)$ is a proper permutation of $(1, \ldots, 2)$, it can be decomposed as a product of N transpositions. Concerning the right hand side of (7), using the procedure above for each transposition, we get in the end that it can be re-written :

$$B_{i_1}^{j_1} \cdots B_{i_n}^{j_n} \quad \eta_{j_1 \ldots j_n} = detB(-1)^N \, .$$

On the other hand, in the considered case, the left-hand side of (7) equals

$$detB(-1)^N \, ,$$

so that (7) holds true once again. This concludes the proof.

Def.5.3. (Ricci's Pseudo tensor). *Let V be a vector space with field $\mathbb{R}$ and dimension $n < +\infty$, endowed with a pseudo-metric tensor* $\mathbf{g}$. **Ricci's pseudo tensor** *is the anti-symmetric $(0, n)$ pseudo tensor ε represented in each canonic basis by components*

$$\varepsilon_{i_1 \ldots i_n} := \sqrt{|detg|} \; \eta_{i_1 \ldots i_n} \, ,$$

where $g = [g_{ij}]$, g_{ij} being the components of g in the considered basis and

$$\eta_{i_1 \ldots i_n} = \epsilon_{\sigma_{i_1 \ldots i_n}} \, ,$$

if $(i_1, \ldots, i_n)$ is a permutated string of $(1, 2, \ldots, n)$, otherwise

$$\eta_{i_1 \ldots i_n} = 0 \, ,$$

and $\epsilon_{\sigma_{i_1 \ldots i_n}}$ is the parity of the permutation $\sigma_{i_1 \ldots i_n} \in \mathcal{P}_n$ defined by:

$$(\sigma(1), \ldots, \sigma(n)) = (i_1, \ldots, i_n) \, .$$

Ricci's pseudo tensor has various applications in mathematical physics in particular when it is used as a linear operator which produces pseudo tensors when acts on tensors. In fact, consider $t \in V^{n\otimes}$ and take an integer $m \leq n$. Fix a basis in V and, in the canonical bases induced by that basis, consider the action of ε on t:

$$t^{i_1 \ldots i_n} \mapsto \tilde{t}_{j_1 \ldots j_{n-m}} := \varepsilon_{j_1 \ldots j_{n-m} i_1 \ldots i_n} t^{i_1 \ldots i_n} \, .$$

We leave to the reader the proof of the fact that the components $\tilde{t}_{j_1,\ldots,j_{n-m}}$ define a anti-symmetric pseudo tensor of order $(0, n-m)$ which is called the **conjugate** pseudo tensor of t.

Example 5.1. As a trivial but very important example consider the **vector product** in a three-dimensional vector space V on the field $\mathbb{R}$ endowed with a metric tensor g. If $u, v \in V$ we may define the pseudo vector of order $(1, 0)$:

$$(u \wedge v)^r := g^{ri} \varepsilon_{ijk} u^j v^k \ .$$

If $\{e_i\}_{i=1,2,3}$ is an orthonormal basis in V, everything strongly semplificates. In fact, the Ricci tensor is represented by components

$$\varepsilon_{ijk} := 0$$

if (i, j, k) is not a permutation of $(1, 2, 3)$ and, otherwise,

$$\varepsilon_{ijk} := \pm 1 \ ,$$

where $+1$ corresponds to *cyclic* permutations of $1, 2, 3$ and -1 to *non-cyclic* permutations (see **Examples 3.1.2**). In such a basis:

$$(u \wedge v)_i = \varepsilon_{ijk} \, u^j v^k \ ,$$

because $g^{ij} = \delta^{ij}$ in each orthonormal bases.

