

MAURO CAPRIOLI
Politecnico di Bari

Appunti delle Lezioni

CAPITOLO 1 - GEODESIA

CAPITOLO 1 - GEODESIA	1
1 - Movimenti della terra	3
2 - Forza di gravità. Equazione del geoide.....	3
3 - Sferoide ed ellissoide	7
4 - Conoscenza del campo della gravità.....	10
4.1 - Deviazione della verticale	11
5 - Equazioni parametriche ellissoide	12
6 - Sezioni normali. Raggi di curvatura dell'ellissoide terrestre	16
7 - Le linee geodetiche	19
8 - Misure sulla superficie di riferimento.....	21
9 - Risoluzione di figure geometriche sulla superficie di riferimento	22
10 - Campo geodetico o sferico (<100 km)	23
11 - Campo topografico	24
12 - Eccesso sferico - Teorema di Legendre.....	25

1 - Movimenti della terra

La Terra è un corpo celeste che fa parte del sistema solare, la cui massa è limitata da una superficie chiusa denominata superficie fisica della terra o litosfera.

I movimenti della terra sino ad oggi constatati e misurati sono numerosi, tra questi quelli più interessanti per gli studi geodetici sono i seguenti:

rotazione intorno all' asse polare;

rivoluzione intorno al Sole.

Il primo si compie in un giorno siderale che é pari a circa 24 ore con velocità angolare pressoché costante. La Terra compie una rotazione completa nel cosiddetto giorno siderale che a causa del moto di rivoluzione intorno al sole in senso contrario a quello di rotazione è più corto del giorno solare medio di 355,91". La velocità angolare ha il valore: $\omega = 2\pi/86164.091 = 7.2921 \cdot 10^{-5} \text{ sec}^{-1}$.

Il secondo si realizza in un anno siderale pari a circa 365 giorni solari, descrivendo una orbita ellittica denominata *eclittica*. Il moto di rivoluzione si completa in un anno *siderale* pari a 365,256360 giorni solari medi.

Tra gli altri movimenti sono da ricordare:

precessione degli equinozi;

mutazione, per cui l'asse polare risulta attualmente inclinato rispetto all'eclittica di circa 23,5° valore variabile con il tempo;

migrazioni del polo terrestre;

traslazione del sistema solare;

spostamento del sistema siderale;

moto mensile della terra attorno al centro di gravità del sistema Terra-Luna.

.....

2 - Forza di gravità. Equazione del geoide

Sulla Terra esiste una direzione di immediata individuazione strumentale denominata "*verticale*". Essa é definita vettorialmente dalla "direzione della forza di gravità" nel generico punto e risulta determinabile in ogni punto in direzione, verso e modulo.

La direzione della verticale viene materializzata attraverso il "filo a piombo" per cui risulta agevole pensare che esista una eventuale superficie di riferimento ad essa collegabile.

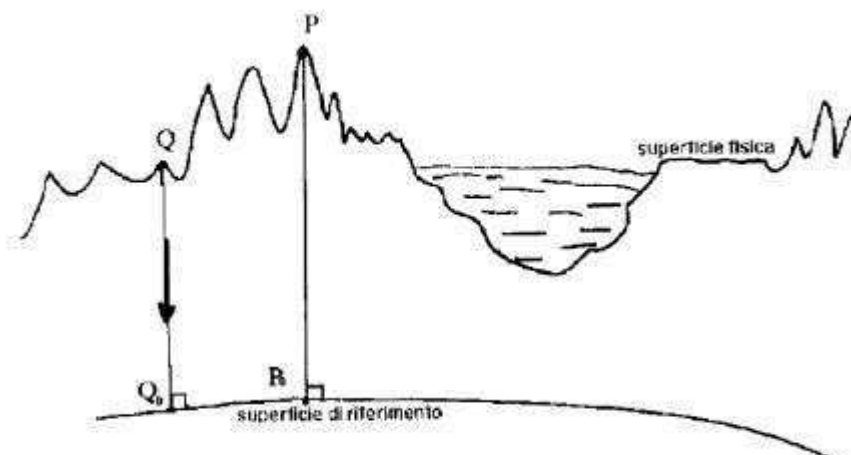


Fig. n. 2.1- Superficie fisica, superficie di riferimento

Val bene ricordare che le più evidenti asperità della superficie fisica terrestre non superano la decina di km (Monte Everest +8,846 km, Fossa delle Marianne circa -11 km), entità del tutto trascurabili rispetto alla dimensione del raggio medio terrestre $\cong 6400$ km. Si può pensare, quindi, di sostituire alla reale superficie terrestre una ipotetica superficie di riferimento, strettamente collegata con la direzione della gravità.

Scopo della geodesia è la ricerca di tale superficie di riferimento. Lo studio verrà effettuato analizzando le forze che agiscono sul generico punto della superficie terrestre.

Osserviamo che il vettore gravità \bar{g} in un punto può essere considerato somma delle diverse forze conseguenti ai diversi movimenti cui viene sottoposto lo stesso punto.

I contributi preponderanti sono forniti dalla attrazione newtoniana e, in misura minore, da quella centrifuga, mentre si può ritenere trascurabile il contributo delle restanti forze (tra le quali anche quelle dovute all'attrazione del Sole e della Luna) poiché assai piccole e variabili nel tempo.

Ricordando la legge di attrazione newtoniana: una massa m concentrata in un punto Q posta a distanza l da un punto P in cui si suppone concentrata la massa m' , esercita su quest'ultima una forza F diretta secondo la congiungente i punti P e Q pari a:

$$F = G \cdot \frac{m \cdot m'}{l^2} \quad (2.1)$$

dove G è la costante di attrazione newtoniana ($G = 6.673 \pm 0.003 \cdot 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-2}$).

Nel caso della Terra la formula suddetta non è direttamente applicabile poiché non è possibile considerare la sua massa concentrata in un punto. Se scomponiamo tale massa in tanti elementi infinitesimi dM , la forza di attrazione newtoniana che ogni elemento dM esercita sulla massa unitaria posta in P sarà data da:

$$dF = G \cdot \frac{dM \cdot 1}{l^2} \quad (2.2)$$

La somma geometrica, ossia la risultante F di tali forze elementari, rappresenta l'attrazione newtoniana esercitata dall'intera massa della Terra sulla massa unitaria concentrata nel punto P .

Le masse legate alla Terra sono sottoposte, inoltre, all'azione della forza centrifuga. Il moto rotatorio causa una accelerazione centripeta, variabile da punto a punto, espressa dalla formula

$$a = \omega^2 \cdot r \quad (2.3)$$

dove r indica la distanza del punto generico dall'asse di rotazione. Ogni elemento di massa m viene, quindi, sottoposto ad una forza centrifuga data da

$$f = m \cdot a \quad (2.4)$$

La risultante delle forze analizzate rappresenta la forza di gravità o semplicemente gravità visto che si riferisce ad una massa unitaria (fig. 2.2)

$$\bar{g} = \bar{F} + \bar{f}$$

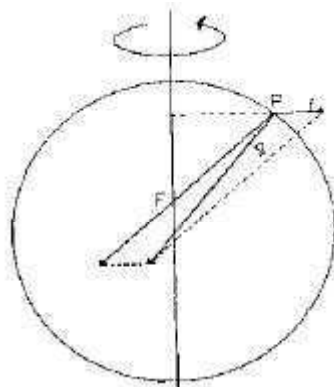


Fig. n. 2.2- Composizione della forza di gravità

Poiché per ogni punto della Terra esiste un ben determinato valore di g si può dedurre che la gravità costituisce un campo di forze che ha la caratteristica di essere conservativo (di ammettere quindi un potenziale), le cui linee di forza sono tangenti in ogni loro punto alla direzione del campo.

Consideriamo il sistema di coordinate cartesiane QXYZ avente origine in un ipotetico baricentro terrestre, l'asse Z coincidente con l'asse di rotazione e gli assi X ed Y coincidenti con quelli principali d'inerzia: dette (x,y,z) le coordinate di P e (a,b,c) quelle dell'elemento Q di massa dM , si ha (fig. 2.3)

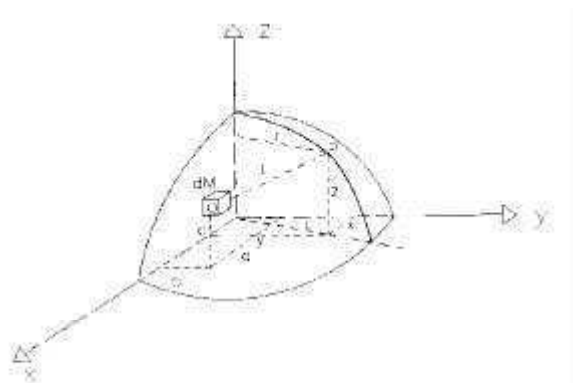


Fig. n. 2.3- Forza infinitesima di attrazione newtoniana

$$dF = -G \cdot \frac{dM}{(x-a)^2 + (y-b)^2 + (z-c)^2} \quad (2.5)$$

essendo:

$$l^2 = (x-a)^2 + (y-b)^2 + (z-c)^2;$$

il segno negativo dovuto al fatto che dF è diretta verso l'interno;

mentre per la forza centrifuga si ottiene:

$$f = \omega^2 \cdot \sqrt{x^2 + y^2} \quad (2.6)$$

essendo $r = \sqrt{x^2 + y^2}$ con evidente dipendenza di tali forze dalle coordinate di P.

Considerando che la dM è pari alla espressione

$$dM = 1 \cdot \delta \cdot da \cdot db \cdot dc \quad (2.7)$$

con δ = densità volumetrica, la (2.5) diventa:

$$dF = -G \cdot \frac{\delta \cdot da \cdot db \cdot dc}{(x-a)^2 + (y-b)^2 + (z-c)^2} \quad (2.8)$$

Essendo ben noto che le forze indicate sono conservative, si può ricavare l'espressione dei loro due potenziali, rispettivamente V e v . Ricordando che una funzione generica in $g = g(x,y,z)$ ammette potenziale quando esiste una funzione scalare $W(x,y,z)$ tale che

$$\frac{\partial W}{\partial x} = g_x \quad \frac{\partial W}{\partial y} = g_y \quad \frac{\partial W}{\partial z} = g_z \quad (2.9)$$

dove g è il gradiente di W ed è legato a quest'ultimo attraverso la relazione vettoriale differenziale

$$dW = \vec{g} \bullet d\vec{P} \quad (2.10)$$

I potenziali di dF ed f sono facilmente ricavabili:

$$dV = G \cdot \frac{dM}{\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}} \quad (2.11)$$

$$v = \frac{1}{2} \cdot \omega^2 \cdot (x^2 + y^2) \quad (2.12)$$

Il potenziale della forza di attrazione newtoniana estesa all'intero corpo terrestre è dato dall'integrale di volume (τ).

$$V = G \cdot \iiint_{\tau} \frac{\delta \cdot da \cdot db \cdot dc}{\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}} \quad (2.13)$$

Il potenziale della gravità W è dato quindi dalla somma:

$$W(xyz) = V(xyz) + v(xyz) \quad (2.14)$$

che diventa:

$$W = G \cdot \iiint_{\tau} \frac{\delta \cdot da \cdot db \cdot dc}{\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}} + \frac{1}{2} \cdot \omega^2 \cdot (x^2 + y^2) \quad (2.15)$$

Ponendo $W = \text{cost}$, si ottiene l'equazione cartesiana di una superficie equipotenziale della gravità. Per diversi valori della costante si ricava una famiglia di superfici equipotenziali. Per convenzione si considera come superficie di riferimento quella coincidente con il livello medio del mare.

Tale superficie è denominata *geoide*.

Il geoide è la superficie equipotenziale della gravità passante per il livello medio del mare.

Il geoide è meno irregolare della superficie terrestre, ed è esprimibile in forma matematica mediante l'espressione $W = \text{cost}$; la sua posizione in corrispondenza dei continenti non è direttamente misurabile; ciò costituisce, pertanto, uno dei compiti fondamentali della geodesia. Le generiche superfici equipotenziali godono delle seguenti proprietà:

- Le linee di forza della gravità (verticali) risultano normali alle superfici equipotenziali;
- Se si indica con dP uno spostamento infinitesimo si ottiene

$$dW = \vec{g} \bullet \vec{dP} \quad (2.16)$$

ossia la derivata del potenziale secondo una direzione dP fornisce la componente del vettore gravità in quella direzione. Se la direzione individuata da dP è tangente alla superficie equipotenziale per P , si ha $dW = 0$ cioè

$$\vec{g} \bullet \vec{dP} = 0$$

da cui si deduce l'ortogonalità di \vec{g} rispetto alla superficie equipotenziale.

A questo punto viene ribadita la definizione di verticale come la direzione della forza di gravità, che è in ogni punto normale alle superfici equipotenziali della gravità.

- Le superfici equipotenziali non sono tra di loro parallele

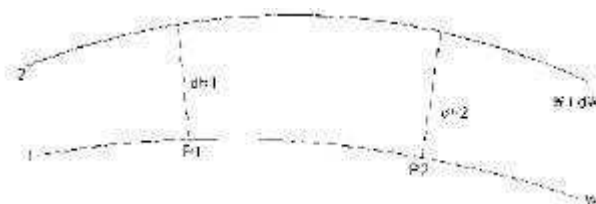


Fig. n. 2.4- Distanza tra due superfici equipotenziali

Se si analizzano due superfici molto vicine tra loro (fig. 2.4) ad esempio la prima a potenziale W ed un'altra a potenziale $W+dW$, si osserva come due punti P_1 e P_2 posti sulla prima superficie distano dall'altra rispettivamente dh_1 e dh_2 .

Essendo:

$$dW = \vec{g}_1 \bullet \vec{dh}_1 = \vec{g}_2 \bullet \vec{dh}_2$$

poiché $\vec{g}_1 \neq \vec{g}_2$ per definizione stessa della gravità, necessariamente \vec{dh}_1 risulterà diverso da \vec{dh}_2 .

Tale constatazione fa sì che la normale alle diverse superfici (cioè la verticale) non è una linea retta ma una curva, che rappresenta la già menzionata linea di forza della gravità.

La tangente in un punto alla generica superficie equipotenziale definisce la direzione dell'orizzontale (che risulta normale alla verticale).

Una piccola porzione di superficie equipotenziale viene denominata *superficie di livello* e può essere rappresentata facilmente da un liquido in quiete, contenuto in un recipiente.

Si può introdurre la definizione di *quota di un punto*, intesa come la distanza, misurata lungo la verticale, dal generico punto P sulla superficie fisica al geoide.

La determinazione di tali quote (che definiremo *ortometriche*) viene sempre riferita al geoide.

Tuttavia, essendo limitata la lunghezza del tratto compreso tra il punto sulla superficie terrestre e il geoide è possibile considerare, in prima approssimazione, tale distanza come un segmento rettilineo anziché curvo.

3 - Sferoide ed ellissoide

Si è precedentemente evidenziata l'equazione del geoide pari a:

$$W = G \cdot \iiint_{\tau} \frac{\delta \cdot da \cdot db \cdot dc}{\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}} + \frac{1}{2} \omega^2 (x^2 + y^2) = W_0 = \text{costante} \quad (3.1)$$

La determinazione analitica di tale equazione e delle superfici di livello in genere, presuppone la risoluzione dell'integrale triplo che non solo presenta ovvie difficoltà di calcolo, ma richiede anche la conoscenza della densità δ per ogni elemento di volume.

La densità della terra, inoltre, non è costante ma aumenta con la profondità con una legge di distribuzione non ancora ben nota ($\delta = 2.7 \text{ gr cm}^{-3}$ per la crosta terrestre, mentre ha valore medio $\delta_m = 5.52 \text{ gr cm}^{-3}$).

Queste difficoltà comportano l'impossibilità di scrivere in maniera rigorosa l'equazione del geoide.

Tuttavia è possibile, dopo aver eseguito una serie di passaggi¹, scrivere l'espressione del potenziale gravitazionale come di seguito riportata:

¹ Consideriamo il generico punto P sulla superficie terrestre e il punto Q in cui si suppone concentrata la massa dM



Fig. n. 3.1 - Triangolo OPQ del Teorema di Carnot **Fig. n. 3.2** - Coordinate cartesiane e polari

La distanza l tra P e Q può essere espressa per mezzo di due raggi vettori σ e σ' e dell'angolo ϑ da esso compresi. Applicando il teorema di Carnot al triangolo OPQ (fig. 3.1) si ottiene

$$l^2 = \sigma^2 + \sigma'^2 - 2\sigma\sigma' \cos \vartheta = \sigma^2 \left(1 + \left(\frac{\sigma'}{\sigma} \right)^2 - 2 \frac{\sigma'}{\sigma} \cos \vartheta \right) \quad (3.2)$$

esprimendo ancora la distanza l in funzione delle coordinate polari (fig. 3.2) ossia

$$x = \sigma \cdot \cos \psi \cdot \cos \lambda \quad y = \sigma \cdot \cos \psi \cdot \sin \lambda \quad z = \sigma \cdot \sin \psi \quad (3.3)$$

si ottiene:

$$l^2 = \sigma^2 + \sigma'^2 - 2\sigma \cdot \sigma' \cdot [\sin \psi \cdot \sin \psi' + \cos \psi \cdot \cos \psi' \cdot \cos(\lambda - \lambda')] \quad (3.4)$$

da cui per confronto della l determinata con il teorema di Carnot e quella in coordinate polari si determina il valore di $\cos \vartheta$ in funzione delle coordinate polari di P e Q:

$$\cos \vartheta = [\sin \psi \cdot \sin \psi' + \cos \psi \cdot \cos \psi' \cdot \cos(\lambda - \lambda')] \quad (3.5)$$

L'elemento di volume è esprimibile attraverso

$$da \cdot db \cdot dc = \sigma'^2 \cdot \cos \psi' \cdot d\psi' \cdot d\lambda' \cdot d\sigma' \quad (3.6)$$

Sostituendo tutti i parametri trovati nell'equazione del potenziale della forza di attrazione newtoniana si ha

$$V = \frac{G}{\sigma} \cdot \iiint_{\tau} \frac{\delta \cdot \sigma'^2 \cos \psi' d\psi' d\lambda' d\sigma'}{\sqrt{1 + \left(\frac{\sigma'}{\sigma} \right)^2 - \left(\frac{\sigma'}{\sigma} \right) \cos \vartheta}}$$

$$U = \frac{GM}{\sigma} \left[1 + \frac{1}{2\sigma^2} \left(\frac{C-A}{M} \right) (1 - 3\sin^2\psi) \right] + \frac{1}{2} \omega^2 \sigma^2 \cos^2\psi \quad (3.7)$$

Ponendo $U=U_0$ = costante si ottiene l'equazione di una superficie di rotazione chiamata *sferoide*.

Una analisi della sua espressione pone in evidenza come il primo termine esprima il potenziale di una massa M concentrata in un punto ed uniformemente distribuita su una superficie di raggio σ ; il secondo termine contiene la differenza tra i momenti d'inerzia equatoriale e polare ed è quindi legato allo schiacciamento terrestre. Tali termini non richiedono la conoscenza della distribuzione interna delle masse. I termini trascurati, inoltre, sono dell'ordine di qualche unità di 10^{-3} , per cui lo sferoide rappresenta una buona approssimazione del geoide. Eseguendo un certo numero di derivazioni e di sviluppi in serie si giunge all'espressione finale dello sferoide in coordinate geocentriche polari (fig. 3.2), passante per un punto prefissato del livello medio del mare che risulta essere data da:

$$\sigma = a \cdot (1 - \alpha \cdot \sin^2\psi) \quad (3.8)$$

dove:

a = semiasse equatoriale

c = semiasse polare

α = schiacciamento = $(a - c)/a = 1 - c/a$

ψ = latitudine geocentrica

Lo sferoide è una superficie di rotazione molto simile ad un ellissoide purché presenti gli stessi valori di a e di c . Infatti partendo dalla generica equazione dell'ellissoide

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1 \quad (3.9)$$

si arriva all'espressione dell'ellissoide data da (3.8)²

V rappresenta una funzione armonica nello spazio esterno alla terra ed è quindi sviluppabile in una serie convergente. Con facili passaggi si giunge all'espressione (3.7).

² ricordando la definizione stessa di α , si ha

$$c = a(1 - \alpha) \Rightarrow c^2 = a^2(1 - \alpha)^2 \cong a^2(1 - 2\alpha)$$

in cui si è trascurato il termine α^2 (infinitesimo di ordine superiore). Sostituendo nella equazione dell'ellissoide l'espressione di c si ha

$$x^2 + y^2 + z^2(1 - 2\alpha)^{-1} = a^2$$

Sviluppando $(1 - 2\alpha)^{-1}$ in serie binomiale

$$(1 - x)^n = 1 + nx + (nx)^2/2! + (nx)^3/3! + \dots$$

arrestandola al primo termine e ricordando che $\sigma^2 = x^2 + y^2 + z^2$ si ha

$$\sigma^2 = a^2 \left(1 - 2\alpha \frac{z^2}{a^2} \right) \quad (3.10)$$

$$\sigma = (1 - \alpha \cdot \sin^2 \psi)$$

che coincide con l'equazione dello sferoide avente gli stessi semiassi, a meno delle approssimazioni eseguite.

4 - Conoscenza del campo della gravità

Non devono meravigliare le semplificazioni introdotte per passare ad una superficie di riferimento più semplice (ellissoide) sostitutiva del geoide.

Lo scopo, infatti, è quello di determinare una superficie utilizzabile per i rilievi topografici, non di determinare il geoide.

Nella ricerca applicata, tuttavia, la posizione del geoide viene determinata misurando gli scostamenti dall'ellissoide.

Il problema della determinazione del campo della gravità terrestre può essere semplificato se si tiene conto che la Terra, con buona approssimazione può essere considerata un'ellissoide di rivoluzione. Con tale ipotesi è possibile suddividere il campo gravitazionale terrestre in una parte detta di *riferimento normale* (corrispondente al campo gravitazionale di un ellissoide di massa M pari a quella della Terra) ed una parte di *disturbo* che rappresenta la deviazione del campo gravitazionale reale da quello normale.

Pertanto si ha

$$W(P) = U(P) + T(P)$$

dove:

$W(P)$ = Potenziale effettivo

$U(P)$ = Potenziale normale

$T(P)$ = Potenziale anormale

Le superfici di livello, cioè quelle a potenziale costante, sono ellissoidi per il campo normale e superfici più complesse per il campo effettivo. Pertanto le prime sono esprimibili in forma matematica semplice, mentre le altre hanno forme analitiche difficilmente risolubili.

In particolare avendo definito *geoide* la superficie equipotenziale avente $W=W_0$ passante per il livello medio del mare, si chiamerà ellissoide di riferimento quella superficie equipotenziale del campo normale avente $U_0=W_0=\text{costante}$.

Poiché σ^2 differisce da a^2 per quantità dell'ordine di α , si può sostituire al rapporto (z^2/a^2) il rapporto (z^2/σ^2) per il quale vale l'uguaglianza $(z^2/a^2) = \sin^2 \psi$, per cui si ottiene:

$$\sigma = a\sqrt{1 - \sin^2 \psi} \quad (3.11)$$

e quindi, con il consueto sviluppo binomiale, si ottiene la formula (3.8)

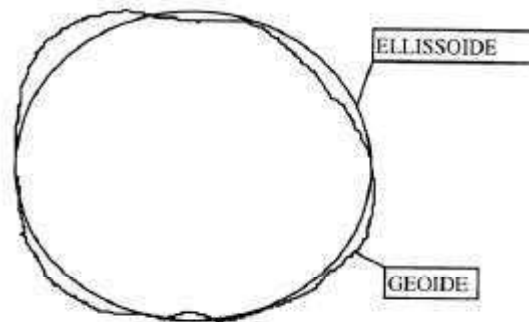


Fig. n. 4.1- Posizione relativa tra geoide ed ellissoide

La conoscenza del potenziale gravitazionale terrestre implica la determinazione del potenziale anomalo; essa è fondamentale per la comprensione di diversi fenomeni geofisici e geodinamici, quali per esempio:

- la deriva dei continenti, dovuta principalmente ai moti di convezione del “mantello”, che sono causa degli eventi sismici e vulcanici;
- la dinamica delle correnti degli oceani, causate dalle variazioni di potenziale sulla superficie del mare, che ha conseguenze meteorologiche su tutto il globo;
- la determinazione delle orbite dei satelliti artificiali, notevolmente influenzate dalla presenza dello scostamento del campo gravitazionale terrestre rispetto a quello normale.

La conoscenza del geoide risulta fondamentale anche per applicazioni di ambito strettamente topografico, quali ad esempio:

lo scostamento della *verticale* (normale al geoide) rispetto alla normale all’ellissoide, denominato *deviazione della verticale*;

la definizione del riferimento per le quote;

la trasformazione dei dislivelli GPS in differenze di altezze ortometriche;

tutti argomenti che saranno trattati in dettaglio successivamente.

4.1 - Deviazione della verticale

A causa delle ondulazioni del geoide la sua normale (cioè la verticale) non coincide in ogni punto con la normale all’ellissoide. L’angolo formato dalla verticale e dalla normale in un punto viene denominato deviazione della verticale. Il suo valore numerico è di qualche decina di secondi sessagesimali quindi non trascurabile per misure di elevatissima precisione. Nasce quindi la necessità di riferirsi ad un doppio sistema di coordinate geografiche uno relativo all’ellissoide che chiameremo di coordinate geografiche ellissoidiche e l’altro relativo al geoide che chiameremo di coordinate geografiche astronomiche o geoidiche. Consideriamo l’ellissoide di riferimento e su esso le coordinate geografiche ellissoidiche e geoidiche di un punto P; indichiamo con φ_l, λ_l le prime, con φ_a, λ_a le seconde e con ε la deviazione della verticale. Relativamente alla stesso punto sulla superficie terrestre esistono delle differenze

$$\Delta\varphi = \varphi_a - \varphi_l \quad \Delta\lambda = \lambda_a - \lambda_l$$

La deviazione della verticale ε in un punto P risulta parallela al piano tangente all’ellissoide nel punto stesso.

Si può scomporre ε in due componenti. Una ε_m nel piano meridiano, ed una ε_v , nel piano del parallelo.

Siano inoltre N e V rispettivamente la normale all’ellissoide e la verticale.

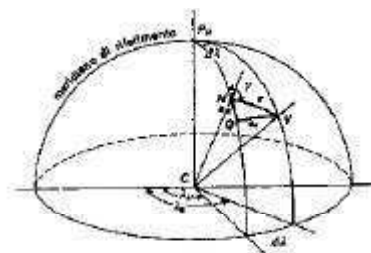


Fig. n. 4.1.1- Deviazione della verticale riferita alle coordinate geografiche

Tra le due componenti di deviazione e le coordinate geografiche esistono le seguenti relazioni

$$\varepsilon_m \cong \varphi_a - \varphi_l \quad (4.1)$$

$$\varepsilon_p \cong (\lambda_a - \lambda_l) \cos \varphi_a \quad (4.2)$$

Gli scostamenti tra geoide ed ellissoide possono essere determinati sia con misure gravimetriche, ossia del valore locale dell'accelerazione di gravità, sia con lo studio delle orbite dei satelliti artificiali. La conoscenza del suddetto scostamento è importante non solo per la determinazione dell'andamento del geoide ma anche per la denuncia, per ogni variazione anomala della verticale e quindi del geoide della presenza di masse perturbanti all'interno della superficie terrestre, per esempio la presenza di liquidi (acqua, petrolio, ecc.) o di gas.

5 - Equazioni parametriche ellissoide

La irregolarità del geoide, di cui si è accennato, non permette la esecuzione di misure e calcoli su tale superficie, pur avendo lo stesso riferimento fisico (la verticale) della reale superficie terrestre. Conviene introdurre, quindi, l'ellissoide terrestre che permette di determinare le coordinate delle proiezioni su di esso dei punti del terreno, una volta che siano state misurate le grandezze (distanze e angoli) necessarie a definire le posizioni reciproche dei punti. Per poter individuare le formule necessarie per i calcoli occorre studiare il comportamento locale dell'ellissoide.

Riferiamo l'ellissoide di rotazione ad un sistema di assi cartesiani ortogonali con l'origine nel suo centro, l'asse Z coincidente con l'asse di rotazione e diretto positivamente verso il polo Nord e gli assi X ed Y contenuti nel piano equatoriale (fig. 5.1)

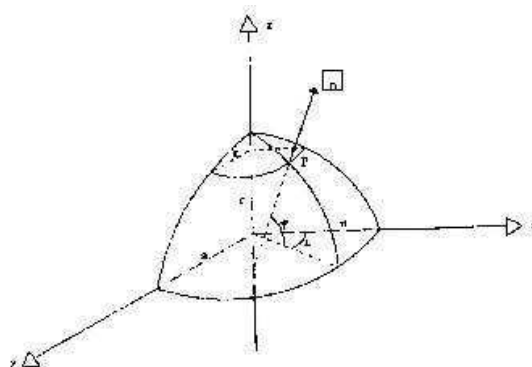


Fig. n. 5.1- Ellissoide di rotazione riferito agli assi X,Y,Z

L'equazione cartesiana dell'ellissoide è data da

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1 \quad (5.1)$$

L'uso delle coordinate cartesiane tuttavia non è conveniente per risolvere figure tracciate su di una superficie; si ricorre, quindi, a rappresentazioni di tipo parametrico introducendo un sistema di coordinate curvilinee. Si definisce:

- latitudine φ

l'angolo che la normale alla superficie dell'ellissoide passante per P forma con il piano equatoriale x,y. La latitudine è misurata positivamente verso il polo nord e negativamente verso il polo sud, varia quindi da +90° a -90°;

- longitudine λ

l'angolo che il piano contenente l'asse polare di rotazione e la normale in P alla superficie (definito piano meridiano), forma con il piano meridiano assunto come riferimento (quello passante per GREENWICH).

Le linee $\varphi = \text{costante}$ della superficie si chiamano *paralleli*, quelle $\lambda = \text{costante}$ si chiamano *meridiani*.

Consideriamo l'ellisse meridiana contenuta in un generico piano, la cui intersezione con il piano equatoriale fornisce l'asse delle ascisse, che indichiamo con r (fig. 5.2)

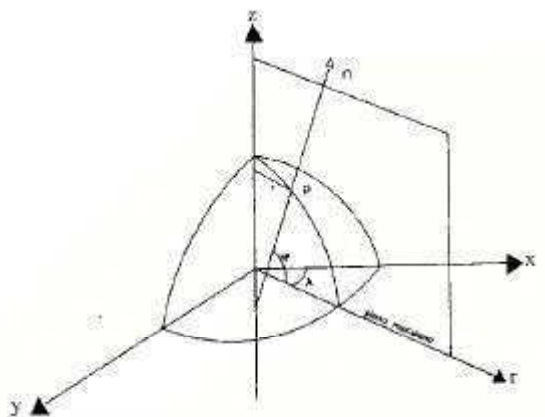


Fig. n. 5.2- Ellisse meridiana

E' facile verificare che essendo $r^2 = x^2 + y^2$ l'equazione dell'ellisse così definita può essere scritta

$$\frac{r^2}{a^2} + \frac{z^2}{c^2} = 1 \quad (5.2)$$

I coseni direttori della normale n sono espressi dalle derivate parziali della funzione rispetto ad r ed a z

$$\cos \varphi = \frac{\partial f}{\partial r} = K \cdot 2 \frac{r}{a^2}$$

$$\cos\left(\frac{\pi}{2} - \varphi\right) = \sin \varphi = \frac{\partial f}{\partial z} = K \cdot 2 \frac{z}{c^2}$$

Attraverso una serie di passaggi si giunge alle equazioni (5.12)³:

³ Facendo il rapporto tra i due coseni direttori si ottiene:

$$\frac{\sin \varphi}{\cos \varphi} = \operatorname{tg} \varphi = \frac{z \cdot a^2}{r \cdot c^2} \quad (5.3)$$

si ha quindi

$$z = r \cdot \operatorname{tg} \varphi \frac{c^2}{a^2}$$

Il termine c^2/a^2 è possibile esprimerlo attraverso un parametro chiamato eccentricità “e”

$$e^2 = \frac{a^2 - c^2}{a^2} = 1 - \frac{c^2}{a^2} \Rightarrow \frac{c^2}{a^2} = 1 - e^2 \quad (5.5)$$

sostituendo l'espressione trovata nella (5.4) si ha

$$\operatorname{tg} \varphi = \frac{z}{r} \cdot \frac{1}{1 - e^2} \Rightarrow z = r \cdot \operatorname{tg} \varphi (1 - e^2)$$

sostituendo l'espressione di z così determinato nella (5.2) si ottiene una equazione nell'unica incognita r

$$\frac{r^2}{a^2} + \frac{r^2 \operatorname{tg}^2 \varphi \cdot (1 - e^2)^2}{c^2} = 1 \quad (5.6)$$

moltiplicando la (5.6) per a^2 si ottiene

$$r^2 + \frac{r^2 \operatorname{tg}^2 \varphi \cdot (1 - e^2)^2}{c^2 / a^2} = a^2$$

sostituendo l'espressione (5.5) di c^2/a^2 si ha

$$r^2 + \frac{r^2 \operatorname{tg}^2 \varphi \cdot (1 - e^2)^2}{1 - e^2} = a^2 \Rightarrow r^2 \operatorname{tg}^2 \varphi (1 - e^2) = a^2$$

ed ancora moltiplicando tutto per $\cos^2 \varphi$ si ha

$$r^2 \cos^2 \varphi + r^2 \sin \varphi \cdot (1 - e^2) = a^2 \cos^2 \varphi$$

$$r^2 \cos^2 \varphi + r^2 \sin \varphi - r^2 e^2 \sin^2 \varphi = a^2 \cos^2 \varphi$$

Essendo $(\cos^2 \varphi + \sin^2 \varphi) = 1$ si può scrivere

$$r^2 - r^2 e^2 \sin \varphi = a^2 \cos^2 \varphi$$

$$r^2 (1 - e^2 \sin^2 \varphi) = a^2 \cos^2 \varphi$$

$$r^2 = \frac{a^2 \cos^2 \varphi}{(1 - e^2 \sin^2 \varphi)} \Rightarrow r = \frac{a \cos \varphi}{\sqrt{(1 - e^2 \sin^2 \varphi)}} \quad (5.7)$$

$$r = \frac{a \cdot \cos \varphi}{\sqrt{1 - e^2 \sin^2 \varphi}} \quad (5.8)$$

Indicando $W = \sqrt{1 - e^2 \sin^2 \varphi}$ la (5.8) diventa

$$r = \frac{a \cdot \cos \varphi}{W} \quad (5.9)$$

sostituendo la (5.8) nella (5.4) si ottiene l'espressione di z

$$z = \frac{a \cdot \cos \varphi \cdot \sin \varphi}{W \cdot \cos \varphi} (1 - e^2) = \frac{a \cdot \sin \varphi (1 - e^2)}{W} \quad (5.10)$$

La (5.8) e la (5.10) sono le equazioni parametriche dell'ellisse meridiana dell'ellissoide di rotazione; per avere le equazioni parametriche dell'ellissoide basta osservare che con riferimento al sistema di coordinate cartesiane (OXYZ) una qualsiasi ellisse meridiana può essere individuata dall'angolo λ , che rappresenta la longitudine; è possibile allora scrivere

$$x = r \cos \lambda \qquad y = r \sin \lambda \quad (5.11)$$

Sostituendo nelle (5.11) l'espressione di r si ottengono le equazioni parametriche dell'ellissoide

$$\begin{cases} x = \frac{a \cos \varphi \cos \lambda}{W} \\ y = \frac{a \cos \varphi \sin \lambda}{W} \\ z = \frac{a(1 - e^2) \sin \varphi}{W} \end{cases} \quad (5.12)$$

Le (5.12) e le loro inverse sono le formule che vengono utilizzate dal sistema GPS per determinare le coordinate di un generico punto (x,y,z) sull'ellissoide WGS-84 e ricavare, quindi, le coordinate φ e λ .

Per tale determinazione occorre conoscere i valori di

a = semiasse equatoriale

e = eccentricità dell'ellissoide.

Si impone, quindi, la scelta di un ellissoide.

Sin dall'antichità l'ipotesi della Terra sferica ha provocato diverse e contrastanti interpretazioni, che si basavano su considerazioni di carattere filosofico.

Gli Egizi, convinti della sfericità della Terra, furono i primi a trattare questi problemi; nella Piramide di Cheope, infatti, il perimetro della base è pari a 1:120 del grado di meridiano egizio, cioè a 931.22 m; l'altezza corrisponde a 148.21 m, con un rapporto fra base e altezza pari a 2π .

Lo studio della determinazione del raggio terrestre (ovvero del semiasse equatoriale) venne affrontato fin dal 200 a.C. da Eratostene, matematico e geografo alessandrino, che effettuò la prima misura del raggio ottenendo un valore molto vicino a quello oggi noto.

Circa un millennio dopo gli arabi per ordine del Califfo Almanun giunsero ad una misura approssimata della circonferenza (ovvero dell'ellisse meridiana) di circa quarantamila km.

Dopo sette secoli, nel 1525, l'astronomo francese Fernel misurò la differenza di latitudine fra Parigi e Amiens, nonché la distanza fra le due città, mediante un contatore applicato alla ruota di una carrozza, ottenendo un valore della lunghezza del meridiano di circa 40.063 km.

Solo dopo la scoperta della legge di gravitazione di Newton si pensò a una forma della terra diversa dalla quella sferica, perché l'ellissoide schiacciato ai poli sembrò essere il solido più attendibile per la maggior parte degli studi dell'epoca.

Nel 1841 furono pubblicati i parametri dell'ellissoide calcolati da Bessel e successivamente si ottennero altri valori definiti attraverso i parametri a ed α , dove a rappresenta il semiasse equatoriale ed α lo schiacciamento (ossia $\alpha = (a - c)/a$).

Si elencano di seguito gli ellissoidi più importanti, perché largamente impiegati, ciascuno dei quali indicato con il nome dello scienziato che ne ha determinato i parametri, mentre Canberra fa riferimento alla località sede nel 1980 del congresso della Associazione Internazionale di Geodesia:

ELLISSOIDI	a	α	
BESSEL (1841)	6.337,397	1:299	<u>Sistema Cassini-Soldner</u>
CLARKE (1880)	6.378,243	1:293,5	
HELMERT (1906)	6.378,140	1:298,3	
HAYFORD (1909)	6.378,388	1:297,0	<u>Sistema Gauss-Boaga</u>
KRASSOVSKY (1942)	6.378,245	1:298,3	
CANBERRA (1980)	6.378,137	1:298,257	

Mentre fino al 1960 la determinazione di questi parametri è stata effettuata attraverso la misura di archi di meridiani, oggi si ricorre, invece, ad osservazioni satellitari che forniscono parametri meccanici quali GM (costante di gravitazione universale per la massa della Terra), J_2 , (momento d'inerzia), ω (velocità angolare) da cui si ricavano le costanti geometriche a, e^2, α .

Attraverso queste nuove metodologie è stato determinato l'ellissoide WGS-84 (World Geodetic System 1984) oggi usato nelle determinazioni satellitari GPS che, come vedremo in seguito, consentono la determinazione di φ e λ di ogni punto della superficie terrestre con approssimazione diversa a seconda dello strumento utilizzato. I parametri meccanici di questo ellissoide sono:

$$\begin{cases} GM = 3986005 \cdot 10^{-8} m^3 s^{-2} \\ J_2 = 108263 \cdot 10^{-8} \dots\dots\dots \\ \omega = 7292155 \cdot 10^{-11} rad \text{ sec}^{-1} \end{cases}$$

da cui si ricavano le costanti geometriche

$$\begin{cases} a = 6378137 m \dots\dots\dots \\ e^2 = 0.0066438002 \dots\dots\dots \\ \alpha = 1/298.257223563 \end{cases}$$

Si noti che nel WGS84 come anche nei più recenti sistemi, l'origine del sistema cartesiano ortogonale è il centro di massa del pianeta, mentre l'asse delle Z va fatto coincidere convenzionalmente con l'asse di rotazione. Ciò perché l'asse di rotazione si muove rispetto alla massa terrestre, così come il polo Nord, descrivendo la curva irregolare "polodia".

6 - Sezioni normali. Raggi di curvatura dell'ellissoide terrestre

Per studiare il comportamento dell'ellissoide nell'intorno di un punto è necessario conoscere i raggi di curvatura nel punto stesso.

Il raggio di una curva qualsiasi, appartenente ad una generica superficie, può essere determinato qualora siano noti i raggi di curvatura delle due sezioni normali principali alla superficie nel punto stesso.

E' opportuno a tale scopo ricercare i raggi principali di curvatura dell'ellissoide in un suo punto generico.

Si consideri il punto P e la normale all'ellissoide in tale punto n; il fascio di piani aventi per costola n interseca l'ellissoide secondo curve piane chiamate *sezioni normali*; ciascuna di queste sezioni ha nel punto P un determinato raggio di curvatura (nel caso di una superficie sferica invece tutte le sezioni normali sono circonferenze aventi lo stesso raggio della sfera).

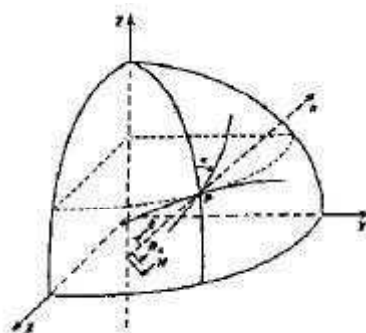


Fig. n. 6.1 - Sezioni normali in un punto dell'ellissoide

I raggi di curvatura delle sezioni normali in un punto dell'ellissoide variano con continuità da un valore minimo ρ , raggio del *Meridiano*, ad un valore massimo N, raggio della *Gran Normale*.

Le sezioni che hanno rispettivamente il massimo e il minimo raggio di curvatura sono chiamate sezioni normali principali ed i loro raggi sono denominati *raggi principali di curvatura*.

Questi due valori estremi sono situati in due piani fra loro ortogonali, che definiscono due direzioni principali, individuanti rispettivamente la direzione del piano della sezione meridiana per il punto P, e quella del piano ad esso normale nello stesso punto P, che realizza a sua volta la sezione denominata sezione primo verticale.