If V is oriented, it is possible to define ε as a *proper tensor* instead a pseudo tensor. In this case one defines the components in a canonical basis associated with a positive-oriented basis ov V as

$$\varepsilon_{i_1 \ldots i_n} := \sqrt{|detg|} \ \eta_{i_1 \ldots i_n} \ ,$$

and

$$\varepsilon_{i_1 \ldots i_n} := -\sqrt{|detg|} \ \eta_{i_1 \ldots i_n} \ ,$$

if the used basis is associated with a basis of V which is negative-orientd. One can prove staightforwardly that the defined components give rise to a tensor of $V^{*n\otimes}$ called **Ricci tensor**.
This alternative point of view is equivalent, in the practice, to the other point of view corresponding to Definition 5.3.
An important feature of Ricci's pseudotensor is the formula, with obvious notations:

$$t_{i_1 \ldots i_p} = \frac{|detg|}{detg} \ \frac{(-1)^{p(n-p)}}{p!(n-p)!} \ \varepsilon_{i_1 \ldots i_p j_1 \ldots j_n} \varepsilon^{j_1 \ldots j_n r_1 \ldots r_p} \ t_{r_1 \ldots t_p} \ ,$$

which holds for *anti-symmetric* tensors $t \in V^{*p\otimes}$, $0 \leq p \leq n = dimV$.
This implies that, if the tensor t is anti symmetric, then its conjugate pseudo tensor $\tilde{t}$ takes the same information than t itself.

Exercises 5.1.
5.1.1 Often, the definition of vector product in $\mathbb{R}^3$ is given, in orthonormal basis, as

$$(u \wedge v)_i = \varepsilon_{ijk}\, u^i v^k \ ,$$

where it is assumed that the basis is *right oriented*. Show that it defines a proper vector (and not a pseudo vector) if a convenient definition of $\wedge$ is given in *left oriented* basis.
5.1.2. Is it possible to define a sort of vector product (which maps pair of vectors in vectors) in $\mathbb{R}^4$ generaluzing the vector product in $\mathbb{R}^3$?
5.1.3. In physics literature one may find the statement "Differently from the impulse $\vec{p}$ which is a **polar vector**, the angular momentum $\vec{l}$ is an **axial vector**". What does it mean?
(Solution. Polar vector = vector, Axial vector = pseudo vector.)
5.1.4. Consider the **parity inversion**, $P \in O(3)$, as the active transformation of vectors of physical space defined by $P := -I$ when acting in components of vectors in any orthonormal basis. What do physicists mean when saying "Axial vectors transform differently from polar vectors under parity inversion"?
(Hint. interpret P as a passive transformation, i.e. a changing of basis and extend the result to the active interpretation.)
5.1.5. Can the formula defining the conjugate pseudo tensor of t:

$$t^{i_1 \dots i_n} \mapsto \tilde{t}_{j_1 \dots j_{n-m}} := \varepsilon_{j_1 \dots j_{n-m} i_1 \dots i_m} t^{i_1 \dots i_n} \ ,$$

be generalized to the case where t is a pseudo tensor? If yes, what sort of geometric object is $\tilde{t}$?
5.1.6. Consider a vector product $u \wedge v$ in $\mathbb{R}^3$ using an orthonormal basis. In that basis there is an anti-symmetric matrix which takes the same information as $u \wedge v$ and can be written down using the components of the vectors u and v. Determine that matrix and explain the tensorial meaning of the matrix.
5.1.7. Prove the formula introduced above:

$$t_{i_1 \dots i_p} = \frac{|detg|}{detg} \ \frac{(-1)^{p(n-p)}}{p!(n-p)!} \ \varepsilon_{i_1 \dots i_p j_1 \dots j_n} \varepsilon^{j_1 \dots j_n r_1 \dots r_p} \ t_{r_1 \dots t_p} \ ,$$

for anti-symmetric tensors $t \in V^{*p\otimes}$, $0 \leq p \leq n = dimV$.

5.3 Tensor densities.

In section 5.2 we have seen that the determinant of the matrix representing a pseudo-metric tensor g transforms, under change of basis with the rule

$$detg' = |detA|^{-2} detg$$

where the pseudo-metric tensor is $g'_{ij} e'^{*i} \otimes e'^{*i} = g_{pq} e^{*p} \otimes e^{*q}$ and $A = [A^i{}_j]$ is the matrix used in the change of basis for controvariant vectors $t^i e_i = t'^p e'_p$, that is $t^i = A^i{}_p t^p$. Thus the assignment of the numbers $detg$ for each basis in $\mathcal{B}$ does not define a scalar because of the presence of the

factor $|detA|^{-1}$. Similar mathematical objects plays a relevant role in mathematical/theoretical physics and thus deserve a precise definition.