Per una sezione normale generica, che forma un angolo α (denominato azimuth) con il meridiano, il raggio di curvatura R_α può essere espresso in funzione del raggio minimo ρ del meridiano e massimo N della Gran Normale, attraverso l'applicazione del Teorema di Eulero, che determina:

$$\frac{1}{R_\alpha} = \frac{\cos^2 \alpha}{\rho} + \frac{\sin^2 \alpha}{N} \quad (6.1)$$

Se consideriamo un arco ds del meridiano, il cui angolo corrispondente è pari a $d\varphi$ (visto che il punto P è individuato su di esso dalla latitudine φ) è possibile scrivere (fig. 6.2)

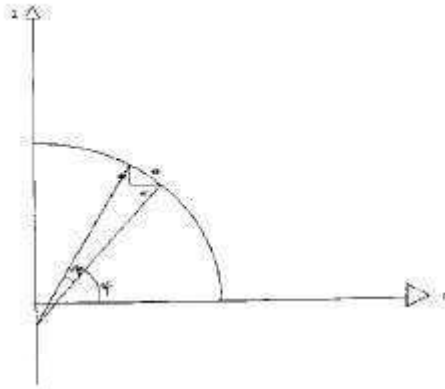


Fig. n. 6.2- Arco di ellisse meridiana

$$\rho d\varphi = ds \quad (6.2)$$

da cui successivamente ricordando le equazioni parametriche di r e z si ricava⁴

⁴ E' lecito ipotizzare l'elemento ds pari all'ipotenusa del triangolo rettangolo formato da dz e dr, per cui

$$ds^2 = dz^2 + dr^2 \Rightarrow ds = \sqrt{dz^2 + dr^2}$$

sostituendo nella (6.2) il valore di ds così trovato si ottiene

$$\rho = \frac{\sqrt{dz^2 + dr^2}}{d\varphi} \quad (6.3)$$

Si considerano ora le espressioni parametriche di r e z calcolate precedentemente

$$r = \frac{a \cdot \cos \varphi}{W} \quad z = \frac{a \cdot \sin \varphi}{W} \cdot (1 - e^2) \quad (6.4)$$

e si calcolano le derivate rispetto a φ , ottenendo

$$\frac{dr}{d\varphi} = \frac{a \cdot (1 - e^2) \cdot \sin \varphi}{W^2} \quad \frac{dz}{d\varphi} = \frac{a \cdot (1 - e^2) \cdot \cos \varphi}{W^3} \quad (6.5)$$

sostituendo le (6.5) nella (6.3) si ottiene

$$\rho = \frac{\sqrt{dz^2 + dr^2}}{d\varphi} = \frac{a \cdot (1 - e^2)}{W^3}$$

$$\rho = \frac{a \cdot (1 - e^2)}{W^3}$$

Il teorema di Meusnier dimostra che se si considera una generica sezione obliqua di una superficie il suo raggio di curvatura è uguale al raggio di curvatura della sezione normale che, nel punto considerato, ha con la sezione obliqua la tangente in comune, moltiplicato per il coseno dell'angolo formato dai due piani.

Nel caso in esame la sezione principale di raggio N contiene la tangente al parallelo di raggio r e forma con il piano di quest'ultimo un angolo uguale alla latitudine del punto, si ha quindi

$$r = N \cos \varphi$$

da cui

$$N = \frac{r}{\cos \varphi} = \frac{a \cos \varphi}{W} \cdot \frac{1}{\cos \varphi} = \frac{a}{W} = \frac{a}{\sqrt{1 - e^2 \cdot \sin^2 \varphi}}$$

Dalle formule precedenti si può ricavare la differenza fra N e ρ

$$\frac{N - \rho}{N} = \frac{e^2 \cos^2 \varphi}{1 - e^2 \sin^2 \varphi} \quad (6.6)$$

Ciò permette di verificare che N è sempre maggiore o uguale a ρ ed inoltre che la differenza massima tra i due raggi si riscontra all'equatore, ossia per $\varphi = 0$, mentre è nulla ai poli, ossia per $\varphi = \pi/2$.

Si noti che la differenza relativa (6.6) fra N e ρ è comunque dell'ordine di e^2 ossia è pari a circa 1/150, cioè dell'ordine di 10^{-2} .

Tenuto conto che il generico raggio di curvatura, di una qualunque sezione normale, ha sempre un valore intermedio tra quello di N e quello di ρ si può dedurre che l'ellissoide terrestre nell'intorno di un punto si discosta poco da una sfera avente come raggio R la media geometrica del raggio di curvatura minimo e massimo cioè:

$$R = \sqrt{\rho \cdot N} = \frac{a \sqrt{1 - e^2}}{1 - e^2 \sin^2 \varphi}$$

La sfera di raggio R , tangente all'ellissoide nel punto in cui il raggio stesso è calcolato, si chiama sfera locale. Il suo raggio, infatti, è funzione di ρ , quindi varia da punto a punto. Poiché questa sfera differisce poco dall'ellissoide, in un largo intorno del punto, può essere facilmente utilizzata per molti calcoli in sostituzione dell'ellissoide.

7 - Le linee geodetiche

Due punti generici, disposti sulla superficie dell'ellissoide, possono essere collegati attraverso una infinità di linee. *La linea geodetica è rappresentata dalla linea di lunghezza minima.*

La *linea geodetica* o più semplicemente *geodetica* su di una superficie gode della proprietà di avere in ogni suo punto la normale principale coincidente con la normale alla superficie.

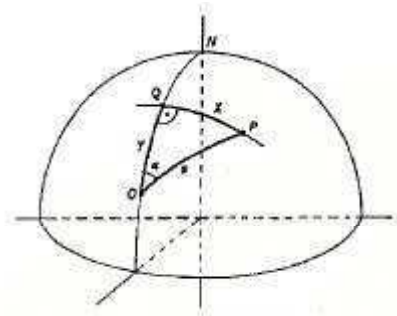


Fig. n. 7.1 - - Terna Euleriana

Possiamo definire (fig. 7.1):

- *coordinate geodetiche polari* (s, α) la lunghezza s dell'arco di geodetica OP e l'angolo α formato dalla tangente alla geodetica e dalla tangente al meridiano. Tale angolo prende il nome di azimut della geodetica in O
- *coordinate geodetiche rettangolari* (X, Y) considerando la geodetica passante per P, normale al meridiano per O che interseca quest'ultimo in Q; la coordinata X è pari alla lunghezza dell'arco QP e la coordinata Y è uguale alla lunghezza dell'arco di meridiano OQ.

Si possono determinare le relazioni che permettono di calcolare X ed Y noti s ed α e viceversa.

Sull'ellissoide il meridiano è una linea geodetica, pertanto il triangolo OPQ è un triangolo i cui lati sono tutte linee geodetiche, è quindi un triangolo geodetico.

L'elemento lineare sull'ellissoide, espresso in funzione delle coordinate geografiche è (fig. 7.2):

$$ds^2 = \rho^2 d\varphi^2 + N^2 \cos^2 \varphi d\lambda^2 \quad (7.1)$$

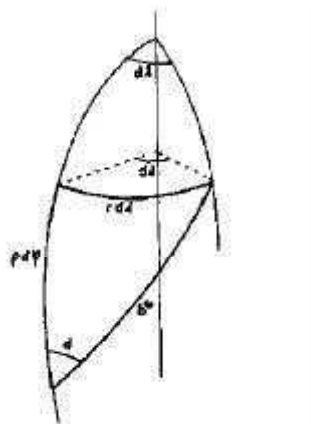


Fig. n. 7.2- Elemento lineare sull'ellissoide

La lunghezza di un arco di linea qualunque viene allora espressa dall'integrale:

$$s = \int_{\varphi_1}^{\varphi_2} \sqrt{\rho^2 + N^2 \cos^2 \varphi \left(\frac{d\lambda}{d\varphi} \right)^2} d\varphi \quad (7.2)$$

La geodetica è quella linea di equazione $s = f(\varphi)$ tale da rendere minimo il valore dell'integrale.

Le geodetiche di una superficie sono (come le rette in un piano) una doppia infinità, tuttavia per due punti ne passa una ed una sola, come una ed una sola è quella passante per un punto ed avente un azimut assegnato.

In generale le geodetiche sono delle curve gobbe, mentre fanno eccezione i meridiani e l'equatore che sono invece delle curve piane.

8 - Misure sulla superficie di riferimento

La realizzazione di un rilievo comporta l'esecuzione di misure di angoli e distanze sulla reale superficie fisica, mentre i calcoli relativi sono svolti utilizzando una prestabilita superficie di riferimento, diversa a seconda dell'ampiezza della zona da rilevare.

Le operazioni di misura relative alle quote sono, invece, sempre riferite al geoide mediante la verticale, poiché tale riferimento è facilmente riproducibile con i vari strumenti (come si vedrà di seguito).

Se si considerano su di una sfera due punti A e B ed, inoltre, il piano normale in A che passa per B, tale piano intersecherà la sfera secondo un arco di cerchio coincidente con quello formato dall'intersezione della sfera con il piano normale in B passante per A. Le due normali in A e B, infatti, sono convergenti nel centro della sfera, i due piani coincidono e, quindi, per due punti della sfera passa una ed una sola sezione normale.

Per l'ellissoide ciò non avviene; volendo collegare due punti appartenenti a tale superficie non esiste un'unica sezione normale che li contenga (fig. 8.1).

Si osserva, infatti, che la normale in A incontra l'asse Z in C_A mentre la normale in B lo incontra in C_B e le due normali non sono in genere complanari. Pertanto il piano che contiene la normale in A ed il punto B intersecherà l'ellissoide secondo una certa curva (s') *sezione normale* (\overline{AB}) ; se invece si considera il piano contenente la normale in B e il punto A, questa intersecherà l'ellissoide secondo un'altra curva (s'') che si chiama *sezione normale reciproca* (\overline{BA}) .

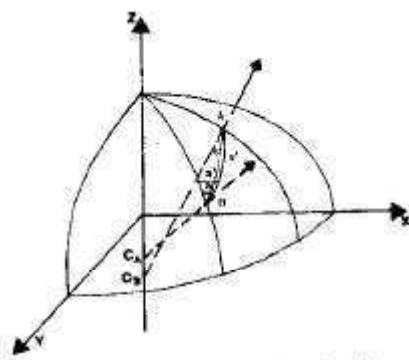


Fig. n. 8.1- Sezioni normali reciproche

Gli scostamenti tra le due sezioni normali reciproche sono funzioni crescenti dell'eccentricità dell'ellisse meridiana; si può dire che più ci si allontana dalla forma sferica più lo sdoppiamento delle sezioni normali diviene sensibile.

Se consideriamo due punti sull'ellissoide A e B, chiamando con s' l'arco di una delle due sezioni normali che li congiungono avente azimuth α' , e con l'arco s di geodetica avente azimuth α , a meno di termini dell'ordine di $(s/N)^8$, ovvero assolutamente trascurabili per archi lunghi centinaia di chilometri, si dimostra che

$$\frac{s'-s}{s} = \frac{1}{360} \cdot \frac{s^4}{N^2 R_\alpha^2} \left(\frac{e^2}{1-e^2} \right)^2 \sin^2 2\alpha \cos^4 \varphi \quad (8.1)$$

Se si ipotizza $s = 1000$ Km sviluppando i calcoli, tale rapporto risulta essere dell'ordine del centimetro (ossia 10^{-8} della distanza considerata). Tenendo presente, invece, che nella misura delle distanze si può raggiungere una precisione relativa di 10^{-6} cioè (± 1 mm/1 km), si può concludere che è *perfettamente giustificato ritenere che misure di distanze eseguite secondo sezioni normali diano gli stessi risultati di misure eseguite secondo archi di geodetiche*.

La differenza dei due azimuth, espressa in radianti è pari a

$$\alpha' - \alpha = \frac{s^2}{12NR_\alpha} \cdot \frac{e^2}{1-e^2} \sin 2\alpha \cdot \cos^2 \varphi \quad (8.2)$$

che, a parità di lunghezza s , risulta essere massima all'equatore e nulla ai poli dove sezioni normali e meridiani coincidono.

Per geodetiche aventi un azimut di $\pi/4$ ($\sin 2\alpha = 1$) e $s = 300$ Km tale differenza sarà rispettivamente $0.26''$ all'equatore e $0.13''$ ai poli. Considerando che la precisione di misura degli angoli raggiunge al massimo qualche decimo di secondo sessagesimale ($\pm 0.2''$) e che inoltre non è possibile, a causa della curvatura terrestre, effettuare misure fra punti distanti più di 200 Km, si può concludere che una misura di azimut anche se effettuata con riferimento ad una sezione normale può sempre considerarsi riferita ad una geodetica.

Quanto esposto costituisce la sostanza dei teoremi della Geodesia operativa, ovvero *qualunque misura di azimut, angolo o distanza eseguita con i mezzi a disposizione dei topografi (sulla reale superficie fisica) può ritenersi, invece eseguita con riferimento ad archi di geodetiche sulla superficie di riferimento (ellissoide).*

9 - Risoluzione di figure geometriche sulla superficie di riferimento

Stabilito che qualsiasi misura di angoli e distanze si può considerare riferita ad archi di geodetiche ellissoidiche, ne deriva che i relativi calcoli, tipo la risoluzione di un triangolo, di un quadrilatero o di una figura poligonale, dovrebbero essere eseguiti con gli algoritmi propri della trigonometria ellissoidica; tale trigonometria è però piuttosto complessa, per cui è opportuno esaminare la possibilità di eseguire i calcoli in maniera più semplice, in relazione al fatto che i triangoli, o le figure che in genere si devono risolvere, hanno lati che raramente eccedono i 60 Km, ovvero sono piccoli rispetto ai raggi di curvatura dell'ellissoide, pari a circa 6400 Km.

Il rapporto s/ρ fra la lunghezza di un arco di geodetica s ed il raggio di curvatura del meridiano ρ , è quindi una quantità inferiore ad uno e non eccede in genere 10^{-2} , questa quantità verrà presa in genere come quantità piccola del primo ordine, e si esprimeranno le approssimazioni delle formule in termini di potenze di tale quantità; sono altresì quantità piccole del 1° ordine i rapporti s/N , s/R_α e così via, come pure è da considerare tale il valore di e^2 .

A conclusione della breve indagine che segue si vedrà che, se gli archi di geodetica che compongono le figure oggetto del calcolo non eccedono i 100 Km, i calcoli eseguiti con gli algoritmi della trigonometria sferica danno risultati praticamente uguali a quelli che si otterrebbero usando la trigonometria ellissoidica; inoltre se le dimensioni non eccedono i 15÷20 Km, i risultati che si ottengono con la trigonometria piana sono praticamente uguali a quelli che si otterrebbero usando algoritmi più raffinati.

Si vedrà in seguito che misure di angoli e di distanze non possono considerarsi *esatte*, per cui la posizione dei punti rilevati risulta caratterizzata da un'incertezza più o meno alta a seconda degli strumenti e dei metodi usati; ne segue che potranno reputarsi praticamente uguali i risultati di due calcoli eseguiti con algoritmi diversi tutte le volte che le loro differenze siano decisamente inferiori alle incertezze derivanti dalle misure.

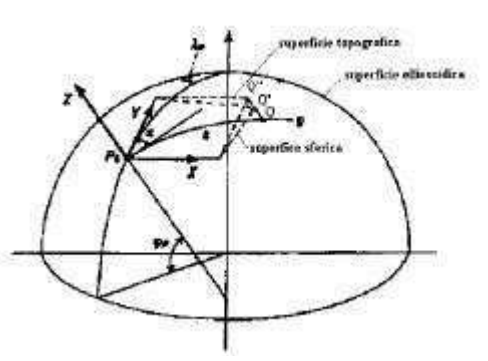


Fig. n. 9.1 - Geodetica riferita alla terna Euleriana

Si consideri (fig 9.1) la terna cartesiana ortogonale P_0XYZ avente il piano XY tangente all'ellissoide nel punto P_0 , di coordinate geografiche φ_p, λ_p , avente cioè l'asse Z diretto secondo la normale all'ellissoide, l'asse Y tangente al meridiano e diretto verso Nord e l'asse X diretto verso Est (terna Euleriana); sia g una geodetica uscente da P_0 secondo l'azimut α , e sia s la lunghezza dell'arco di geodetica compreso fra l'origine ed un punto generico Q ; le formule di PUISEUX-WEINGARTEN, che si derivano dalle formule di FRENET, danno, mediante uno sviluppo in serie, le coordinate cartesiane del punto Q della geodetica in funzione della lunghezza dell'arco s

$$\begin{cases} X = s \cdot \sin \alpha (1 - s^2 / 6 \rho N + \dots) \\ Y = s \cdot \cos \alpha (1 - s^2 / 6 \rho N + \dots) \\ Z = -s^2 / 2 R_\alpha \dots \end{cases} \quad (9.1)$$

in cui vengono indicati i termini dello sviluppo dell'ordine di s^3/R^3 e trascurati (\longleftrightarrow) quelli del tipo s^4/R^4 perché quantità piccola e superiore al secondo ordine.

10 - Campo geodetico o sferico (<100 km)

Si consideri una sfera di raggio $R = \sqrt{\rho N}$ tangente all'ellissoide nello stesso punto P_0 . Anche per la sfera gli sviluppi sono applicabili purché al posto di ρ ed N si ponga il raggio R della sfera; per un punto $Q'(X', Y', Z')$ posto a distanza s dall'origine sull'arco di cerchio massimo avente l'azimut α si hanno, nella stessa approssimazione, le coordinate

$$\begin{cases} X' = s \cdot \sin \alpha (1 - s^2 / 6 R^2) \\ Y' = s \cdot \cos \alpha (1 - s^2 / 6 R^2) \\ Z' = -s^2 / 2 R \dots \end{cases}$$

Calcolando i valori delle due prime differenze $(X-X')$ e $(Y-Y')$ per $\varphi = 0$ ed $\varphi = \pi/2$ e per $\alpha = 0$ ovvero $\alpha = \pi/2$ in modo da ottenere i valori massimi, si trova che per $s=100$ Km tali differenze non superano 27 mm e si possono reputare decisamente inferiori alle incertezze di posizione derivanti dalle misure.

Si può concludere che i problemi riguardanti figure geometriche ellissoidiche comprese nell'intorno di un punto di raggio fino a 100 km, possono essere risolti con gli algoritmi della trigonometria sferica, assumendo come raggio della sfera la media geometrica dei raggi principale di curvatura dell'ellissoide nel punto considerato.

Per le quote, invece, occorre precisare che nel campo geodetico non è trascurabile la differenza fra la curvatura $1/R_\alpha$ della sezione normale (uguale alla curvatura della geodetica) e la curvatura $1/\sqrt{\rho N}$ della sfera locale.

Sviluppando i calcoli si ottiene:

s	1Km	10Km	20Km	50Km	10Km
$(z - z')$	0.13mm	1.3cm	5.4cm	0.33m	1.3m

Gli scarti indicati risultano molto elevati anche per piccole distanze rispetto alla precisione ottenibile nella misura dei dislivelli; da ciò possiamo concludere che i calcoli relativi alle quote dei punti, o meglio alla misura dei dislivelli, devono comunque essere riferiti ad una superficie di riferimento avente curvatura ellissoidica per distanze superiori a 20 Km.

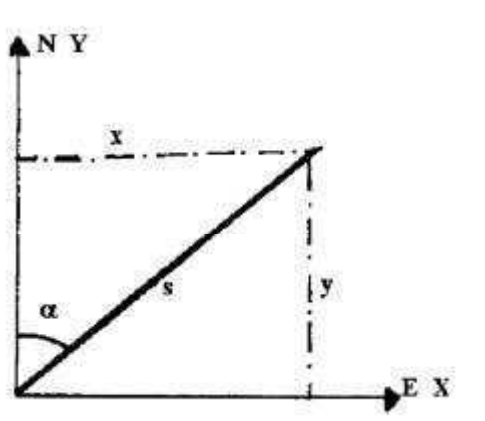


Fig. n. 10.1- Coordinate planimetriche nel campo topografico

11 - Campo topografico

Per valori di s di qualche decina di chilometri le coordinate planimetriche (fig.n.11.1) del punto $Q''(X'', Y'', Z'')$ si possono calcolare con le

$$\begin{cases} X'' = s \cdot \sin \alpha \\ Y'' = s \cdot \cos \alpha \\ Z'' = 0 \dots \dots \dots \end{cases}$$

potendosi ritenere trascurabile l'apporto del secondo termine fra parentesi tonda delle prime due formule (9.1) come si vede dalla seguente tabella

s	10Km	15Km	20Km	30Km	50Km
$s^2/6\rho N$	$0.4 \cdot 10^{-6}$	$0.9 \cdot 10^{-6}$	$1.6 \cdot 10^{-6}$	$3.7 \cdot 10^{-6}$	$10.4 \cdot 10^{-6}$

Non si può fissare un limite ben definito all'intorno di un punto, chiamato campo topografico, in cui è possibile usare le formule della trigonometria e geometria piana, poiché questo dipende dalla precisione delle misure; si può dire che l'intorno ha un raggio di almeno 15 Km se si tiene conto di una precisione di 10^{-6} (1 millimetro fra punti distanti un chilometro).

Come si è già detto, invece, nelle misure dei dislivelli si deve tenere conto della curvatura della superficie di riferimento; infatti lo scostamento fra il piano tangente in un punto P_0 all'ellissoide, sul quale tutti i punti hanno $z = 0$, e punti dell'ellissoide (o anche per semplicità della sfera locale) che a distanza s da P_0 hanno $z = s^2 / \sqrt{\rho N}$, non può essere trascurato anche per valori di s molto bassi come risulta dalla seguente tabella

s	0.1 Km	0.5 Km	1.0 Km	5.0 Km	10.0 Km
$\frac{s^2}{\sqrt{\rho N}}$	0.0008 m	0.02 m	0.08 m	2.0 m	7.8 m

Poiché si possono misurare differenze di quota fra punti distanti 100 m con la precisione dei decimi di millimetro, si può constatare che in tali casi è necessario tenere conto della curvatura della superficie di riferimento, anche per distanze molto piccole

12 - Eccesso sferico - Teorema di Legendre

Prima di dare l'enunciato di questo teorema è necessario precisare che la somma dei tre angoli A, B e C di un triangolo sferico è superiore a π di una quantità, che si denota con \mathcal{E} , chiamata eccesso sferico, e cioè

$$\hat{A} + \hat{B} + \hat{C} - \pi = \mathcal{E}$$

Si dimostra facilmente (con il metodo dei fusi sferici) che l'eccesso sferico è numericamente valutabile, in radianti, facendo il rapporto fra l'area S del triangolo sferico ed il quadrato del raggio della sfera e cioè (fig. 12.1):

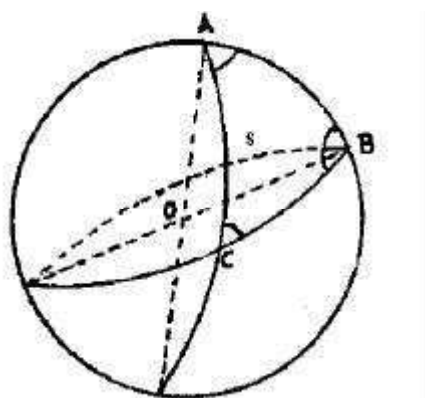


Fig. n. 12.1- Fusi sferici

$$F_A + F_B + F_C - 2S = 2\pi R^2$$

essendo :

$$\frac{F_A}{\hat{A}} = \frac{2\pi R^2}{\pi} = \frac{F_B}{\hat{B}} = \frac{F_C}{\hat{C}}$$

$$2R^2(\hat{A} + \hat{B} + \hat{C}) - 2S = 2\pi R^2$$

$$\mathcal{E} + \pi = S / R^2 + \pi$$

$$\mathcal{E} = S / R^2$$

Enunciato di Legendre: sia dato un triangolo sferico i cui lati l siano una piccola frazione del raggio R della sfera e si assuma il rapporto l/R come quantità piccola del 1° ordine; commettendo un errore dell'ordine di $(l/R)^4$ gli angoli del triangolo piano che ha i lati della stessa lunghezza dei lati del triangolo sferico si possono derivare dagli angoli di quest'ultimo sottraendo a ciascuno di essi un terzo dell'eccesso sferico.