Def.5.4. (Tensor densities.) *Let V be a finite-dimensional vector space with field $\mathbb{R}$ and $\mathcal{B}$ the class of all bases of V.*
If $S \in \mathcal{A}_{\mathbb{R}}(V)$, a **tensor densities of S with weight** *$w \in \mathbb{Z} \setminus \{0\}$ is a mapping $d : \mathcal{B} \to S$ such that, if $B = \{e_i\}_{i \in I}$ and $B' = \{e'_j\}_{j \in I}$ are two bases in $\mathcal{B}$ with $e_k = A^i{}_k e'_i$ then*

$$d(B') = |detA|^w d(B) \, .$$

where $A = [A^i{}_k]$. Furthermore:
(a) *the various tensorial properties enjoyed by all $d(B)$ are attributed to d. (So, for instance, if a $d(B)$ is symmetric (and thus all $d(B)$ with $B \in \mathcal{B}$ are symmetric), d is said to be symmetric);*
(b) *if $B \in \mathcal{B}$, the* **components of** *d with respect to the canonical bases induced by B are the components of $d(B)$ in those bases.*

If $\mathbf{g}$ is a pseudo-metric tensor on V a trivial example of a density with weight w in, for instance $S = V \otimes V^* \otimes V$, can be built up as follows. Take $t \in V \otimes V^* \otimes V$ and define

$$d_t(\{e_i\}_{i \in I}) := (\sqrt{|detg|})^{-w} t \, ,$$

where g is the matrix of the coefficients of $\mathbf{g}$ in the canonical basis associated with $\{e_i\}_{i \in I}$. In components, in the sense of (b) of the definition above:

$$(d_t)^i{}_j{}^k = (\sqrt{|detg|})^{-w} t^i{}_j{}^k \, .$$

To conclude we give the definition of pseudo tensor density which is the straightforward extension of the definition given above.

Def.5.5. (Pseudo-tensor densities.) *Let V be a finite-dimensional vector space with field $\mathbb{R}$ and $\mathcal{B}$ the class of all bases of V.*
If $S \in \mathcal{A}_{\mathbb{R}}(V)$, a **pseudo-tensor densities of S with weight** *$w \in \mathbb{Z} \setminus \{0\}$ is a mapping $d : \mathcal{B} \to S$ such that, if $B = \{e_i\}_{i \in I}$ and $B' = \{e'_j\}_{j \in I}$ are two bases in $\mathcal{B}$ with $e_k = A^i{}_k e'_i$ then*

$$d(B') = \frac{detA}{|detA|} |detA|^w d(B) \, .$$

where $A = [A^i{}_k]$. Furthermore:
(a) *the various tensorial properties enjoyed by all $d(B)$ are attributed to d. (So, for instance, if a $d(B)$ is symmetric (and thus all $d(B)$ with $B \in \mathcal{B}$ are symmetric), d is said to be symmetric);*
(b) *if $B \in \mathcal{B}$, the* **components of** *d with respect to the canonical bases induced by B are the components of $d(B)$ in those bases.*

Remark. *It is clear that the sets of tensor densities and pseudo-tensor densities of a fixed space S and with fixed weight form linear spaces with composition rule which reduces to usual linear composition rule of components.*

There is an important property of densities with weight -1 which is very useful in integration theory on manifolds. The property is stated in the following theorem.