La applicazione di tale enunciato comporta la necessità di dover calcolare l'eccesso sferico, e quindi, l'area del triangolo sferico. Si può dimostrare che, a meno di errori dell'ordine di $(l/R)^4$, cioè nello stesso ordine di approssimazione del teorema di Legendre, l'area del triangolo sferico si può calcolare con le formule della trigonometria piana riferendosi ad un triangolo piano avente gli stessi lati del triangolo sferico e gli angoli depurati dell'eccesso sferico.

$$\varepsilon = S' / R^2 + T_4$$

con S' area del triangolo piano.
Sviluppando i calcoli si ottiene:

S	15Km	30Km	60Km
ε	0.5"	2"	8"

Si può notare che il teorema di Legendre non è più necessario quando i lati del triangolo siano inferiori a 15 km. In tal caso l'eccesso sferico è di circa 0.5" e quindi la correzione da apportare a ciascun angolo è di circa 0.16", inferiore all'approssimazione con cui vengono effettuate le misure degli angoli.

Si dimostra così per altra via, che i calcoli relativi a figure geometriche contenute nel campo topografico (fino a 15 km) possono essere eseguiti, solo *per le coordinate planimetriche*, impiegando gli algoritmi della trigonometria piana.

CAPITOLO 2 - CARTOGRAFIA

CAPITOLO 2 - CARTOGRAFIA.....	27
1-Definizioni dei vari tipi di carte.....	29
2 - Posizione del problema della cartografia.....	31
3 - Carta di Mercatore	40
4 - Proiezione Stereografica	42
5 – Rappresentazione cartografica di Gauss	44
6 - Il Sistema UTM. (Universal Transverse Mercator)	46
7 – IL sistema GAUSS-BOAGA	47
8- La Carta Catastale.....	52
9 - Le Carte Regionali	53
10 - Le Carte Tematiche.....	54
11 - Cartografia Numerica	54
11.1 – CAD - AM/FM	55
11.2 - GIS, LIS, Sistemi Informativi Territoriali (S.I.T.)	56

Si definisce CARTA TOPOGRAFICA la rappresentazione figurativa sul piano della superficie fisica terrestre, secondo norme e segni convenzionali, dalla quale sia possibile ricavare elementi geometrici (distanze, dislivelli, angoli, aree, ecc.) con prefissate tolleranze.

Ciò è possibile se la carta è dotata di un SISTEMA DI RIFERIMENTO che consenta di ricavare le COORDINATE PLANOALTIMETRICHE di ogni punto e/o oggetto rappresentativo (linee, poligoni, aree, ecc.).

Vedremo in seguito che in Italia esistono tuttora cartografie riferite a SISTEMI DIVERSI, anche se con notevoli sforzi si sta cercando di introdurre un unico SISTEMA NAZIONALE (Gauss- Boaga) ed addirittura uno MONDIALE (WGS84).

Il prodotto finale di un generico rilievo (topografico, fotogrammetrico, satellitare, ecc.) è la costruzione di una rappresentazione grafica (o carta), ad una determinata scala $s = 1/n$, nella quale siano riportati tutti gli elementi osservati e classificati sul terreno mediante segni convenzionali, imitativi per quanto consentito dalla scala, ovvero puramente convenzionali, rispettando determinati limiti di tolleranza. Si ottengono in tal modo le carte topografiche *rilevate* che sono la fonte e la base geometrica di ogni carta di diversa utilizzazione.

Da queste si possono ricavare le carte *derivate*, ottenute riducendo la scala di rappresentazione con opportuni procedimenti di riduzione fotografica, con la conseguente “eliminazione” di molti particolari e l’eventuale aumento dei segni convenzionali.

Si noti, che il procedimento inverso, cioè l’ingrandimento di una carta topografica non consente la costruzione di una nuova carta, a scala più grande, non solo perché l’ultima sarebbe priva dei particolari che le competono in funzione della scala prescelta, ma soprattutto perché la nuova carta manterrebbe la accuratezza, ovviamente minore, relativa alla scala della rappresentazione da cui è stata generata.

Negli ultimi anni si stanno sviluppando metodi di produzione della CARTOGRAFIA NUMERICA, i quali consistono in elaborati cartografici sviluppati in “ambiente informatico”.

Con tale denominazione ci si riferisce alla produzione di elaborati da parte di “PLOTTERS” che realizzano in forma grafica i contenuti di “FILES” (archivi numerici) memorizzati mediante computers. Poiché è possibile individuare ogni punto mediante le sue tre coordinate in un sistema di riferimento spaziale (tre numeri) associando a ciascun punto uno, o eventualmente più, codici qualitativi (casa, strada, confine, linea elettrica, curva di livello, ecc.), si intuisce la definizione “numerica” attribuita alla cartografia prodotta con il metodo ortofotogrammetrico dai “restitutori” analitici (o digitali) governati ed assistiti da computers, dei quali è ben nota la continua evoluzione di velocità, potenza di calcolo, capacità di memoria, ecc.

La forma grafica su carta diviene, quindi, una delle numerose e possibili forme di visualizzazione dell’oggetto tridimensionale costituito dalla superficie fisica del terreno, con tutte le opere antropiche su di essa esistenti. Sono molto semplici da realizzare con i moderni CAD (Computer Aided Design) viste ed eventuali grafici in prospettiva, assonometria, sezioni, ecc. di porzioni limitate di prodotti cartografici, con eventuale animazione da punti di vista diversi, ecc.

La potenza di elaborazione della cartografia numerica ha reso ancora più facile la possibilità di variare la scala della rappresentazione grafica che, si ribadisce, è connessa alla scala di rilievo prescelta (scala nominale), con i relativi limiti di incertezza nel posizionamento dei punti e, quindi, con i conseguenti limiti nelle dimensioni dei particolari rappresentabili in funzione del graficismo scelto.

L’eventuale utilizzo di una scala di rappresentazione più grande di quella scelta per il rilievo, comporta gli stessi gravi inconvenienti del citato ingrandimento fotografico.

L'eventuale riduzione di scala causerà, invece, la necessità di “sfortimento” dei numerosi particolari che, ovviamente, non competono alla scala più piccola da utilizzare.
In relazione ai contenuti delle carte possiamo distinguere:

A) CARTE TOPOGRAFICHE: (rilevate o derivate) dette anche “DI BASE”. Tali carte hanno lo scopo di dare del terreno la più completa rappresentazione possibile, compatibile con la scala, di tutti i particolari di interesse generale (orografia, idrografia, morfologia, viabilità, manufatti, limiti amministrativi, toponomastica, ecc.), comprensiva delle indicazioni altimetriche e dei riferimenti relativi al sistema di coordinate piane utilizzate e/o al reticolato geografico. La caratteristica fondamentale di una carta topografica generale viene stabilita dal rispettare geometricamente determinate tolleranze, cioè dalla possibilità di effettuare misure grafiche sulla carta di angoli, distanze, e dislivelli, con prefissati limiti d'incertezza.

Tali misure sulla carta sono caratterizzate dall'incertezza, nella determinazione grafica della posizione dei punti, esprimibile mediante:

- l'errore medio planimetrico (compreso tra ± 0.2 e ± 0.5 mm alla scala della carta), ricavabile da una copia stampata dalla carta, che scaturisce dall'errore di graficismo, ossia l'incertezza connessa con il disegno cartografico di ogni linea. Di conseguenza una carta topografica in scala 1:1000 avrà, per esempio, un errore medio planimetrico compreso tra 20 e 50 cm.
- l'errore medio altimetrico, compreso tra ± 0.02 e ± 0.2 mm alla scala della carta, per le quote numeriche scritte “punti quotati” e tra ± 0.1 e ± 0.5 mm per quelle ricavate dalle curve di livello, la cui equidistanza è in genere stabilita in 1/1000 del denominatore della scala.

La scelta della scala di una rappresentazione cartografica è legata, quindi, alla combinazione dei due elementi che determinano la precisione di una carta:

- il graficismo, ossia l'incertezza (± 0.2 mm / ± 0.5 mm) connessa con il disegno cartografico nel tracciamento di una linea, limite che viene mantenuto “convenzionalmente” anche nella moderna Cartografia Numerica (scala nominale);
- la dimensione del più piccolo particolare fisico del terreno (o manufatto) che si vuole evidenziare graficamente con simbolo imitativo in scala, e quindi la necessità di “spazio-grafico” necessario per contenerli (ad esempio un marciapiede, un tombino, un palo elettrico, ecc.).

Le fasi necessarie alla costruzione di una carta possono così distinguersi:

- ripresa effettuata da una camera istallata nella carlinga di un aeroplano la quale fornisce una serie di “fotogrammi” disposti secondo strisciate. I fotogrammi sono scattati ad intervalli di tempo tali che le zone del territorio sono ricoperte, con una percentuale elevata, nella direzione di volo (vedere *intervallo di scatto* e *quota di volo*);
- rete di punti di appoggio indispensabili per la conoscenza delle coordinate spaziali di un certo numero di punti sul terreno, necessari per l'orientamento dei modelli;
- restituzione fotogrammetrica il complesso di operazioni necessarie per l'acquisizione degli elementi fisici e antropici (case, strade, fiumi, ecc...), ed il successivo tracciamento computerizzato della carta, sia planimetrica che altimetrica e con l'attribuzione delle relative “codifiche” legate al data-base del SIT da realizzare;
- ricognizione da effettuare direttamente sul terreno per la individuazione dei particolari “nascosti”;
- collaudo effettuato in corso d'opera sin dalle prime fasi del lavoro, allo scopo di :

1. garantire la corrispondenza alla NORMATIVA
2. garantire il rispetto delle “tolleranze” relative alla scala “prescelta”
3. verificare i dati metrici (angoli, distanze, dislivelli)

B) CARTE TEMATICHE: carte speciali compilate per un determinato scopo; possono essere ottenute dalle carte generali rilevate o derivate inserendovi, dopo averle eventualmente private di alcuni particolari, la rappresentazione di altri elementi, rilevati di proposito sul terreno, riguardanti speciali fenomeni. A volte sono allestite per mettere in maggiore evidenza cartografica un certo aspetto della rappresentazione.

Sono carte tematiche le carte urbanistiche, geologiche, magnetiche, forestali, stradali (tra cui le note ottime carte del Touring Club Italiano), ecc. La differenza più evidente tra le carte tematiche e le topografiche è da ricercare nei limiti di precisione geometrica della carta, che per le tematiche sono alquanto elastici, e soprattutto dalla mancanza del SISTEMA DI RIFERIMENTO che, invece, viene espressamente indicato nelle CARTE TOPOGRAFICHE. Si possono avere, quindi, carte tematiche con assenza di contenuto geometrico e via via carte tematiche con maggiori/minori qualità di metriche, ma sempre inferiori o al più paragonabili alle caratteristiche della carta topografica da cui la tematica è stata tratta.

Le considerazioni che seguono si riferiscono alla costruzione delle carte topografiche di base.

2 - Posizione del problema della cartografia

La rappresentazione sul piano di una determinata porzione della superficie fisica terrestre comporta problemi diversi a seconda della estensione della zona interessata dalla carta.

Se si tratta di una porzione sufficientemente ristretta, tale cioè da essere compresa nel “campo topografico” del suo punto centrale, la superficie di riferimento per le operazioni planimetriche è, in questo caso, il piano tangente all’ellissoide terrestre nel punto stesso: di conseguenza la rappresentazione cartografica è immediata in quanto il piano della carta coincide col suddetto piano tangente e su di esso vengono tracciate figure simili alle corrispondenti figure esistenti nella realtà, con rapporti di similitudine costante per tutta la carta e pari alla scala $s = 1/n$.

E’ il caso risolto nell’esercizio grafico “Piano quotato e Piano curve di livello” svolto nel corso di Topografia, nella cui rappresentazione grafica sono riportati:

- angoli, uguali a quelli misurati in campagna,
- distanze topografiche pari a quelle misurate, con la riduzione indicata dalla scala,
- coordinate planimetriche, ridotte in scala, riferite ad un sistema locale.

Si noti che anche in questo caso le quote, sia che vengano trascritte con caratteri numerici per i punti quotati, sia rappresentate dalle curve di livello, sono sempre riferite alla superficie assunta come riferimento altimetrico, cioè al Geoide.

Più complesso è invece il problema di rappresentare porzioni più estese della superficie fisica terrestre perché in tal caso la superficie di riferimento da adottare per le misure planimetriche non può essere il piano tangente. Si è già visto (fig. n.2.1) che la complessità di rappresentazione della superficie fisica ha imposto l’adozione di una ben precisa superficie di riferimento (il Geoide) sulla quale vengono proiettati, secondo la direzione della verticale, tutti i punti che interessano il rilievo, o punti caratteristici del terreno.

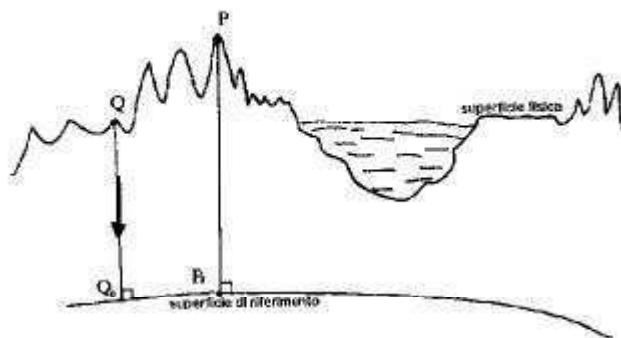


Fig. n. 2.1- Superficie fisica del terreno, geoide e rappresentazione del terreno

Si suppone quindi di eseguire le misure di angoli e di distanze sul geoide, in modo da ricavare la posizione di un punto qualsiasi in coordinate curvilinee. A causa delle approssimazioni che ne derivano abbiamo già visto come sia lecito, solo per le coordinate planimetriche, sostituire al geoide alcune superfici più facilmente individuabili con espressioni analitiche: l'ellissoide o la sfera locale. In ogni caso è sempre possibile individuare la posizione di un punto qualunque sulla superficie di riferimento (e quindi sulla superficie fisica terrestre) attraverso le sue coordinate geografiche ellissoidiche: latitudine ϕ e longitudine λ .

A questo punto sarebbe già possibile una rappresentazione in scala molto piccola del geoide conservando un supporto curvilineo: il mappamondo, in cui in genere $s = 1/10.000.000$

E' naturale che quando si voglia una rappresentazione in scala più grande si dovrà ricorrere ad un supporto piano, e quindi alla necessità di stabilire una corrispondenza biunivoca tra coordinate curvilinee ϕ e λ coordinate piane x, y

$$\begin{cases} x = f(\phi, \lambda) \\ y = f(\phi, \lambda) \end{cases} \quad (2.1)$$

in modo che ad ogni punto dell'ellissoide terrestre corrisponda un punto del piano di rappresentazione e viceversa.

E' noto dalla geometria che tale corrispondenza può aversi in modo rigoroso solo tra superfici aventi la stessa curvatura totale (prodotto degli inversi dei raggi principali di curvatura). Superfici aventi la stessa curvatura totale si dicono applicabili; in particolare se una superficie è applicabile al piano essa si dice svilupabile. Nel caso dell'ellissoide terrestre i raggi principali di curvatura sono ρ (raggio del meridiano) e N (Gran Normale), e quindi

$$\frac{1}{\rho} \cdot \frac{1}{N} = \frac{1}{R^2} \neq 0 \quad \text{dove } R = \text{raggio della sfera locale}$$

quantità variabile in funzione di ϕ diversa da 0 (curvatura totale del piano). E' evidente, quindi, che non è possibile rappresentare la superficie terrestre sul piano senza introdurre delle deformazioni (esempio della palla di gomma).

Per esprimere e poter valutare analiticamente tali deformazioni si fa uso dei loro moduli e precisamente:

- modulo di deformazione lineare $m_l = ds'/ds$ dove ds' è la distanza misurata sulla carta moltiplicata per n denominatore della scala e ds è la distanza misurata sull'ellissoide.

Le carte in cui $m_l = 1$ si dicono equidistanti.

- modulo di deformazione superficiale $m_s = S'/S$ dove S' è l'area misurata sulla carta per il quadrato del denominatore della scala e S è l'area misurata sull'ellissoide.

Le carte in cui $m_s = 1$ si dicono equivalenti.

- modulo di deformazione angolare $m_\alpha = \alpha' - \alpha$ dove α' è un angolo misurato sulla carta, α è l'angolo corrispondente sull'ellissoide.

Le carte in cui $m_\alpha = 0$ si dicono conformi.

Nell'impossibilità di ottenere una carta che possieda contemporaneamente i tre requisiti ($m_l = 1; m_s = 1; m_\alpha = 0$) si studiano i diversi sistemi per costruire delle carte che approssimino quanto più è possibile la similitudine tra carta e superficie fisica terrestre proiettata sulla superficie di riferimento (ellissoide o "sfera locale"). Esistono infiniti modi di creare tali corrispondenze biunivoche, sia per via geometrica, sia per via analitica; per valutare la possibilità d'uso di ciascuno di essi si dovranno studiare:

- le espressioni analitiche delle formule di corrispondenza,
- i valori assunti dai moduli di deformazione,
- la posizione in piano del reticolato geografico (cioè le linee sul piano che corrispondono ai meridiani e paralleli).

A) RISOLUZIONE GEOMETRICA

Le corrispondenze biunivoche si possono determinare per via geometrica individuando diversi sistemi di proiezione cartografica:

1. Proiezioni prospettiche

Si approssima il geoide con la sfera locale nel punto situato al centro della zona da rappresentare, e si proiettano i punti della superficie terrestre sul piano tangente alla sfera locale. A seconda della posizione del centro di proiezione P si hanno le seguenti proiezioni prospettiche (fig. n.2.2):

- centrografica: P al centro della sfera;
- stereografica: P all'antipolo del punto di tangenza;
- scenografica: P sulla normale al piano tangente fuori della sfera;
- ortografica: P sulla normale al piano tangente all' ∞ .

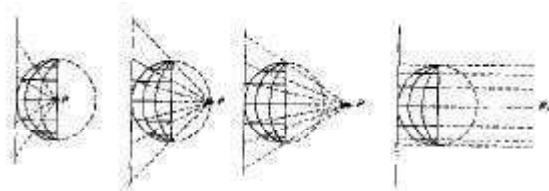


Fig. n. 2.2- Proiezioni prospettiche

Si vede subito che le deformazioni sono contenute solo all'intorno del punto e che si può, eventualmente, estendere la zona di rappresentazione se il piano diventa secante.

2. Proiezioni cilindriche

Il geoide viene approssimato con l'ellissoide (o anche con la sfera locale). La proiezione avviene su di un cilindro tangente lungo l'equatore (diretta) (fig.n. 2.3), o su di un cilindro tangente lungo un meridiano (inversa) (fig.n. 2.4), proiettando ogni punto o dal centro della sfera o dal centro del parallelo del punto.

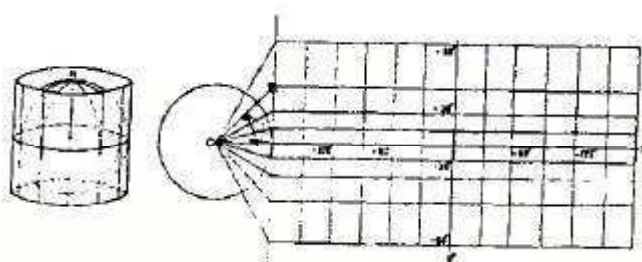


Fig. n. 2.3- Proiezione cilindrica diretta

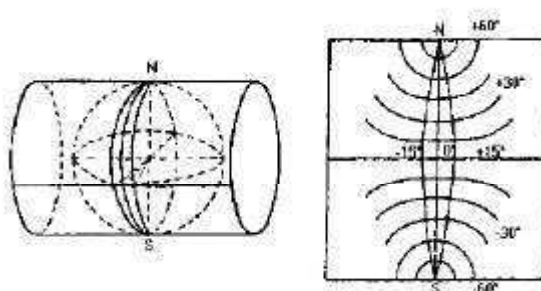


Fig. n. 2.4- Proiezione cilindrica inversa

Successivamente il cilindro, (che ha curvatura totale nulla), tagliato lungo una generatrice, viene sviluppato sul piano. Anche in questo caso le deformazioni sono contenute solo a cavallo della linea di tangenza ed eventualmente si possono usare più cilindri secanti per ridurre le deformazioni ai margini della zona interessata.