Theorem 5.1. *Let V be a vector space on $\mathbb{R}$ with dimension $n < +\infty$. There is a natural isomorphism G from the space of scalar densities of weight -1 and the space of antisymmetric covariant tensors of order n. In components, using notation as in Definition 5.3,*

$$G : \alpha \mapsto \alpha \eta_{i_1 \ldots i_n} \ .$$

Proof. Fix a canonical basis of $V^{*n\otimes}$ associated with a basis of V. Any nonvanishing tensor t in space, $\Lambda_n(V)$, of antisymmetric covariant tensors of order $n = dimV$ must have the form $t_{i_1 \ldots i_n} = \alpha \eta_{i_1 \ldots i_n}$ in components because different nonvanishing components can be differ only for the sign due to antisymmetry properties of t. Therefore the dimension of $\Lambda_n(V)$ is 1 which is also the dimension of the space of scalar densities of weight -1. The application G defined above in components from $\mathbb{R}$ to $\Lambda_n(V)$ is linear and surjective and thus is injective. Finally, re-adapting straightforwardly the relevant part of the discussion used to define ϵ, one finds that the coefficient α in $t_{i_1 \ldots i_n} = \alpha \eta_{i_1 \ldots i_n}$ transforms as a scalar densities of weight -1 under change of basis. $\square$

6 Appendix: Square Roots of operators and Polar Decomposition Theorem.

We racall some basic definitions and results which should be known by the reader from elemetary courses of linear algebra.

If $A : V \to V$ is a (linear) operator on any finite-dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, an **eigenvalue** $\lambda \in \mathbb{K}$ of A is a scalar such that

$$(A - \lambda I)u = 0$$

for some $u \in V \setminus \{0\}$. In that case u is called **eigenvector** associated with λ. The set $\sigma(A)$ containing all of the eigenvalues of A is called the **spectrum** of A. The **eigenspace** E_λ associated with $\lambda \in \sigma(A)$ is the subspace of V spanned by the eigenvectors associated with λ.

Proposition A.1. *Let V be a real (complex) finite-dimensional equipped with a (resp. Hermitean) scalar product $(|)$. For every operator $A : V \to V$ there exists exactly one of operator $A^\dagger : V \to V$, called the* **adjoint operator** *of A, such that*

$$(A^\dagger u|v) = (u|Av) \,,$$

for all $u, v \in V$.

Proof. Fix $u \in V$, the mapping $v \mapsto (u|Av)$ is a linear functional and thus an element of V^*. By Theorem 4.2 there is a unique element $w_{u,A} \in V$ such that $(u|Av) = (w_{u,A}|v)$ for all $v \in V$. Consider the map $u \mapsto w_{u,A}$. It holds, if a, b are scalars in the field of V and $u, u' \in V$

$$(w_{au+bu',A}|v) = (au+bu'|Av) = \overline{a}(u|Av)+\overline{b}(u'|Av) = \overline{a}(w_{u,A}|v)+\overline{b}(w_{u',A}|v) = (aw_{u,A}+bw_{u',A}|v).$$

Hence, for all $v \in V$:

$$(w_{au+bu',A} - aw_{u,A} - bw_{u',A}|v) = 0 \,,$$

The scalar product is nondegenerate by definition and this implies

$$w_{au+bu',A} = aw_{u,A} + bw_{u',A} \,.$$

We have obtained that the mapping $A^\dagger : u \mapsto w_{u,A}$ is linear, in other words it is an operator. The uniqueness is trivially proven: if the operator B satisfies $(Bu|v) = (u|Av)$ for all $u, v \in V$, it must holds $((B - A^\dagger)u|v) = 0$ for all $u, v \in V$ which, exactly as we obtained above, entails $(B - A^\dagger)u = 0$ for all $u \in V$. In other words $B = A^\dagger$. $\square$

There are a few simple properties of the adjoint operator whose proofs are straightforward. Below A, B are operators in a real (complex) finite-dimensional vector space V equipped with a (resp. Hermitean) scalar product $(|)$ and a, b belong to the field of V.

(1) $(A^\dagger)^\dagger = A$,

(2) $(aA+bB)^\dagger = \overline{a}A^\dagger + \overline{b}B^\dagger$,
(3) $(AB)^\dagger = B^\dagger A^\dagger$,
(4) $(A^\dagger)^{-1} = (A^{-1})^\dagger$ (if A^{-1} exists).