3. Proiezioni coniche

Il geoide, approssimato all'ellissoide o alla sfera locale; viene proiettato sul cono che gli è tangente lungo il parallelo alla latitudine media della fascia da rappresentare con centro di

proiezione di ogni punto P al centro dell'ellissoide o al centro del parallelo del punto. Il cono viene poi sviluppato sul piano (fig. n. 2.5)

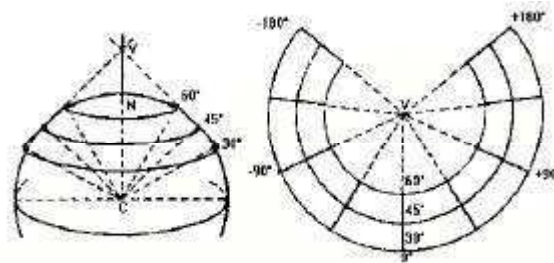


Fig. n. 2.5- Proiezione conica

Si vedrà più innanzi come queste proiezioni con opportune modifiche siano state utilizzate per le rappresentazioni cartografiche.

B) RISOLUZIONE ANALITICA

E' possibile risolvere il problema di individuare la corrispondenza biunivoca tra punti dell'ellissoide e punti del piano per via analitica. Si tratta di individuare tra le infinità di soluzioni che risolvono le

$$\begin{cases} x = f(\varphi, \lambda) \\ y = f(\varphi, \lambda) \end{cases}$$

quelle che più ci soddisfano, imponendo determinate condizioni.

Ovviamente l'imposizione di alcune condizioni limita il verificarsi di altre, cioè è impossibile scrivere delle equazioni che abbiano come risultato per i moduli di deformazione rispettivamente i valori 1, 1, 0 ma è possibile ottenere corrispondenze in cui siano verificati valori molto prossimi a quelli desiderati, con dei compromessi tra le varie deformazioni.

Considerando il modulo di deformazione lineare:

$$m_1 = m = ds'/ds$$

si avrà (fig.n.2.6)

$$(ds')^2 = dx^2 + dy^2$$

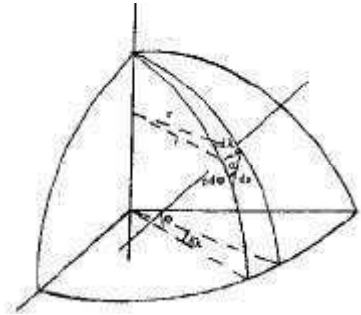


Fig. n. 2.6- Moduli di deformazione lineare

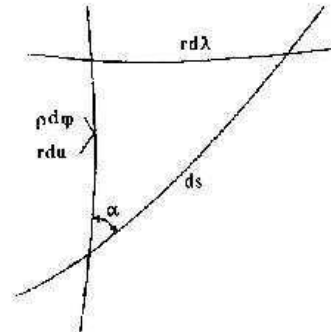


Fig. n. 2.7- Elemento ds sull'ellissoide

$$ds^2 = \rho^2 d\varphi^2 + r^2 d\lambda^2$$

per cui

$$m^2 = \frac{dx^2 + dy^2}{\rho^2 d\varphi^2 + r^2 d\lambda^2}$$

Per unificare i coefficienti del denominatore conviene introdurre la

$$U = \int_0^\varphi \frac{\rho}{r} d\varphi \qquad du = \frac{\rho}{r} d\varphi$$

denominata *latitudine isometrica* o crescente. Infatti al crescere di ρ sappiamo che diminuisce r (ai poli $r=0$ e $\rho/r = \infty$ e quindi cresce il rapporto ρ/r).

Esprimendo il modulo di deformazione in funzione della U avremo

$$m^2 = \frac{dx^2 + dy^2}{r^2 (du^2 + d\lambda^2)}$$

Essendo

$$\begin{cases} x = f(u, \lambda) \\ y = f(u, \lambda) \end{cases}$$

si ottiene:

$$(ds')^2 = dx^2 + dy^2 = (\partial x / \partial u)^2 du^2 + 2(\partial x / \partial u \cdot \partial x / \partial \lambda) du d\lambda + (\partial x / \partial \lambda)^2 d\lambda^2 +$$

$$(\partial y / \partial u)^2 du^2 + 2(\partial y / \partial u \cdot \partial y / \partial \lambda) du d\lambda + (\partial y / \partial \lambda)^2 d\lambda^2$$

da cui

$$(ds')^2 = [(\partial x / \partial u)^2 + (\partial y / \partial u)^2] du^2 + 2[(\partial x / \partial u \cdot \partial x / \partial \lambda) + (\partial y / \partial u \cdot \partial y / \partial \lambda)] du d\lambda + [(\partial x / \partial \lambda)^2 + (\partial y / \partial \lambda)^2] d\lambda^2$$

$$(ds')^2 = Edu^2 + 2F du d\lambda + G d\lambda^2$$

con

$$\begin{aligned} E &= (\partial x / \partial u)^2 + (\partial y / \partial u)^2 \\ F &= (\partial x / \partial u \cdot \partial x / \partial \lambda) + (\partial y / \partial u \cdot \partial y / \partial \lambda) \\ G &= (\partial x / \partial \lambda)^2 + (\partial y / \partial \lambda)^2 \end{aligned}$$

da cui si ricava

$$m^2 = \frac{Edu^2 + 2F du d\lambda + G d\lambda^2}{r^2 (du^2 + d\lambda^2)}$$

e dividendo numeratore e denominatore per du^2 si ottiene:

$$m^2 = \frac{E + 2F \left(\frac{d\lambda}{du} \right) + G \left(\frac{d\lambda^2}{du^2} \right)}{r^2 \left(1 + \frac{d\lambda^2}{du^2} \right)}$$

esprimendo il rapporto $\frac{d\lambda}{du}$ in funzione dell'azimut α dell'elemento ds sull'ellissoide (fig.n. 2.7), essendo $\rho d\phi = r du$ ed ancora

$$r \cdot d\lambda = r \cdot du \cdot tg\alpha \quad \text{cioè} \quad \frac{d\lambda}{du} = tg\alpha$$

sostituendo si ha:

$$m^2 = \frac{E + 2F tg\alpha + G tg^2\alpha}{r^2 (1 + tg^2\alpha)}$$

essendo

$$1 + tg^2\alpha = 1/\cos^2\alpha$$

si ottiene:

$$r^2 m^2 = E \cos^2\alpha + 2F \sin\alpha \cos\alpha + G \sin^2\alpha \quad (2.2)$$

Quindi m_1 dipende dai parametri dell'equazione della superficie da rappresentare (ρ, r) da quelli della rappresentazione cartografica (E, F, G) ed inoltre dalla posizione del punto P sull'ellissoide (U, λ) e dall'azimut α dell'elemento ds uscente da P ; *quindi m_1 varia da un punto all'altro della superficie ed anche con la direzione.*

Ponendo:

$$\begin{array}{lll} \alpha = 0 & \text{si ha} & m^2 = E/r^2 \\ \alpha = \pi/2 & \text{si ha} & m^2 = G/r^2 \end{array}$$

rispettivamente quadrati del modulo di deformazione lineare secondo le direzioni di un meridiano e di un parallelo.

Analoghe considerazioni sul modulo di deformazione angolare e sul modulo di deformazione superficiale portano ad espressioni analitiche che consentirebbero, una volta fissato il tipo di proiezione cartografica e, quindi, le formule di corrispondenza, di ricavare l'entità delle deformazioni indotte nella rappresentazione.

Quando in essa sono presenti, in maniera più o meno sensibile, tutti i tipi di deformazioni descritti, si ottengono carte "AFILATTICHE" che non sono né conformi né equivalenti ma tendono a trovare un compromesso tra le diverse deformazioni, avendole limitate il più possibile.

E' evidente che in questo caso non si può scrivere un'equazione generale valida per l'intera classe delle "afilattiche": sarà solo possibile scrivere l'equazione di una singola rappresentazione (Mercatore, U.P.S., ecc.) imponendo il rispetto di particolari vincoli, validi solo per quella rappresentazione.

Se imponiamo, per esempio, che per una certa rappresentazione cartografica il modulo di deformazione lineare vari in funzione della posizione del punto P ma non della direzione α dell'elemento lineare uscente da P , avremo una trasformazione in cui a figure infinitesime sull'ellissoide corrispondono figure simili sul piano: quindi la rappresentazione sarà conforme. Cioè l'angolo α formato da due linee qualsiasi uscenti da P sull'ellissoide risulta uguale all'angolo α' formato dalle trasformate piane di tali linee: in particolare il reticolato geografico formato dalle trasformate piane dei meridiani e dei paralleli risulta ortogonale.

Si potranno ottenere, quindi, delle equazioni generali per le rappresentazioni conformi imponendo che il modulo di deformazione, espresso dall'equazione (2.2), sia indipendente dall'azimut α dell'elemento lineare uscente dal punto considerato.

In termini analitici tale condizione si esprime annullando la derivata prima rispetto ad α dell'espressione (2.2) di m :

$$dm^2/d\alpha = 0$$

Cioè

$$-2E \cdot \sin\alpha \cdot \cos\alpha + 2F \cdot \cos^2\alpha - 2F \cdot \sin^2\alpha + 2G \sin\alpha \cdot \cos\alpha = 0$$

essendo:

- $2 \cdot \sin\alpha \cdot \cos\alpha = \sin 2\alpha$
- $-\cos^2\alpha + \sin^2\alpha = \cos 2\alpha$

si ottiene:

$$(G - E)\sin 2\alpha - 2F \cos 2\alpha = 0 \quad \Leftrightarrow \quad (G - E)\tan 2\alpha - 2F = 0$$

da cui perché il primo membro sia nullo dovrà essere:

$$\begin{cases} G - E = 0 \\ F = 0 \end{cases}$$

che individuano le espressioni analitiche delle rappresentazioni conformi.
Ricordando le espressioni di E, F, G si ha:

$$\begin{cases} \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 = \left(\frac{\partial x}{\partial \lambda}\right)^2 + \left(\frac{\partial y}{\partial \lambda}\right)^2 \\ \left(\frac{\partial x}{\partial u} \cdot \frac{\partial x}{\partial \lambda}\right) + \left(\frac{\partial y}{\partial u} \cdot \frac{\partial y}{\partial \lambda}\right) = 0 \end{cases} \quad (2.3)$$

Possiamo trasformare la seconda delle (2.3)

$$\left(\frac{\partial x}{\partial u} \cdot \frac{\partial x}{\partial \lambda}\right) = -\left(\frac{\partial y}{\partial u} \cdot \frac{\partial y}{\partial \lambda}\right)$$

$$\left(\frac{\frac{\partial x}{\partial u}}{\frac{\partial y}{\partial \lambda}}\right) = -\left(\frac{\frac{\partial y}{\partial u}}{\frac{\partial x}{\partial \lambda}}\right) = H$$

avendo indicato con H il valore comune dei due rapporti.
Si ottiene

$$\begin{cases} \frac{\partial x}{\partial u} = H \frac{\partial y}{\partial \lambda} \\ \frac{\partial y}{\partial u} = -H \frac{\partial x}{\partial \lambda} \end{cases}$$

e sostituendo nella prima delle (2.3)

$$H^2 \left(\frac{\partial y}{\partial \lambda}\right)^2 + H^2 \left(\frac{\partial x}{\partial \lambda}\right)^2 = \left(\frac{\partial x}{\partial \lambda}\right)^2 + \left(\frac{\partial y}{\partial \lambda}\right)^2$$

Cioè

$$H = \pm 1$$

e quindi

$$\begin{cases} \frac{\partial x}{\partial u} = \pm \frac{\partial y}{\partial \lambda} \\ \frac{\partial y}{\partial u} = \pm \frac{\partial x}{\partial \lambda} \end{cases} \quad \underline{\text{EQUAZIONI DI RIEMANN - CAUCHY}} \quad (2.4)$$

Le (2.4) sono le equazioni generali delle rappresentazioni conformi. Esse esprimono x e y , coordinate piane, in funzione di U e λ coordinate isometriche sull'ellissoide.

Essendo queste ultime equazioni differenziali alle derivate parziali avranno diverse possibilità di integrazione a seconda delle funzioni arbitrarie che esprimono le condizioni al contorno e quindi il particolare tipo di rappresentazione prescelta.

Le relazioni trovate esprimono la condizione necessaria e sufficiente affinché la variabile complessa $(x+iy)$ sia funzione della variabile complessa $(U + i\lambda)$.

Quindi potremo scrivere:

$$(x + iy) = f(U + i\lambda)$$

Ciò significa che una qualsiasi rappresentazione conforme della superficie ellissoidica sul piano si realizza scegliendo una qualunque funzione $f(\cdot)$, purché regolare, della variabile complessa formata con i parametri isometrici della superficie.

Poiché si è costretti a rappresentare porzioni limitate della superficie terrestre, onde evitare eccessive deformazioni, si avrà che il termine $i\lambda$ (λ espresso in radianti) si potrà considerare come incremento e quindi sviluppare la funzione in serie di Taylor:

$$x + iy = f(U) + i\lambda \frac{df}{dU} + i^2 \frac{\lambda^2}{2!} \left(\frac{d^2 f}{dU^2} \right) + i^3 \frac{\lambda^3}{3!} \left(\frac{d^3 f}{dU^3} \right) + \dots$$

Quindi trascurando i termini successivi dello sviluppo ed eguagliando nei due membri le parti reali e i coefficienti di quelle immaginarie (essendo $i^2 = -1; i^3 = -i; i^4 = 1; i^5 = i, \dots$) si ha:

$$x + iy = f(U) + i\lambda f'(U) - \frac{\lambda^2}{2!} f''(U) - i\frac{\lambda^3}{3!} f'''(U) + \frac{\lambda^4}{4!} f^{IV}(U) + i\frac{\lambda^5}{5!} f^V(U) + \dots$$

e quindi

$$\begin{cases} x = f(U) - \frac{1}{2!} \frac{d^2 f(U)}{dU^2} \lambda^2 + \frac{1}{4!} \frac{d^4 f(U)}{dU^4} \lambda^4 + \dots \\ y = \frac{df(U)}{dU} \lambda - \frac{1}{3!} \frac{d^3 f(U)}{dU^3} \lambda^3 + \frac{\lambda^5}{5!} \frac{d^5 f(U)}{dU^5} + \dots \end{cases} \quad (2.5)$$

che rappresentano le formule di corrispondenza di una qualsiasi rappresentazione conforme.

Le condizioni al contorno consentono di definire la $f(U)$ per il particolare tipo di rappresentazione scelta e quindi anche i termini relativi alle derivate di $f(U)$.

Si sceglierà il grado di approssimazione e quindi il numero di termini della serie da considerare in funzione dell'ampiezza in longitudine della zona da rappresentare e della precisione richiesta nella misura delle coordinate dei punti sulla carta realizzata.

3 - Carta di Mercatore

Considerando la proiezione cilindrica pura (fig.n.3), con cilindro tangente lungo l'equatore, si ottiene che meridiani e paralleli si proiettano secondo linee rette, parallele tra loro e a distanza costante i primi, mentre risultano a distanza crescente con la latitudine i secondi. La rappresentazione così ricavata è afilattica e le deformazioni sono crescenti al crescere della latitudine.

Mercatore (1569) modificò tale proiezione in modo da ottenere una carta conforme.

Infatti la proiezione dei paralleli verrà ottenuta da una espressione analitica funzione della latitudine isometrica.

Se nelle (2.5) si assume $f(U)=aU$ cioè

$$f(U) = a \int_0^{\varphi} \frac{\rho}{r} d\varphi \quad \left(\text{dove } a = \text{semiasse equatoriale dell'ellissoide prescelto} \right)$$

si avrà:

$$\begin{cases} N = X = aU \\ E = Y = a\lambda \end{cases}$$

dove E = l'arco di equatore compreso tra il meridiano origine e il meridiano passante per il punto P essendo nulle le derivate di ordine successivo di $df(U)/dU$.

Si avrà che un meridiano ($\lambda = cost$) è rappresentato dalla equazione $E=cost$ cioè una retta parallela all'asse delle N , ed analogamente un parallelo ($\varphi = cost$) dalla $N=cost$, parallela all'asse delle E .

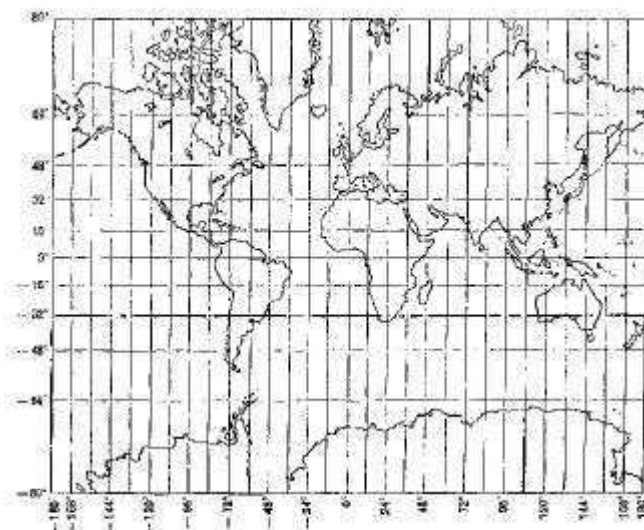


Fig. n. 3.1- Il reticolato geografico nella proiezione di Mercatore

Esaminando il modulo di deformazione lineare lungo il meridiano si ottiene:

$$m = \frac{ds'}{ds} = \frac{dx}{\rho d\varphi} = \frac{adu}{\rho d\varphi} = \frac{a \frac{\rho}{r} d\varphi}{\rho d\varphi} = \frac{a}{r} = \frac{(1 - e^2 \sin^2 \varphi)^{1/2}}{\cos \varphi}$$

funzione crescente con la latitudine.

Si noti che il modulo di deformazione lineare è funzione della sola φ , cioè è costante su ogni parallelo, in particolare $m = 1$ sull'equatore mentre $m = \infty$ ai poli, quindi avremo che la scala della rappresentazione dovrà variare passando da un parallelo all'altro (fig.n. 3.2).

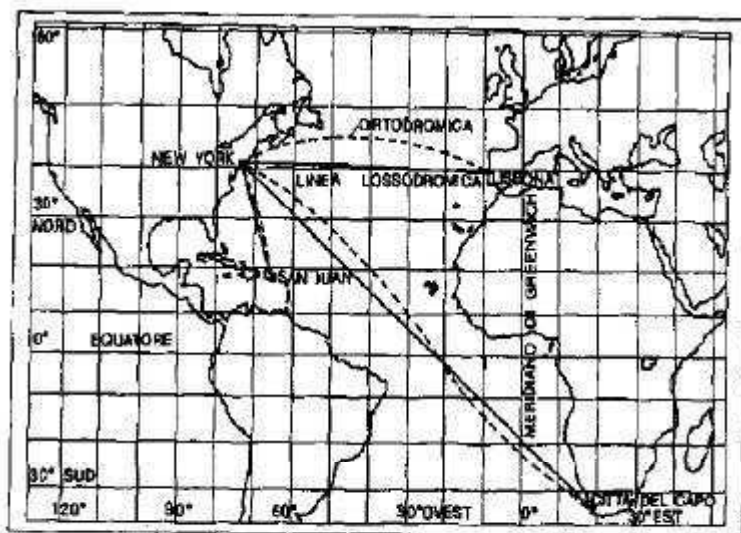


Fig. n. 3.2- Proiezione di Mercatore: linea lossodromica e linea ortodromica

I vantaggi di tale tipo di rappresentazione, che viene utilizzata molto per la navigazione marittima e aerea si riferiscono non solo alla possibilità di rappresentare con continuità l'intera superficie del globo (tra -80° e $+80^\circ$ di latitudine) in un unico sistema di coordinate piane, ma soprattutto per la caratteristica di individuare sulla carta la congiungente tra due punti P e Q come una retta che interseca tutti i meridiani secondo un angolo costante θ (angolo di rotta) e quindi, poiché la carta è conforme, anche nell'ellissoide la linea corrispondente alla retta PQ interseca i meridiani secondo lo stesso angolo costante.

Si tenga presente che la "lossodromia" non corrisponde sull'ellissoide alla linea di minor percorso (che coincide con la geodetica¹ dei due punti P e Q) per distanze molto grandi si ricorre ad una spezzata di lossodromie che approssimi la "ortodromia" coincidente con la trasformata della geodetica.

4 - Proiezione Stereografica

La rappresentazione U.P.S. (Universal Polar Stereographic) (fig.n.4.1), è utilizzata per rappresentare le due calotte polari: è la sola proiezione geometrica pura che si utilizza. Il piano di proiezione è tangente a un polo e i punti sono proiettati con centro nell'altro polo.

Si vede subito che è una carta conforme perché i meridiani, che sono trasformati in rette uscenti da uno stesso punto, formano angoli uguali alle differenze di longitudine e incontrano secondo angoli retti, i paralleli che coincidono con circonferenze concentriche.

¹ Si definisce "geodetica" quella linea sulla superficie ellissoidica che gode della proprietà di avere in ogni punto la normale principale coincidente con la normale alla superficie (curva di minima distanza tra due punti).

Fig. n. 4.1- Proiezione stereografica²

Essendo:

$$P_1 M_1 = 2R \cdot \operatorname{tg}(\frac{\pi}{4} - \frac{\varphi}{2})$$

Da cui si trae:

$$\begin{aligned} x &= 2Rtg(\pi/4 - \varphi/2)\cos\lambda \\ y &= 2Rtg(\pi/4 - \varphi/2)\sin\lambda \end{aligned}$$

per

- $\lambda = \cos t \Rightarrow y = x \tan \lambda$ cioè una retta
- $\varphi = \cos t \Rightarrow x^2 + y^2 = r^2$ cioè una circonferenza

Su questa carta la ortodromia (ossia la trasformata della geodetica tra due punti) si può considerare rettilinea e permette di individuare i diversi angoli di rotta da mantenere per seguire un percorso minimo, corrispondente ai diversi angoli che la retta forma con i meridiani (fig. n. 4.2).

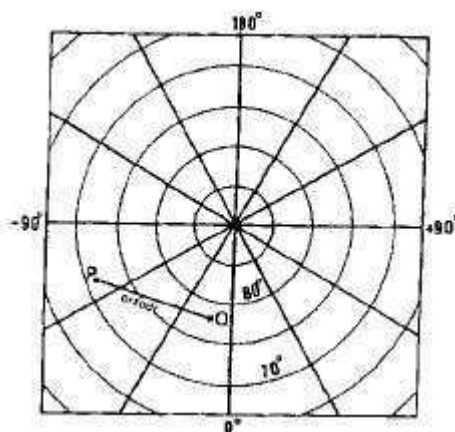


Fig. n. 4.2- Rappresentazione U.P.S

² Si noti in fig.n. 4.1 la possibilità di ridurre le deformazioni, che crescono allontanandosi dal punto di tangenza, utilizzando un piano secante anziché tangente.

La rappresentazione cartografica è ottenuta dalla proiezione cilindrica inversa, cioè con il cilindro tangente all'ellissoide lungo un meridiano. Si ricava una proiezione del tipo di fig.n.5.1 nella quale è riportato il reticolato geografico relativo al semiellissoide compreso tra le longitudini -90° e $+90^\circ$. Per ridurre le fortissime deformazioni ai margini occorre limitare il valore della differenza di longitudine dei punti rispetto al meridiano di tangenza per cui la rappresentazione viene effettuata per "FUSI" di una limitata ampiezza (in genere 6°), per ognuno dei quali si assume come meridiano di riferimento quello coincidente con il meridiano di tangenza ovvero meridiano centrale del fuso (fig. n. 5.2). Con la limitazione di ampiezza del fuso si ha il vantaggio di ridurre le deformazioni in limiti accettabili, come si vedrà più innanzi.

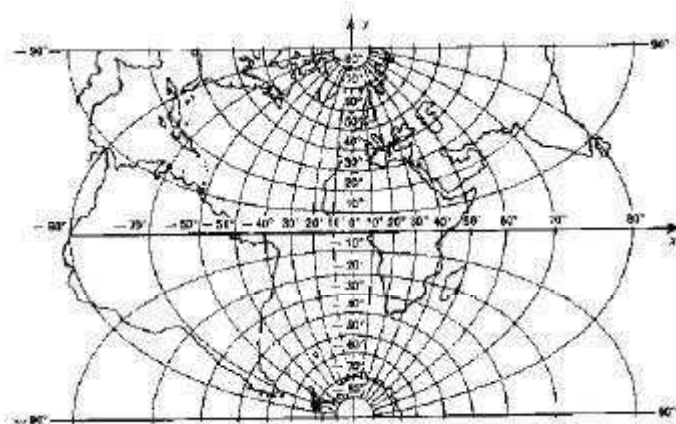


Fig. n. 5.1- Reticolato geografico nella proiezione di Gauss

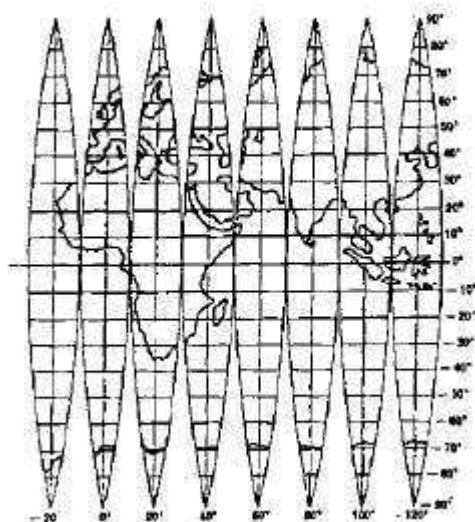


Fig. n. 5.2- I fusi nella carta di Gauss

Per ricavare analiticamente le formule della rappresentazione imponremo la condizione che il meridiano di tangenza si sviluppi con la sua lunghezza reale lungo l'asse delle N, indicando con B l'arco di meridiano compreso tra l'equatore ed il parallelo del punto considerato, si ha:

$$f(U) = B = \int_0^{\varphi} \rho d\varphi \cong R \operatorname{tg} \varphi$$

Avendo determinato la $f(U)$ si possono ricavare tutte le sue derivate ed ottenere le espressioni dalle formule di corrispondenza.

$$\begin{cases} N = x = B + a_2 \lambda^2 + a_4 \lambda^4 + a_6 \lambda^6 \dots\dots \\ E = y = a_1 \lambda + a_3 \lambda^3 + a_5 \lambda^5 + \dots\dots\dots \end{cases}$$

dove, a_1 , a_2 ecc. sono coefficienti che dipendono solo dalla latitudine. Analizzando le espressioni analitiche che si ottengono per $\varphi = \cos t$ (paralleli) e $\lambda = \cos t$ (meridiani) si ritrovano le curve già viste nella proiezione geometrica che hanno le convessità verso gli assi e si intersecano tra loro con angoli di $\pi/2$ essendo la carta conforme. Le curve relative ai meridiani e paralleli individuano sulla carta il “reticolato geografico” che coincide con i margini secondo cui sono tagliati i fogli che compongono una carta, assicurando ovviamente la continuità tra fogli contigui (fig. n. 5.3).

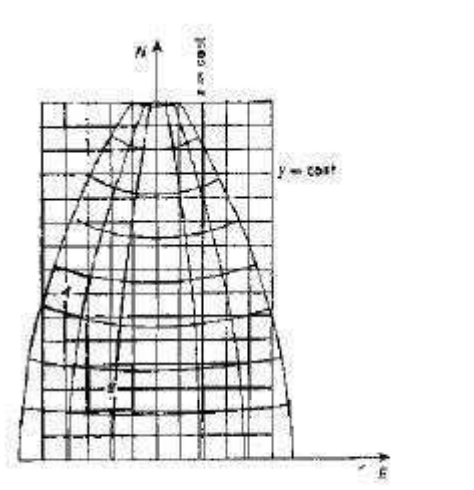


Fig. n. 5.3- A-Taglio dei fogli della cartografia italiana secondo le trasformate di due meridiani e di due paralleli; B-Secondo il reticolato chilometrico solo per le carte a grande scala

Si può ricavare, anche in questo tipo di rappresentazione, il modulo di deformazione lineare essendo:

$$m = \frac{\sqrt{E}}{r} = \frac{\left[\left(\frac{\partial x}{\partial u} \right)^2 + \left(\frac{\partial y}{\partial u} \right)^2 \right]^{1/2}}{r}$$

Si tratta, cioè, di derivare le espressioni di x e di y . Con diversi passaggi analitici, che prevedono anche sviluppi in serie, si perviene all'espressione:

$$m = 1 + \frac{1}{2\lambda^2 \cos^2 \varphi} \quad \text{o anche} \quad m = 1 + \frac{y^2}{2N^2}$$

formule che sono valide per distanze infinitesime.

Per distanze finite tra due punti di coordinate piane $P_1(x_1, Y_1)$ e $P(x_2, y_2)$ si avrà il valore medio del modulo di deformazione espresso da:

$$m_{1,2} = 1 + \frac{y_1^2 + y_1 y_2 + y_2^2}{6N_m \rho_m}$$

Si noti che m è sempre maggiore di 1 ($m > 1$) e che va aumentando al crescere della distanza del meridiano centrale del fuso.

Si possono ricavare anche:

- γ l'angolo di convergenza dei meridiani (cioè l'angolo destrorso formato dalla direzione del meridiano in un punto P e la parallela per P al meridiano centrale):

$$\text{tg} \gamma = \Delta x / \Delta y = \lambda \text{sen} \varphi \left[1 + (\lambda^2 / 3) \cos^2 \varphi (1 + t^2 + 3\eta^2 + 2\eta^4) + \dots \right]$$

- ε l'angolo di riduzione alla corda (cioè l'angolo tra la tangente alla curva, trasformata della geodetica tra P_1 e P_2 e la corda che li congiunge) fig. n. 5.4.

$$\varepsilon = \frac{1}{6N^2} (x_2 - x_1) \cdot (2y_1 + y_2)$$

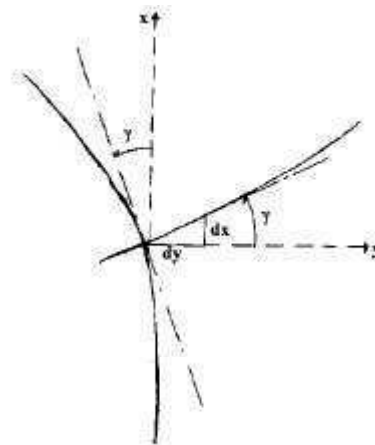


Fig. n. 5.4- Rappresentazione della convergenza tra meridiano e parallelo

6 - Il Sistema UTM. (Universal Transverse Mercator)

La carta di Gauss, con il sistema di rappresentazione policilindrico per fusi è stata adottata in diversi stati compresi tra -80 e +80 di latitudine (ai poli è stata adottata la U.P.S.).

Naturalmente ogni fuso ha il suo sistema di riferimento per le coordinate piane e, in particolare, nel **SISTEMA U.T.M.** si è convenuto di assumere per il meridiano centrale del

fuso rettificato una coordinata y o E di 500 km (per evitare ascisse negative) mentre le ordinate x o N sono riferite all'equatore rettificato.

E' stato adottato quindi un sistema valido in tutto il globo, per esprimere in cifre la posizione planimetrica di un punto qualsiasi.

A tale scopo si indica dapprima il fuso: numeri da 1 a 60 procedendo da W verso E, il numero 1 è compreso tra i meridiani 180° e 174° W da Greenwich, il numero 31 da 0° a 6° E.

L'Italia è, quindi, compresa nei fusi 32°, 33° e per una parte della penisola salentina anche dal 34°.

Ogni fuso è diviso in 20 "zone" ciascuna avente 8° di latitudine indicate da lettere maiuscole (l'Italia che si estende in latitudine da circa 36° a 48° è interessata dalla zona S da 32° a 40° e dalla T da 40° a 48°).

Ciascuna zona è suddivisa in quadrati di 100 Km di lato, individuati da una coppia di lettere maiuscole (fig. n. 15); quindi in definitiva un punto è individuato da

- **numero del fuso;**
- **lettera della zona;**
- **coppia di lettere del quadrato di 100 km**
- **coordinate piane N e E**

la coordinata N esprime la distanza dall'equatore mentre la E definisce la distanza dalla falsa origine con l'approssimazione relativa alla scala della carta (per l'I.G.M. 1:25000 è pari a ± 5 m).

Volendo, quindi, ricavare misure di angoli e distanze sulla carta, note le coordinate piane N, E di un punto generico, mediante il reticolato chilometrico riportato sulla carta, si procede come segue:

- le distanze si calcolano con la $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ dividendo per il modulo di deformazione lineare e moltiplicando per il denominatore della scala;
- gli angoli si ricavano con le formule di trigonometria piana (differenza di angoli di direzione) applicando le riduzioni alle corde e la convergenza del meridiano.

7 – IL sistema GAUSS-BOAGA

La carta ufficiale dello Stato è una carta rilevata alla scala 1:25.000 dall'I.G.M., disegnata nella rappresentazione conforme di Gauss con formule di trasformazione ricavate dal prof G. Boaga (1918-1961). All'Italia competono due fusi: fuso 32 (fuso Ovest o 1° fuso) e fuso 33 (fuso Est o 2° fuso) del sistema U.T.M. e, per ovviare alla discontinuità dei due sistemi di riferimento, è stata creata una "zona di sovrapposizione" di 30' (dal meridiano 12° al 12° 30') nella quale le coordinate dei punti sono riferite ad entrambi i sistemi; mentre il fuso Est si estende per altri 30' per includere anche l'estremità della penisola salentina che ricadrebbe nel 34° (fig.n.7.2).

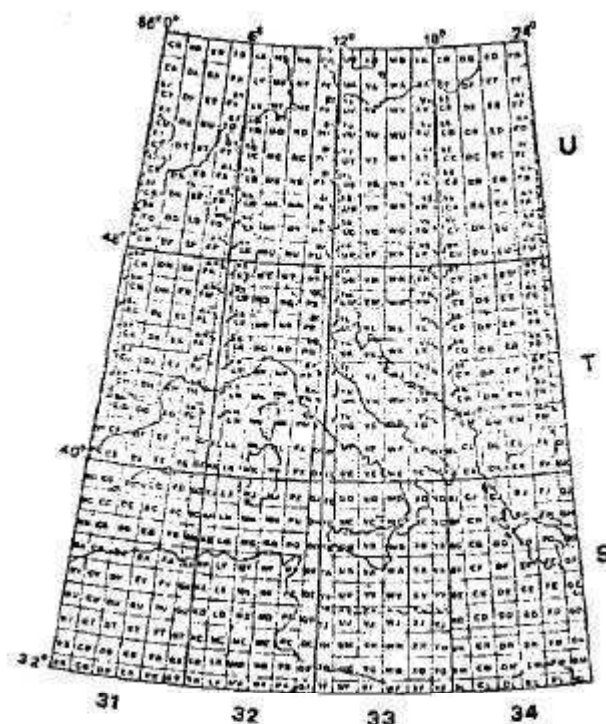


Fig. n. 7.1- Zone di suddivisione dei fusi in quadrati di 100 Km contraddistinti da una coppia di lettere maiuscole

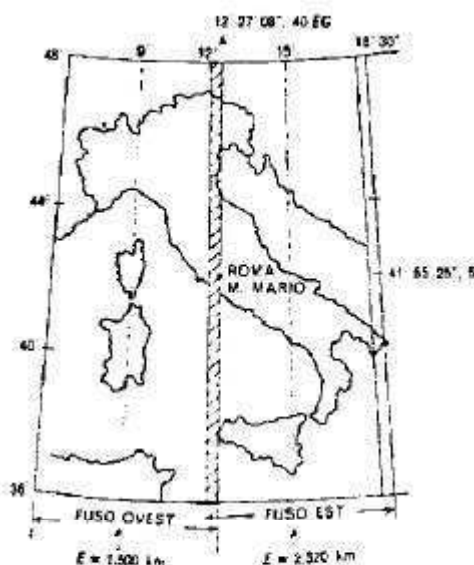


Fig. n. 7.2- I fusi della cartografia italiana nel sistema nazionale Gauss-Boaga

Ai meridiani centrali dei due fusi è stata attribuita nel sistema *Gauss-Boaga* rispettivamente l'ascissa 1500 km (per il meridiano 9°) e 2520 km (per il 15°) (falsa origine) in modo tale che: la prima cifra della coordinata E indica immediatamente a quale fuso appartiene il punto (fig. n.7.2).

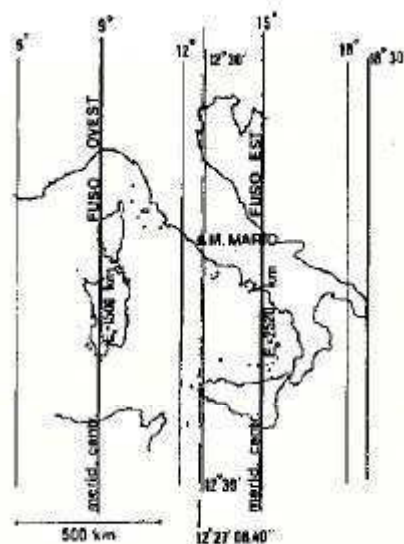


Fig. n. 7.3- I due fusi della cartografia ufficiale italiana

La carta è stata riportata anche nelle scale 1:100.000 e 1:50.000 utilizzando sempre supporti grafici di dimensioni 40x40 cm; naturalmente alla scala 1:100000 queste dimensioni corrispondono ad una differenza di 20' per la latitudine φ e 30' per la longitudine λ ; alla scala 1:50.000 si avranno 10' di latitudine e 15' di longitudine, mentre alla scala 1:25.000 si avranno 5' in latitudine e 7'30" in longitudine. Il "Foglio" in scala 1:100.000 è contraddistinto da un numero che va da 1 a 277 (fig. n.7.4), comprende 4 "quadranti" in scala 1:50.000 individuati dai numeri romani I, II, III, IV ed, infine, ogni quadrante è suddiviso in 4 "tavole" in scala: 1:25.000 indicate dalle sigle NO, NE, SO, SE (fig. n.6.5). Essendo stabilita l'origine della copertura ed il "taglio" (ampiezza) dei fogli in φ e λ di conseguenza ai quattro vertici di ogni carta alle diverse scale risultano attribuite coordinate geografiche ben precise (che scaturiscono dal modulo del foglio).

Ciò vale anche per le carte 1:25000 nelle quali le coordinate dei quattro vertici (NO, SO, NE, SE) sono state ricavate dal Prof. Boaga nel 1940 mediante il calcolo, applicando le formule analitiche della *Carta di Gauss* con coefficienti spinti fino ad a_{20} , a_{21} , ecc

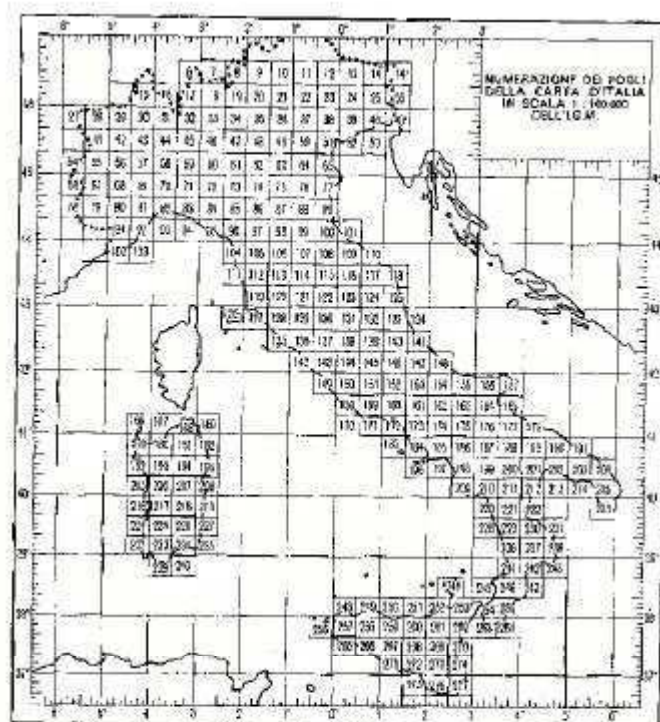


Fig. n. 7.4- Suddivisione della carta d'Italia alla scala 1:100.000



Fig. n. 7.5- Suddivisione del foglio alla scala 1:100.000 della carta d'Italia

Nella tavoletta al 25.000, anch'essa tagliata secondo meridiani e paralleli di posizione nota, vengono riportati:

- i reticolati chilometrici relativi ai due sistemi, U.T.M. e Gauss-Boaga, e le indicazioni sui margini che permettono di ricavare anche le coordinate geografiche (φ, λ), di ciascun punto della carta;
- la “legenda” dei simboli grafici che sostituiscono la rappresentazione in scala dei particolari, quando le dimensioni degli stessi sono paragonabili all'errore di graficismo (per esempio la larghezza delle strade);
- le indicazioni altimetriche relative a punti quotati e curve di livello, con le relative precisioni (± 1.5 m per i punti quotati e ± 2.5 m per le curve di livello);
- il valore dell'angolo di convergenza dei meridiani e della declinazione magnetica per il punto centrale della carta.

Si tenga presente che per la carta 1:25.000 è stata apportata nel sistema Gauss-Boaga una “contrazione” al piano di rappresentazione, ottenuta moltiplicando per una costante tutte le coordinate dei punti corrispondenti ai vertici delle tavolette (di coordinate note) ed anche ai Vertici Trigonometrici (la cui distanza calcolata $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ va divisa per la costante di contrazione per ottenere la distanza reale).

La costante è pari a 0.9996, il che equivale a rendere praticamente equidistante la carta. Infatti il modulo di deformazione lineare che prima era compreso tra 1 (lungo il meridiano di tangenza) e 1.0008 (agli estremi del fuso) viene trasformato in 0.9996 sull'asse e 1.0004 ai margini del fuso, quindi misure grafiche *nell'ambito di una tavoletta* si possono considerare, nei limiti di precisione conseguibile, non affette da deformazioni, essendo le stesse inferiori al graficismo ($14\text{km} \cdot 0.0004 = 5.6\text{m}$ all'estremità del fuso).

Quando invece i punti si trovano su due tavolette diverse per ricavare la distanza con la formula $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ si dovrà tener conto del modulo di deformazione lineare mediante l'espressione:

$$m_{1,2} = 0,9996 + \frac{y_1^2 + y_1 y_2 + y_2^2}{6N_m \rho_m}$$

Valida anche per il sistema U.T.M.

Si noti la leggera differenza tra reticolato U.T.M. e reticolato Gauss-Boaga dovuta al diverso “orientamento” dell'ellissoide di HAYFORD.

Per l'orientamento dell'ellissoide vengono stabilite per convenzione:

- la coincidenza della normale all'ellissoide e della verticale al geoide nel punto di emanazione;
- la coincidenza della quota ellissoidica e della quota geoidica nel punto di emanazione;
- la coincidenza nella direzione del nord ellissoidico e del nord geoidico (astronomico). quest'ultima operazione si realizza fissando l'azimut di una direzione di riferimento (Monte Soratte per il sistema Gauss-Boaga)

Per il primo sistema è stato scelto come punto di emanazione un punto in prossimità di Potsdam (Germania) mentre per il secondo è stato utilizzato il punto corrispondente all'Osservatorio Astronomico di M. Mario (ROMA), di cui sono note le coordinate astronomiche e quindi la deviazione della verticale, che viene, quindi, “convenzionalmente” annullata in quel punto.