In a real finite-dimensional vector space V equipped with a scalar product $(|)$, a linear operator $A : V \to V$ is said to be **symmetric** if $(Au|v) = (u|Av)$ for all $u, v \in V$. It is simply proven that A is symmetric if and only if $A = A^\dagger$.

In a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$, a linear operator $A : V \to V$ is said to be **Hermitean** if $(Au|v) = (u|Av)$ for all $u, v \in V$. It is simply proven that $\sigma(A) \subset \mathbb{R}$ if A is Hermitean. It is simply proven that A is Hermitean if and only if $A = A^\dagger$.

In a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$, a linear operator $A : V \to V$ is said to be **unitary** if $(Au|Av) = (u|v)$ for all $u, v \in V$. It is simply shown that every unitary operator is bijective (the space has finite dimension and the operator is injective, $(|)$ being positive). It is simply proven that $\lambda \in \sigma(A)$ entails $|\lambda| = 1$ if A is unitary. It is simply proven that A is unitary if and only if $AA^\dagger = I$ or equivalently $A^\dagger A = I$.

In a real (complex) finite-dimensional vector space V equipped with a (resp. Hermitean) scalar product $(|)$, a linear operator $A : V \to V$ is said to be **normal** if $AA^\dagger = A^\dagger A$. It is clear that symmetric, Hermitean and unitary opertors are normal.

In a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$, If V is that above, and $U \subset V$ is a subspace, the **orthogonal** of U, $U^\perp$, is the subspace of V made of all the vectors which are orthogonal to U, i.e., $v \in U^\perp$ if and only if $(u|v) = 0$ for all $u \in U$. If $w \in V$, the decomposition $w = u + v$ with $u \in U$ and $v \in U^\perp$ is uniquely determined, and the map $P_U : w \mapsto u$ is linear and it is called **orthogonal projector** onto U.

It is possible to show that an operator $P : V \to V$ (V being a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$) is an orthogonal projector onto some subspace $U \subset V$ if and only if both the conditions below hold
(1) $PP = P$,
(2) $P = P^\dagger$.
In that case P is the orthogonal projector onto $U = \{Pv \mid v \in V\}$.
Another pair of useful results concerning orthogonal projectors is the following. Let V be a space as said above, let U, U' be subspaces of V, with P, P' are the corresponding orthogonal projectors P, P'.
(a) U and U' are **orthogonal** to each other, i.e., $U' \subset U^\perp$ (which is equivalent to $U \subset U'^\perp$) if and only if $PP' = P'P = 0$.
(b) $U \subset U'$ if and only if $PP' = P'P = P$.

If V is as above and it has finite dimension n and $A : V \to V$ is normal, there exist a well-known spectral decomposition theorem (the finite-dimensional version of the "spectral theorem").

Proposition A.2 (**Spectral decomposition for normal operators in complex spaces**.) *Let V be a complex finite-dimensional vector space equipped with a Hermitean scalar product* $(|)$.

If the operator $A : V \to V$ is normal (i.e., $A^\dagger A = AA^\dagger$), the following expansion holds:

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda \ ,$$

where P_λ is the orthogonal projector onto the eigenspace associated with λ. Moreover the mapping $\sigma(A) \ni \lambda \mapsto P_\lambda$ satisfies the following two properties:
(1) $I = \sum_\lambda P_\lambda$,
(2) $P_\lambda P_\mu = P_\mu P_\lambda = 0$ *for* $\mu \neq \lambda$.
A **spectral measure**, *i.e. a mapping $B \ni \mu \mapsto P'_\mu$ with $B \subset \mathbb{C}$ finite, P'_μ othogonal projectors and:*
(1)' $I = \sum_{\mu \in B} P'_\mu$,
(2)' $P'_\lambda P'_\mu = P'_\mu P'_\lambda = 0$ *for* $\mu \neq \lambda$,
coincides with $\sigma(A) \ni \lambda \mapsto P_\lambda$ if
(3)' $A = \sum_{\mu \in B} \mu P'_\mu$.