Si noti, inoltre, che il reticolato chilometrico (parallelo agli assi di riferimento) risulta convergente rispetto ai bordi della tavoletta (meridiani e paralleli) di un angolo pari a γ (convergenza del meridiano).

	ellissoide	orientamento	Rappresentazione cartografica	Falsa origine Est	Contrazione
Gauss-Boaga (Roma 40)	HAYFORD	M. Mario	Gauss policilindrica	1500 km 1° 2520 km 2°	0.9996
U.T.M. (ED50)	HAYFORD	Potsdam	Gauss policilindrica	500 km	0.9996
U.T.M.-WGS84	WGS84	ETRF 89 ³ (IGM 95)	Gauss policilindrica	500 km	0.9996

8- La Carta Catastale

Le mappe catastali utilizzano la rappresentazione di CASSINI-SOLDNER (fig.n.8.1).

Preso il meridiano OQ rettificato come asse delle X ed il parallelo passante per 0 come asse delle ascisse, il punto P viene individuato nel piano cartesiano dalle geodetiche QP ed OQ rettificate.

Chiamate con \bar{X}_p, \bar{Y}_p , le coordinate geodetiche rettangolari di P (ricavabile dalle formule del CAP I), si ha per le coordinate cartografiche:

$$\begin{cases} X = \bar{X}_p \\ Y = \bar{Y}_p \end{cases}$$

Come si vede, le immagini dei meridiani e dei paralleli sull'ellissoide di BESSEL sono, sulla carta, delle curve, eccezione fatta per la coppia che, rettificata, costituisce gli assi cartesiani

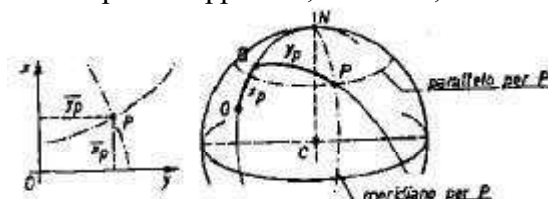


Fig. n. 8.1 - Rappresentazione di Cassini-Soldner

Le deformazioni di questa carta, che viene ottenuta imponendo $m_s=0$ (carte equivalenti), si confondono praticamente con l'errore di graficismo entro un raggio di 70 Km dal centro di sviluppo (origine del sistema cartesiano). In considerazione di questo, sul territorio italiano sono stati individuati oltre 800 centri che rendono la carta policentrica.

Nella cartografia catastale questi centri coincidono generalmente con un vertice I.G.M. di primo ordine; la deformazione è nulla lungo il parallelo, è dell'ordine del 6 per diecimila lungo il meridiano, entro il campo sopraindicato. Per ricoprire tutto il territorio italiano sono state usate 35 zone con altrettante diverse origini!

La carta viene utilizzata a scopo fiscale.

³ Rete geodetica europea di compensazione globale

Le mappe catastali sono in scala 1:2000; solo in zone di scarso interesse è stata usata la scala 1:4000. I centri urbani sono invece rappresentati su allegati alla mappa in scala 1:1000 o anche, dove vi è eccessivo frazionamento, in scala 1:500.

La carta catastale risulta, però, una carta particolare. In essa, infatti; è generalmente rappresentata solo la planimetria. La planimetria inoltre non è nemmeno quella del terreno con la sua morfologia, bensì rappresenta i limiti delle “particelle”, cioè delle porzioni di terreno della stessa destinazione, qualità e classe site nello stesso Comune, appartenenti ai diversi proprietari, individuate dalle seguenti indicazioni: COMUNE, FOGLIO, PARTICELLA.

E’ perciò impossibile usare le mappe per la progettazione poiché le informazioni in esse contenute sono prevalentemente geometrico-fiscali (in funzione della superficie e del reddito).

Esse non hanno nemmeno rilevante contenuto giuridico, poiché il Catasto italiano non è probatorio.

Secondo il codice civile si ricorrerà ad esse per eventuali azioni di riconfinazione od apposizione di termini, solo dopo aver esaminato i dati forniti in ordine

1. dai titoli di proprietà (atti notarili ed eventuali planimetrie allegate, ecc.)
2. dal possesso (segni evidenti quali muretti, recinzioni, siepi, alberi, ecc.)
3. dai dati della “mappa catastale”

Si può immaginare il contenzioso che scaturisce da tale situazione che, invece è stato risolto negli altri Stati Europei con l’introduzione di atti catastali aventi contenuto giuridico.

Degli altri enti cartografici ufficiali italiani, diremo solo che:

- l’ I.I.M. (Istituto Idrografico della Marina) stampa e rileva carte delle coste e carte per la navigazione marittima e fluviale;
- il Servizio Geologico si occupa, peraltro con gravi carenze di fondi, della Carta Geologica Italiana, che ha come supporto i fogli al 100.000 I.G.M.;
- il CIGA (Sezione Cartografica dello S.M. Aeronautica) cura le speciali carte per la navigazione aerea.

9 - Le Carte Regionali

Sia pure con un certo ritardo, rispetto alla media dei paesi europei, anche l’Italia si avvia alla redazione di una carta tecnica fondamentale. Rimangono, allo stato attuale, differenze talvolta cospicue fra i prodotti cartografici di una regione rispetto alle altre, ma, per contro esiste il comune denominatore della scala (che è in gran parte al 5.000) e quello del sistema geometrico di rappresentazione prescelto che è il Gauss-Boaga.

Tra le diverse iniziative ricorderemo quelle del Friuli-Venezia-Giulia, dell’Emilia Romagna, della Liguria, della Campania che hanno scelto la carta disegnata, restituita fotogrammetricamente, i cui lavori, iniziati già negli anni 70 - 80 sono stati conclusi con l’avvio dei successivi SISTEMI INFORMATIVI TERRITORIALI (di cui si parlerà in un paragrafo successivo). Tale procedura è stata seguita anche dalla Cassa per il Mezzogiorno, che ha avuto l’incarico di redigere la carta tecnica d’Italia meridionale, alla scala 1:5000, che interessa anche la PUGLIA della quale sono stati realizzati solo pochissimi elementi. La carta tecnica regionale è naturalmente inquadrata sempre nella rappresentazione nazionale Gauss-Boaga.

Dopo molte discussioni si è deciso il taglio secondo il reticolato geografico, come sottomultiplo della nuova carta I.G.M. al 50.000.

Sul bordo della carta, che è limitato da 1'30" in latitudine, e da 2'30" in longitudine, ed è disegnata su foglio UNI A1 (59.4 x 84,1 cm² che ripiegato in altezza assume le dimensioni UNI A4, pari a 21 x 29.7 cm²) sono riportati i consueti elementi già visti per le carte I.G.M. La carta regionale è oggi già utilizzata come supporto per molte carte tematiche, così come diremo fra breve.

10 - Le Carte Tematiche

Le molteplici attività umane, rivolte ad ordinare, modificare, razionalizzare l'ambiente esterno, trovano come supporto indispensabile le carte tecniche, cioè quelle carte fondamentali a grande scala di cui abbiamo già più volte parlato; ma le attività dell'uomo sono estremamente diversificate fra di loro: nasce da ciò la necessità che siano approntate delle carte che possano fornire, oltre alle consuete informazioni qualitativo-quantitative sulla morfologia del territorio, altre indicazioni particolari.

Esse hanno come base la cartografia ordinaria, a scala piccola, media o grande a seconda delle esigenze come preciseremo più avanti e recano, sovrainpresse, informazioni speciali quali ad esempio le risorse minerarie del territorio, le attività economiche della popolazione che ci vive, la densità delle abitazioni, la natura dei suoli e così via.

Il numero dei soggetti che possono essere rappresentati per mezzo di carte tematiche è naturalmente inesauribile; se si pensa poi che la possibilità di scelta di scale differenti, di raffigurare fenomeni generalizzando o schematizzando, scegliendo simboli o tecniche di rappresentazione e riproduzione diverse è vastissima, si comprende subito come non sia possibile elencare tutte le carte tematiche oggi usate nel mondo.

Particolare attenzione andrà posta, una volta scelta la carta che servirà da supporto per le informazioni tematiche, alla simbologia che verrà utilizzata.

Colori particolari, campiture, schermature, spessore di linee, segni convenzionali, andranno scelti con cura e coordinati con la scala della carta di base.

Indispensabile è qui la "legenda", che dovrà in dettaglio servire da chiave per la lettura della carta; le sovrainpressioni d'altronde non dovranno ledere la geometria e l'intelligibilità delle informazioni fondamentali del supporto cartografico.

Non è inutile sottolineare come sia indispensabile una equa correlazione temporale fra base cartografica ed informazioni tematiche: una carta non aggiornata non serve, o quantomeno falsa il contenuto sovrainpresso.

Altra osservazione importante è quella per cui la simbologia tematica può non rispettare la scala della carta; è sempre però necessario che vi sia una relazione armonica fra le carte e le caratteristiche visive (spessori, dimensioni, ecc.) dei segni usati per le informazioni tematiche.

Sulle carte tematiche un determinato fenomeno può essere rappresentato nella sua estensione e nella sua posizione sia qualitativamente che quantitativamente.

La rappresentazione può essere assoluta (l'oggetto od il fenomeno è definito per mezzo di simboli opportunamente graduati) oppure relativa.

In questo caso si rappresenta la densità tramite colori diversi, e con diversi toni per lo stesso colore.

Le carte della densità d'una popolazione, indicanti il numero di abitanti per Km² o quelle delle precipitazioni, ne sono un tipico esempio.

11 - Cartografia Numerica

Il *supporto informatizzato*, basato cioè su tecnologie informatiche (computer e relative periferiche, memorie di massa e relativo software) ha oggi largamente sostituito carte, tabulati ed altri veicoli tradizionali delle informazioni geografiche.

Al suo interno, le informazioni sono raccolte in archivi, *file* residenti tipicamente su dischi magnetici o magneto-ottici, oppure, quelle in sola lettura, su *CD-ROM* (Compact Disk - Read Only Memory).

Gli archivi organizzati sono detti basi di dati (database) e, se descrivono oggetti e fenomeni del mondo reale di cui è riportata la posizione spaziale, sono definiti database geografici.

La sostituzione del supporto cartaceo con quello informatizzato ha portato a rivedere criteri di raccolta, immagazzinamento, utilizzo e presentazione delle informazioni geografiche; per la diffusione e fruizione, che ora avviene in forma capillare, si è aperta una nuova era. Questo sconvolgimento è forse paragonabile, con i dovuti distinguo, a quello che comportò l'introduzione della stampa a caratteri mobili per la diffusione del libro.

La carta infatti è nel contempo “*contenitore*” e “*supporto di presentazione*” delle informazioni geografiche, aspetti questi che nel supporto informatizzato non sono più coincidenti, in quanto l'utilizzo (analisi, visualizzazione, plottaggio) avviene solitamente operando una selezione sui dati contenuti nel database.

Questo fatto nuovo comporta che nei database informatizzati il contenuto informativo non è più limitato, in termini di qualità e dovizia di dettagli sugli oggetti descritti, dall'impossibilità di poterli disegnare sulla carta.

Nel supporto informatizzato si possono immagazzinare le informazioni con l'accuratezza (accuracy in inglese) massima (a grande scala) *consentita dalle metodologie di acquisizione* (ad esempio dalla restituzione fotogrammetrica) e, tramite un'operazione di selezione, si possono scegliere gli oggetti che si intende (analizzare, vedere sul monitor o plottare su carta, modificare, ecc.); rimane sempre valido il criterio di scelta della scala (scala nominale) da intendere, questa volta, come scala del dettaglio e dei particolari che costituiscono gli oggetti da rappresentare.

La *carta*, ovviamente, continua ad avere la sua validità sia come supporto di presentazione delle informazioni di base e tematiche, sia dei risultati delle analisi operate nel sistema informativo geografico; essa esprime, in altre parole, una particolare visione del database.

Probabilmente nella cartografia prodotta dai database si dovrà indulgere sui segni convenzionali che saranno usati, in quanto questi verranno scelti in modo da poter essere generati “automaticamente” e quindi non potranno avere le forme “artistiche” che tradizionalmente hanno assunto nell'attività cartografica (ad esempio l'indicazione delle scarpe con le “barbette”).

11.1 – CAD - AM/FM

Una delle prime applicazioni dell'informatica nel campo delle informazioni geografiche è stata quella in cui nella fase del disegno e restituzione grafica sono stati introdotti *editor* grafici e disegnatori automatici (*plotter*), realizzando la cosiddetta automazione cartografica; in pratica, ci si è serviti di strumenti moderni per produrre la cartografia nella forma classica.

Questa applicazione va sotto il nome di CAD (Computer Aided Design), in cui i dati geografici rappresentati hanno una struttura semplice, finalizzata principalmente alla sola riproduzione grafica automatizzata.

Successivamente, oltre alla produzione del dato grafico, l'archiviazione elettronica ha consentito di gestire gli oggetti, rappresentati dalle primitive geometriche utilizzate nel disegno, con associati spesso anche database alfanumerici, e ha trovato ampia applicazione,

specialmente nel settore delle reti tecnologiche, nella cosiddetta AM/FM (Automated Mapping /Facilities Management).

11.2 - GIS, LIS, Sistemi Informativi Territoriali (S.I.T.)

La rivoluzione vera e propria si è operata quando lo strumento informatico ha consentito di effettuare analisi spaziali su oggetti del mondo reale attraverso l'esame di un suo modello informatizzato, opportunamente creato. Individuato un fenomeno del mondo reale o "entità" (casa, strada, fiume, ecc.), la sua rappresentazione avviene tramite un "oggetto", che ne è la trasposizione digitale nel database.

Il sistema popolato da questi oggetti va sotto il nome, in inglese, di Geographic Information System (GIS). Nei GIS il dato geografico è costituito, oltre che dalla *componente geometrica* (forma dell'oggetto rappresentata in scala con alta precisione), dalla *componente alfanumerica* associata (strutturata in un database contenente gli attributi dell'oggetto), ed anche dalla componente di relazione fra gli oggetti (topologia), necessaria per l'effettuazione delle analisi spaziali. L'espressione inglese Land Information System (LIS) intende invece un sistema informativo geografico prevalentemente orientato alla gestione degli aspetti amministrativi e legali del territorio.

In Italiano, l'espressione Geographic Information System (GIS) è generalmente tradotta con Sistema Informativo Territoriale (SIT), evidenziando la presenza, nel sistema informativo geografico, sia di dati prettamente fisici (strade, fiumi, case, vegetazione ecc.), su cui operare analisi spaziali, sia di altri, non necessariamente georeferenziati, cioè di coordinate note, ma correlati agli oggetti geografici attinenti la gestione del territorio sotto tutti gli aspetti.

I SIT contengono dati di base, corrispondenti al contenuto informativo della grande scala, con l'aggiunta di altre informazioni, distribuite su vari livelli informativi, riferite alla pianificazione (piani regolatori e di intervento), allo studio e all'uso del suolo, alle reti tecnologiche comunali (gas, acqua, fogna, luce, ecc.), ecc.

Di tutti i dati geografici in forma digitale acquisiti in Italia, nell'ambito della Pubblica Amministrazione, solo una minima parte sono conformati ed utilizzati in un'ottica di sistema informativo, pur essendo proprio la P.A. l'utente privilegiato di tale potente ed affidabile tecnologia di informazione, conoscenza, decisione.

I SIT sono definiti come sistemi in grado di acquisire, gestire ed aggiornare la conoscenza degli oggetti geografici, attraverso i quali gestire ed amministrare i corrispondenti oggetti del mondo reale di cui sono rappresentazione, o di fornire supporto alle decisioni in ambito territoriale. Le principali componenti dei SIT sono la base dei dati (database) geografici, le apparecchiature informatiche (hardware), il software, le risorse umane e tutta l'organizzazione in grado di gestire l'acquisizione, l'aggiornamento, effettuare l'analisi ed il trattamento dei dati. Il database, a sua volta, può essere scomposto in una componente di base (topografia di base), più stabile nel tempo ed in una tematica, più attinente allo specifico settore di applicazione del SIT, che è soggetta ad aggiornamenti più frequenti.

Il database è la componente fondamentale dei sistemi informativi geografici, in quanto è quella più costosa (arriva al 75-80% del valore complessivo), e quella di più lunga durata (10-15 anni di integrazioni ed aggiornamenti prima della sostituzione). La base dati va progettata attentamente, in quanto "il miglior sistema informativo è quello che contiene i dati strettamente necessari e di accuratezza adeguata per la propria applicazione". L'hardware ed il software, che apparentemente rivestono un ruolo predominante, in realtà hanno una limitata importanza in termini economici e di durata (3-6 anni). Negli ultimi anni, il costo dell'hardware è notevolmente diminuito, mentre il software ha mantenuto gli stessi livelli,

probabilmente a causa dei grossi investimenti che le società hanno dovuto effettuare per sviluppare e rendere disponibili “software-tools” con capacità di analisi sempre maggiori. Una particolare attenzione merita la componente umana, e l’ambito in cui questa opera, in relazione alla specifiche conoscenze tecniche che gli addetti all’acquisizione, gestione, interrogazione ed aggiornamento dei dati devono possedere.

Capitolo 3 - TELERILEVAMENTO

Capitolo 3 - TELERILEVAMENTO	58
1 - Le Tecniche di Telerilevamento	59
2 - La correzione degli errori	67
3 - Elaborazione delle misure, interpretazione e classificazione	67
4 - Le Piattaforme	69
5 - Satelliti per le risorse terrestri.....	70
5.1 – Landsat.....	70
5.2 - Satellite SPOT	77
5.3 - EROS	79
5.4 - IKONOS.....	80
5.5 - QUICKBIRD	81
5.6 - I satelliti SAR ERS-1 e 2	82
5.7 - Il satellite J-ERS	82
5.8 - Il satellite RADARSAT.....	82
6 - La Geodesia Spaziale.....	83
7 - Satellite Laser Ranging (SLR).....	85
8 - Sistema Inerziale.....	86
9 – Conclusioni	87

Il Telerilevamento (Remote Sensing) è definibile come quell'insieme di tecniche, strumenti e mezzi interpretativi che permettono di estendere e migliorare le capacità percettive dell'occhio, fornendo all'osservatore informazioni qualitative e quantitative su oggetti posti a distanza, e quindi sull'ambiente circostante.

Storicamente si può collocare l'inizio del Telerilevamento con la nascita e lo sviluppo della tecnica fotografica, che, unitamente alle proprietà delle ottiche di varia focale, ha permesso di estendere le possibilità di percezione per un osservatore, nonché di registrare in modo permanente le osservazioni.

Con il telerilevamento interessa raccogliere informazioni qualitative e quantitative provenienti da superfici poste in genere lontano dall'osservatore: lontano vuol dire, nel nostro caso, una distanza che può variare da qualche metro (proximal sensing) fino a migliaia di chilometri (remote sensing), come nel caso di osservazioni effettuate dai satelliti geostazionari (36000 Km).

L'energia elettromagnetica che trasporta le informazioni utili nel campo del telerilevamento è (nell'uso pratico) limitata alle bande del visibile, infrarosso, microonde; qualche esperienza viene condotta anche in banda ultravioletta, nonostante i limiti imposti dal comportamento dell'atmosfera in questa regione dello spettro elettromagnetico.

Qualunque superficie esterna di un corpo, a temperatura superiore allo zero assoluto¹ emette radiazioni elettromagnetiche proprie che dipendono dalla temperatura del corpo e dalla natura della superficie. La stessa superficie riflette, assorbe o si lascia attraversare dalle radiazioni elettromagnetiche provenienti dall'esterno; emissione, riflessione, assorbimento e trasmissione delle radiazioni sono fenomeni strettamente collegati fra loro, come vedremo.

Lo spettro delle radiazioni elettromagnetiche con i nomi delle varie bande in cui è suddiviso, appare nella fig. 1.1.

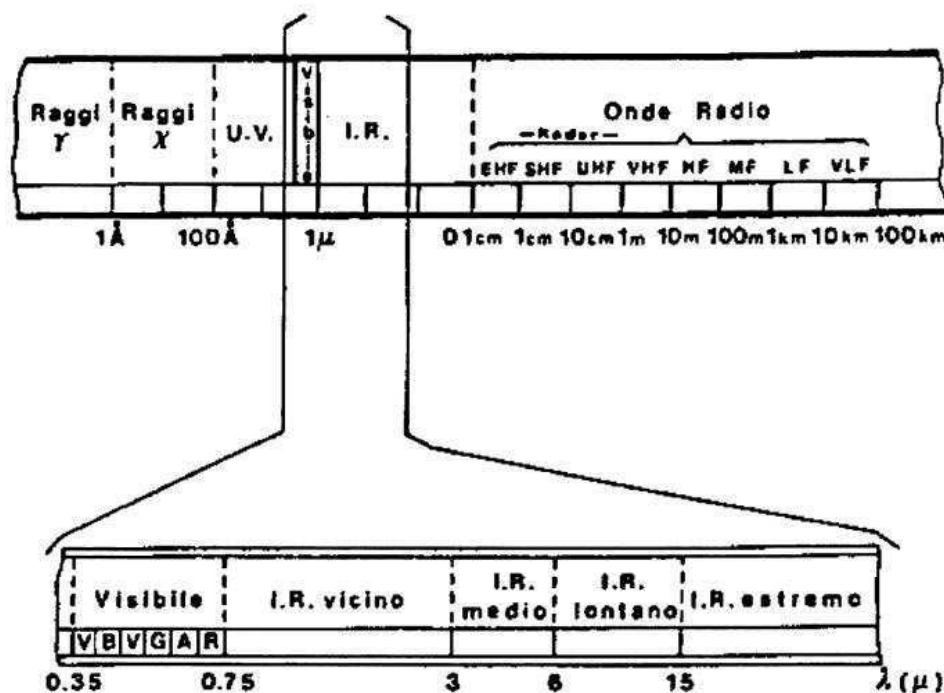


Fig. n. 1.1 - Spettro elettromagnetico

¹ Per zero assoluto si intende lo zero della scala Kelvin, pari a -273.14°C ($K = ^\circ \text{C} + 273.14$).