The polar decomposition theorems has many applications in mathematics and physics. In the following we make use of some definitions introduced above. We need two relevant definitions.

Def.A.1. *If V is a real (complex) vector space equipped with a (resp. Hermitean) scalar product $(|)$, an operator $A : V \to V$ is said to be* **positive** *(or* positive semidefined*) if*

$$(u|Au) \geq 0 \quad \textit{for all } u \in V \ .$$

A positive operator A is said to be **strictly positive** *(or* positive defined*) if*

$$(u|Au) = 0 \quad \textit{entails } u = 0 \ .$$

A straightforward consequence of the given definition is the following lemma.

Lemma A.1. *Let V be a complex vector space equipped with a Hermitean scalar product $(|)$. Any positive operator $A : V \to V$, is Hermitean.*
Moreover, if $dimV < \infty$, a normal operator $A : V \mapsto V$
(a) *is positive if and only if $\sigma(A) \subset [0, +\infty)$;*
(b) *is strictly positive if and only if $\sigma(A) \subset (0, +\infty)$.*

Proof. As $(v|Av) \geq 0$, by complex conjugation $(Av|v) = \overline{(v|Av)} = (v|Av)$ and thus

$$((A^\dagger - A)v|v) = 0$$

for all $v \in V$. In general we have:

$$2(Bu|w) = (B(u+w)|(u+w)) + i(B(w+iu)|(w+iu)) - (1+i)(Bw|w) - (1+i)(Bu|u) \ .$$

So that, taking $B = A^\dagger - A$ we get $(Bu|w) = 0$ for all $u, w \in V$ because $(Bv|v) = 0$ for all $v \in V$. $(Bu|w) = 0$ for all $u, w \in V$ entails $B = 0$ or $A^\dagger = A$.
Let us prove (a). Suppose A is positive. We know that $\sigma(A) \subset \mathbb{R}$. Suppose there is $\lambda < 0$ in $\sigma(A)$. Let u be an eigenvector associated with λ. $(u|Au) = \lambda(u|u) < 0$ because $(u|u) > 0$ since $u \neq 0$. This is impossible.
Now assume that A is normal with $\sigma(A) \subset [0, +\infty)$. By Proposition A.2:

$$(u|Au) = \left(\sum_\mu P_\mu u \middle| \sum_\lambda \lambda P_\lambda \sum_\nu P_\nu u \right) = \sum_{\mu,\lambda,\nu} \lambda \left(P_\nu^\dagger P_\lambda^\dagger P_\mu u \middle| u \right) = \sum_{\mu,\lambda,\nu} \lambda \left(P_\nu P_\lambda P_\mu u | \, u \right)$$

because, if P is an orthogonal projector, $P = P^\dagger$. Using Proposition A.2 once again, $P_\nu P_\lambda P_\mu = \delta_{\nu\mu}\delta_{\mu\lambda} P_\lambda$ and thus

$$(u|Au) = \sum_\lambda \lambda \left(P_\lambda u | \, u \right) = \sum_\lambda \lambda \left(P_\lambda u | \, P_\lambda u \right) \ ,$$

where we have used the property of orthogonal projectors $PP = P$. $\lambda \left(P_\lambda u | \, P_\lambda u \right) \geq 0$ if $\lambda \geq 0$ and thus $(u|Au) \geq 0$ for all $u \in V$.
Concerning (b), assume that A is strictly positive (so it is positive and Hermitean). If $0 \in \sigma(A)$ there must exist $u \neq 0$ with $Au = 0u = 0$. That entails $(u, Au) = (u, 0) = 0$ which is not allowed. Therefore $\sigma(A) \subset [0, +\infty)$. Conversely if A is normal with $\sigma(A) \subset (0, +\infty)$, A is posive by (a). If A is not strictly positive, there is $u \neq 0$ such that $(u|Au) = 0$ and thus, using the same procedure as above,

$$(u|Au) = \sum_\lambda \lambda \left(P_\lambda u | \, P_\lambda u \right) = 0 \, .$$

Since $\lambda > 0$ and $\left(P_\lambda u | \, P_\lambda u \right) \geq 0$, it must be $\left(P_\lambda u | \, P_\lambda u \right) = 0$ for all $\lambda \in \sigma(A)$. This means $P_\lambda u = 0$ for all $\lambda \in \sigma(A)$. This is not possible because, using (1) in Proposition 4.2, $0 \neq u = Iu = \sum_\lambda P_\lambda u = 0$. $\square$

Def.A.2. *If V is a complex vector space equipped with a Hermitean scalar product $(|)$, Let $A : V \to V$ a positive operator. If $B : V \to V$ is another positive operator such that $B^2 = A$, B is called a* **square root of** *A.*

Notice that square roots, if they exist, are Hermitean by Lemma A.1. The next theorem proves that the square root of a positive operator exist and is uniquely determined.

Theorem A.1. *If V is a finite-dimensional complex vector space equipped with a Hermitean scalar product $(|)$. Every positive operator $A : V \to V$ admits a unique square root indicated by $\sqrt{A}$. $\sqrt{A}$ is Hermitian and*

$$\sigma(\sqrt{A}) = \{ \sqrt{\lambda} \ | \ \lambda \in \sigma(A) \} \, .$$

Proof. A is Hermitean by Lemma A.1. Using Proposition A.2

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda \,.$$

Since $\lambda \geq 0$ we can define the linear operator

$$\sqrt{A} := \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda \,.$$

By Proposition A.2 we have

$$\sqrt{A}\sqrt{A} = \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda \sum_{\mu \in \sigma(A)} \sqrt{\mu} P_\mu = \sum_{\lambda\mu} \sqrt{\lambda\mu} P_\lambda P_\mu \,.$$

Using property (2)

$$\sqrt{A}\sqrt{A} = \sum_{\lambda\mu} \sqrt{\lambda\mu}\delta_{\mu\nu} P_\lambda = \sum_{\lambda} (\sqrt{\lambda})^2 P_\lambda = \sum_{\lambda} \lambda P_\lambda = A \,.$$

Notice that $\sqrt{A}$ is Hermitean by construction:

$$\sqrt{A}^\dagger := \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda^\dagger = \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda = \sqrt{A} \,.$$

Moreover, if $B = \{\mu = \sqrt{\lambda} \ |\lambda \in \sigma(A)\}$ and $P'_\mu := P_\lambda$ with $\mu = \sqrt{\lambda}$, it holds

$$\sqrt{A} := \sum_{\mu \in B} \mu P'_\mu \,,$$

and $B \ni \mu \mapsto P'_\mu$ satisfy the properties (1)',(2)',(3)' in Proposition A.2. As a consequence it coincides with the spectral measure associated with $\sqrt{A}$,

$$\sqrt{A} := \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda$$

is the unique spectral decomposition of A,

$$\sigma(\sqrt{A}) = \{\mu = \sqrt{\lambda} \ |\lambda \in \sigma(A)\} \,,$$

and thus $\sqrt{A}$ is positive by Lemma A.1. $\sqrt{A}$ is a Hermitean square root of A with the requested spectral properties.
Let us pass to prove the uniqueness property. Suppose there is another square root S of A. As S is positive, it is Hermitean with $\sigma(S) \subset [0, +\infty)$ and it admits a (unique) spectral decomposition

$$S = \sum_{\nu \in \sigma(S)} \nu P'_\mu \,.$$

Define $B := \{\nu^2 \mid \nu \in \sigma(S)\}$. It is simply proven that the mapping $B \ni \lambda \mapsto P'_{\sqrt{\lambda}}$ satisfy:
(1)' $I = \sum_{\lambda \in B} P'_{\sqrt{\lambda}}$,
(2)' $P'_{\sqrt{\lambda}} P'_{\sqrt{\mu}} = P'_{\sqrt{\mu}} P'_{\sqrt{\lambda}} = 0$ for $\mu \neq \lambda$,
(3)' $A = S^2 = \sum_{\lambda \in B} \lambda P'_{\sqrt{\lambda}}$.
Proposition 4.2 entails that the spectral measure of A and $B \ni \lambda \mapsto P'_{\sqrt{\lambda}}$ coincides: $P'_{\sqrt{\lambda}} = P_\lambda$ for all $\lambda \in \sigma(A) = B$. In other words

$$S = \sum_{\nu \in \sigma(S)} \nu P'_\nu = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda = \sqrt{A}\,.$$

□

Theorem A.2. (**Polar Decomposition of operators**.) *Let V be a finite-dimensional complex vector space equipped with a Hermitean scalar product* $(|)$. *Every injective (or surjective) operator $A : V \to V$ can be uniquely decomposed as a product of two operators:*

$$A = UM\,,$$

where M is strictly positive and U is unitary.

Proof. First we prove the uniqueness of the decomposition. Assume that $A = UM = U'M'$ where M, M are strictly positive and U, U' are unitary. $A^\dagger = MU^\dagger = M'U'^\dagger$, where we have used the fact that M and M' are hermitean by Lemma 4.1. Thus

$$A^\dagger A = MU^\dagger UM = M'U'^\dagger U'M'\,.$$

But $U^\dagger U = U'^\dagger U' = I$ because U and U' are unitary. As a consequence

$$M^2 = M'^2 = A^\dagger A\,.$$

The latter operator is positive: $(u|A^\dagger Au) = (Au|Au) \geq 0$ and thus $A^\dagger A = M^2 = M'^2$ admits a unique square root. In particular $\sqrt{M'^2} = \sqrt{M^2}$. As M and M' are posive, by uniqueness of square roots, they have to coincide with the $\sqrt{M'^2}$ and $\sqrt{M^2}$ respectively. Hence $M = M'$. From $UM = U'M'$ it also follows $U = U'$.
To conclude the proof, let us build up operators U and M.
We know that $A^\dagger A$ is positive, we can reinforce this property proving that it also is strictly positive. By Lemma A.1, it is sufficient to show that $0 \notin \sigma(A^\dagger A)$ or, in other words, if $A^\dagger Au = 0$, $u = 0$. let us prove that fact. If $A^\dagger Au = 0$, $(v|A^\dagger Au) = 0$ for all $v \in V$ and thus $(Av|Au) = 0$ for all $v \in V$. Since A is injective and $dimV < \infty$, A is surjective and $(Av|Au) = 0$ for all $v \in V$ can equivalently be re-written $(w|Au) = 0$ for all $w \in V$ which ientails $u = 0$. We have proven that $\sigma(A^\dagger A) \subset (0, +\infty)$, as a consequence $|A| := \sqrt{A^\dagger A}$ is strictly positive too by Theorem A.1 and Lemma A.1. Since $\sigma(|A|) \not\ni 0$, $|A|$ is injective and surjective and $|A|^{-1}$ is well-defined.
To conclude, we define

$$M := |A|$$

and

$$U := A|A|^{-1} \, .$$

By construction $UM = A$. Moreover: M is strictly positive (and thus Hermitean) by construction and U is unitary, indeed

$$UU^\dagger = (A|A|^{-1})(A|A|^{-1})^\dagger = A|A|^{-1}(|A|^{-1})^\dagger A^\dagger = A\,|A|^{-1}|A|^{-1}\,A^\dagger$$

where we have used the fact that $|A|^{-1}$ is Hermitean because $|A|$ is so. But $|A|^{-1}|A|^{-1} = (|A|^2)^{-1} = (A^\dagger A)^{-1} = A^{-1}(A^\dagger)^{-1}$ and thus

$$UU^\dagger = A\,A^{-1}(A^\dagger)^{-1}\,A^\dagger = II = I \, .$$

□