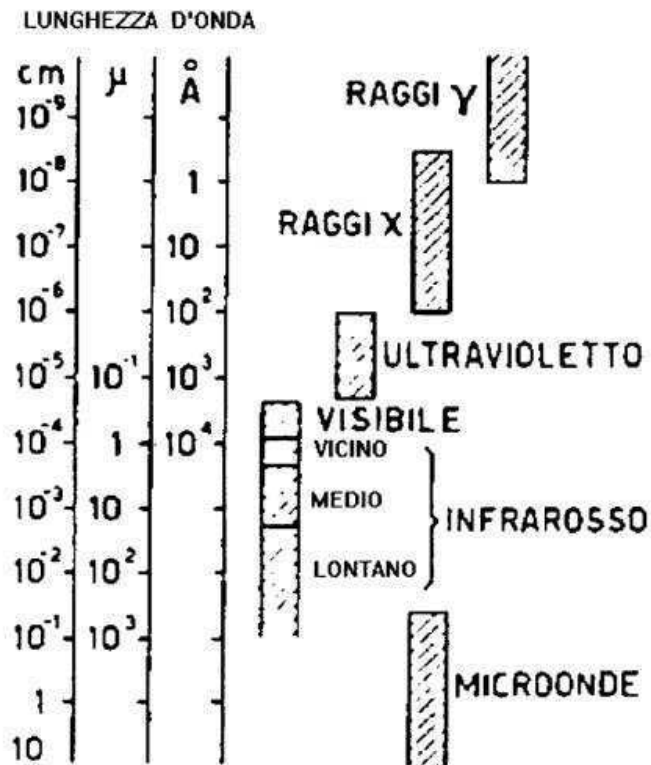


Fig. n. 1.2 - Lunghezze d'onda

Fra una superficie S qualunque e le radiazioni elettromagnetiche su essa incidenti (fig. 1.3) siano:

- E_i = l'energia incidente sulla superficie S
- E_r = l'energia riflessa dalla superficie S
- E_a = l'energia assorbita dalla superficie S
- E_t = l'energia trasmessa dalla superficie S

Tale energia incidente E_i verrà in parte riflessa, in parte assorbita, in parte trasmessa secondo lo schema seguente ed in funzione dei parametri sotto definiti:

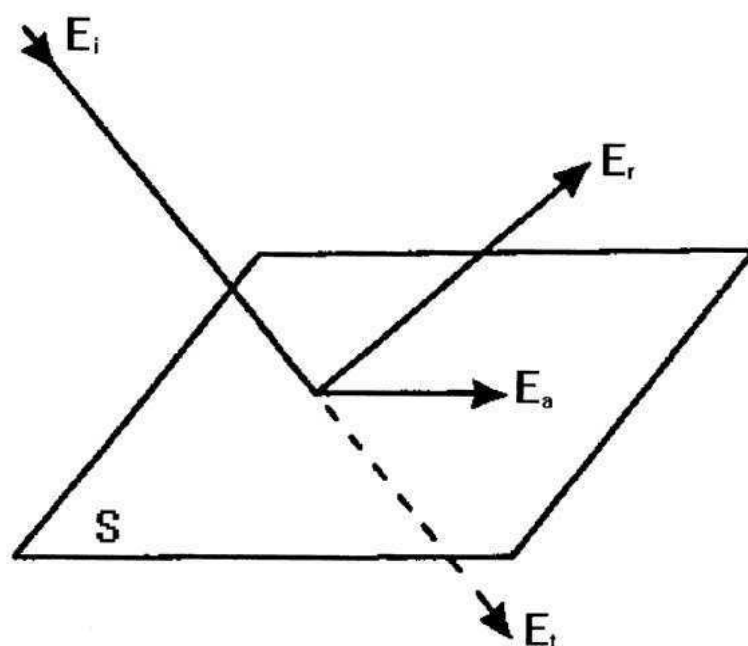


Fig. n. 1.3 - Composizione dell'energia

- 1) Coefficiente di riflessione $\rho = E_r / E_i$ $0 \leq \rho \leq 1$
- 2) Coefficiente di trasmissione $\tau = E_t / E_i$ $0 \leq \tau \leq 1$
- 3) Coefficiente di assorbimento $\alpha = E_a / E_i$ $0 \leq \alpha \leq 1$

Questi tre coefficienti o parametri dipendono strettamente dalla natura fisica delle superfici e dal loro grado di rugosità o di lucidatura. Ad esempio una medesima superficie può riflettere molto nella luce verde e poco nella luce rossa e blu: in questo caso la superficie illuminata con luce bianca, apparirà a noi come essenzialmente verde; la neve è bianca nel visibile e quasi nera nell'I.R. termico. Dalla composizione delle varie percentuali con cui una superficie riflette le luci blu, verde, rossa nasce il cosiddetto colore della medesima: una superficie bianca rifletterà bene sia il rosso che il verde e blu.

Per ogni superficie si può perciò costruire un grafico, che ci informa sulle capacità di riflessione in funzione della lunghezza d'onda della radiazione incidente: questo grafico, caratteristico di ogni superficie, si chiama firma o risposta spettrale (fig. 1.4a e 1.4b).

Per esempio:

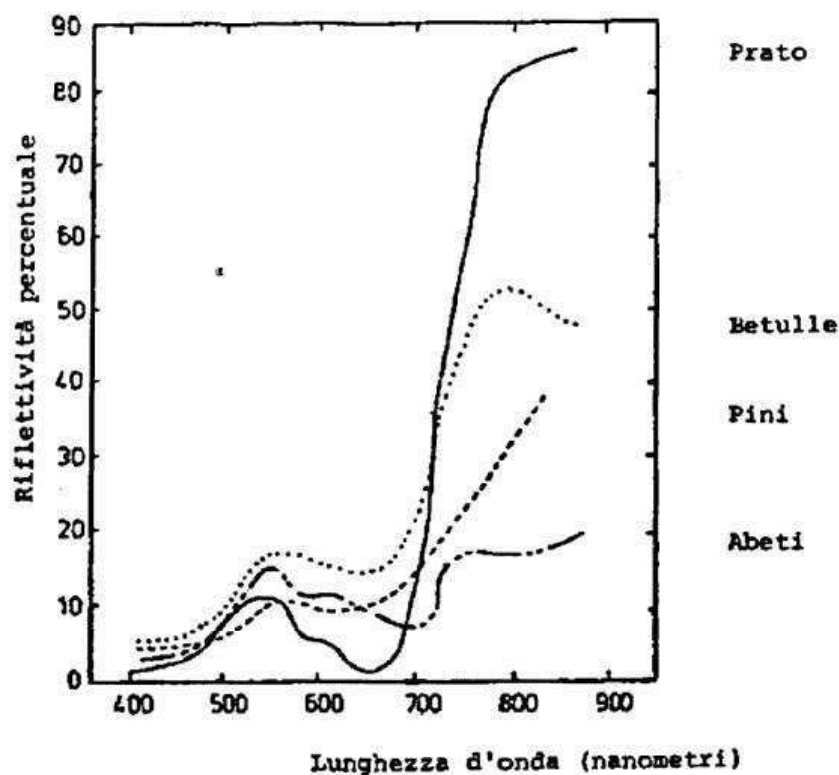


Fig. n. 1.4a - Capacità di riflessione in funzione della lunghezza di onda di alcuni vegetali

Questi grafici, corrispondenti al comportamento medio di alcuni vegetali indicano che tali superfici riflettono la luce soprattutto nelle bande del verde e dell'infrarosso, e poco nelle bande del blu e rosso.

Dall'equazione $\rho + \tau + \alpha = 1$ per corpi opachi si deduce che la vegetazione, riflettendo poco le radiazioni blu e rossa, ne assorbe una quantità maggiore, come in effetti accade in virtù dei meccanismi di fotosintesi.

Ovviamente non solo la vegetazione ha dei comportamenti riflessivi variabili, ma anche ogni altro tipo di superficie; negli esempi che seguono sono riportati degli esempi di suoli con diverso contenuto di umidità.

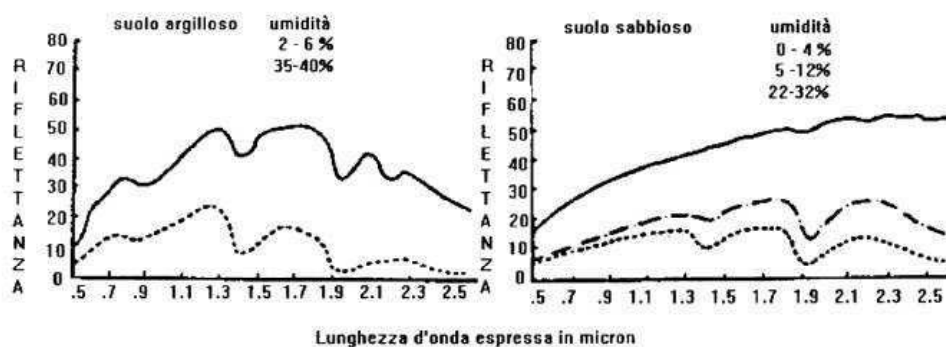


Fig. n. 1.4b - Capacità di riflessione dei suoli

Queste considerazioni stanno alla base delle tecniche sviluppate per il riconoscimento delle superfici con metodi tipici delle teleosservazioni:

è chiaro, infatti, che se è possibile esplorare in varie lunghezze d'onda la luce riflessa da una superficie, ottenendo un grafico come quelli riportati, e se si dispone di una statistica di comportamento spettrale sufficientemente vasta, si può pensare di riconoscere la natura dell'oggetto investigato.

A questo scopo sono stati sviluppati e costruiti strumenti atti a compiere misure in porzioni strette e contigue dello spettro elettromagnetico, suddividendo sia la luce visibile che quella invisibile in "bande", determinate dall'intervallo di lunghezza d'onda che coprono.

Una considerazione importante è questa: appare chiaro che più strette sono le bande considerate - più piccolo cioè è l'intervallo di lunghezza di onda che esse abbracciano - più precisa risulterà l'indagine, essendo esse numerose; d'altra parte diminuendo l'ampiezza delle bande lo strumento di misura raccoglie meno energia specifica, aumentando contemporaneamente il "rumore" associato all'informazione.

Si tratterà quindi per ogni problema di trovare un compromesso opportuno fra larghezza di banda, numero di bande, costo dello strumento, e risultati desiderati.

E' importante anche tenere presente che aumentando il numero delle bande aumenta di conseguenza il tempo necessario ad elaborare i dati, e quindi aumenta il costo globale.

I satelliti per le risorse terrestri, *non effettuano "fotografie" della superficie terrestre*, ma misurano la quantità di energia riflessa dai vari corpi presenti sulla superficie del suolo: il problema è di riuscire a stabilire una corrispondenza tra la quantità e la qualità dell'energia riflessa e la natura o lo stato dei corpi o delle superfici dai quali tale energia riflessa proviene.

E' questo il compito dell'analisi spettrale, ovvero il significato di "firma spettrale".

Quando l'energia elettromagnetica emessa dal sole colpisce un corpo sulla superficie del nostro pianeta, viene in parte assorbita e in parte riflessa. La riflessione può essere speculare. E' quello che avviene se la radiazione colpisce uno specchio d'acqua tranquilla.

La riflessione speculare non dà informazioni sulla natura della superficie riflettente.

La riflessione diffusa da superficie scabre o irregolari, invece, avviene uniformemente in tutte le direzioni e contiene informazioni "spettrali" sul colore e sulla natura della superficie riflettente.

Nel telerilevamento siamo interessati proprio alle misure delle caratteristiche di "riflessione diffusa" dei suoli e delle superfici in genere.

La percentuale dell'energia radiante incidente che viene riflessa (riflettanza) è determinata dalla struttura geometrica delle superfici, dalla natura e dalla composizione dei corpi (influiscono sulla riflettanza, ad esempio, il contenuto in acqua di un terreno o di una vegetazione oppure, al contrario, il contenuto di particelle solide in sospensione in un lago o in uno stagno) e dalla presenza o meno di pigmenti. Ad esempio la clorofilla assorbe fortemente l'energia radiante sulle bande di lunghezze d'onda intorno a $0.45\ \mu\text{m}$ (blu) e $0.65\ \mu\text{m}$ (rosso); riflette invece la radiazione verde, intorno alla lunghezza d'onda di $0.55\ \mu\text{m}$, che noi percepiamo visivamente come colore delle piante.

In pratica è possibile analizzare il valore della riflettanza di un corpo in relazione alle varie lunghezze d'onda dello spettro elettromagnetico, mediante uno strumento chiamato **spettroradiometro**, con il quale è possibile tracciare una curva "riflettanza-lunghezza d'onda", caratteristica (almeno nelle condizioni di laboratorio) di un determinato corpo e di una determinata superficie.

Nella realtà, tali curve di riflettanza spettrale permettono di riconoscere e individuare i tipi, le condizioni e le caratteristiche dei terreni, delle aree coperte da vegetazione e degli specchi d'acqua, da cui sono state rilevate, solo quando sono abbastanza caratteristiche e facilmente distinguibili dalle altre curve rilevabili da un dato territorio. In questi casi le curve possono essere chiamate "firme spettrali".

Vediamo alcuni esempi nelle fig. 1.5 e 1.6.

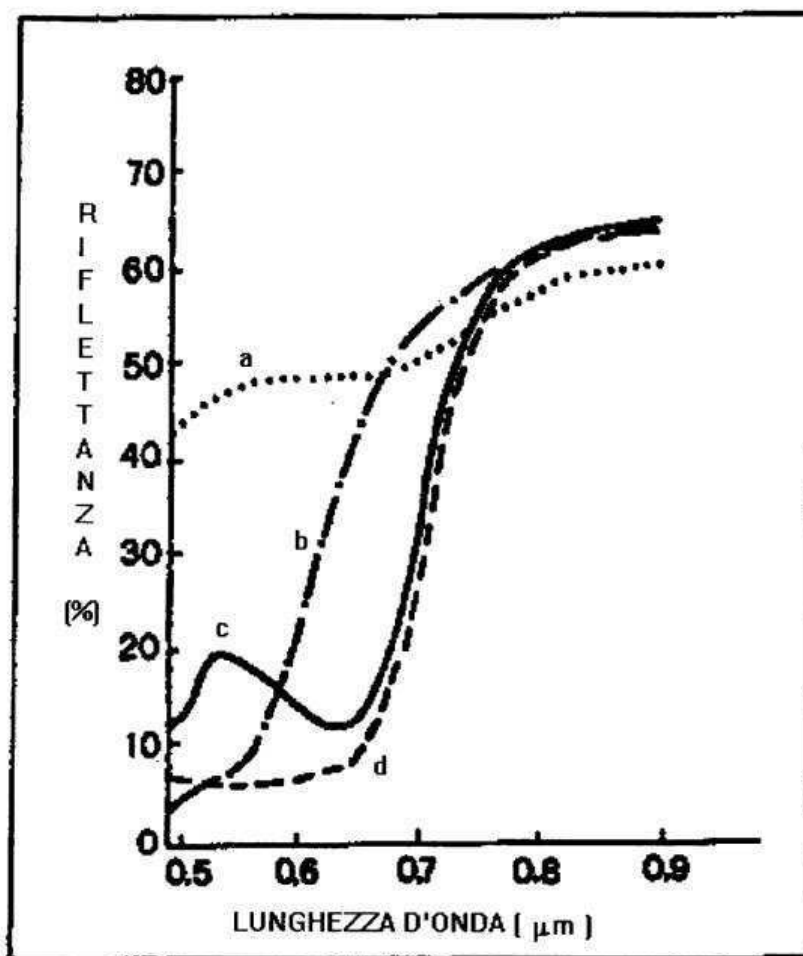


Fig. n. 1.5 - Variabilità della riflettanza delle foglie di Coleus (nell'intervallo di lunghezza d'onda tra 0,5 μm e 0,9 μm) in funzione della loro diversa pigmentazione: a, senza pigmenti; b, solo antociani; c, solo clorofilla; d, clorofilla e antociani.

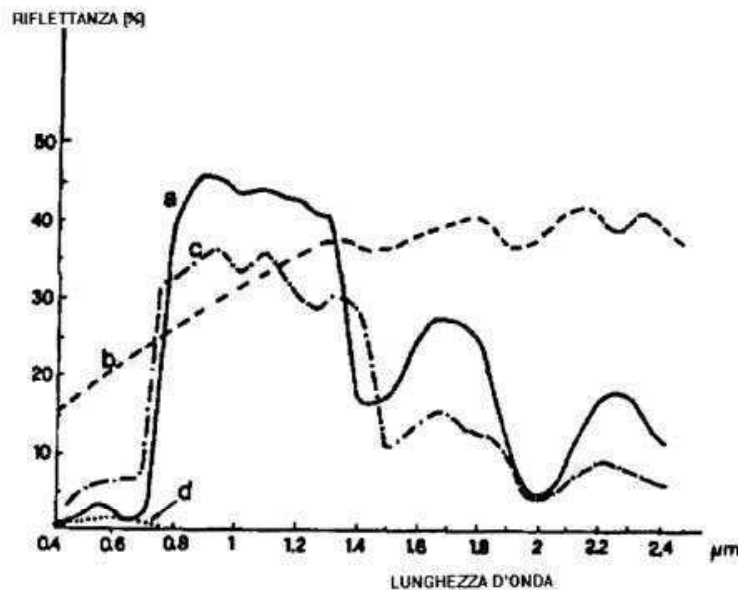


Fig.n. 1.6 - Curve tipiche di riflettanza spettrale del suolo erborato (a), del suolo asciutto grigio-bruno (b) del grano alto circa 60 cm (c) e dell'acqua (d).

Quando la raccolta di dati che utilizzi tecniche di telerilevamento avviene attraverso l'atmosfera terrestre, il risultato delle misure è influenzato dalle proprietà di trasmissione dell'atmosfera interposta fra superficie in osservazione e sensore.

In genere l'atmosfera non offre un contributo utile alla descrizione delle superfici teleosservate, agendo come un velo spesso ineliminabile, e limita le caratteristiche del sensore impiegato.

La trasparenza di questo "velo atmosferico" è molto diversa da zona a zona dello spettro elettromagnetico e, anche dove essa è più propizia alle riprese, i suoi valori possono subire rapidi mutamenti al variare delle condizioni meteorologiche.

Nella realtà operativa quindi le curve di riflettanza spettrale non sono così caratteristiche come quelle illustrate, bensì variano notevolmente a causa della variabilità delle condizioni ambientali degli strati di atmosfera attraversati (temperatura, pressione, umidità, ecc.).

Si deve tener conto, inoltre, del fatto che ai sensori che misurano le riflettanze spettrali giunge l'energia riflessa non da ciascun singolo punto di un campo osservato, ma dall'insieme dei punti che costituiscono l'area elementare di rilevazione, o "*pixel*" (*contrazione di Picture-Element*) nella direzione e nell'istante di osservazione.

Se la vegetazione, la composizione, il colore e l'umidità del suolo non sono uniformi entro il pixel osservato, la curva di riflettanza spettrale relativa a tale pixel è composita; in questo caso riconoscere una firma spettrale può essere difficile o impossibile.

I fattori che producono variazioni nelle curve di riflettanza spettrale possono essere statici, come la pendenza e l'esposizione del terreno, o dinamici. I fattori dinamici provocano differenze di risposta spettrale di una stessa area elementare nel corso del tempo; tra di essi: lo stadio fenologico delle colture erbacee, le loro condizioni fitosanitarie e il grado di copertura del terreno, l'umidità superficiale del suolo, la trasparenza atmosferica e la posizione del sole. Le variazioni dell'angolo di elevazione del sole provocano differenze nella riflettanza della "scena" vista dai sensori spaziali, che dipendono anche dal tipo di suolo. Ad esempio la riflettanza della sabbia è più sensibile alle variazioni di illuminazione solare che non le riflettanze dei terreni coperti da vegetazione.

2 - La correzione degli errori

I dati-immagine, acquisiti dalle stazioni riceventi, sono tagliati in blocchi variabili per ciascun tipo di satellite.

Come si è detto, i dati trasmessi a terra sono affetti da errori di vario tipo: sono errori geometrici e radiometrici.

La posizione di ogni pixel osservato deve essere esattamente individuata geograficamente.

Gli errori geometrici non consentono l'esatta individuazione di tale posizione.

Alcuni degli errori geometrici più importanti sono:

- errore causato dalla curvatura della terra
- errore dovuto alla rotazione della terra e al simultaneo spostamento della traccia a terra del satellite durante un ciclo di osservazione
- errori dovuti alle variazioni di assetto del satellite
- errore dovuto alle variazioni di altezza del satellite, che provoca un effetto di variazione di scala durante l'acquisizione di un'immagine.

Come si è detto, questi errori vengono compensati dalla stazione ricevente per mezzo di opportuni programmi di correzione computerizzati.

Malgrado ciò restano errori residui che devono essere eliminati o ridotti a cura dell'utente.

I numeri codificati relativi a un pixel generico non rappresentano sempre correttamente i livelli di energia relativi alle varie lunghezze d'onda, a causa degli errori radiometrici.

Una prima causa di errore è l'aggiunta di una componente di radiazione dovuta a diffusione di energia radiante da parte dell'atmosfera, che diminuisce il contrasto delle immagini.

Per eliminare questa che potremmo chiamare "componente luminosa dovuta all'aria", si suppone che i pixel cui si associa il valore minimo di riflettanza abbiano in effetti riflettanza zero.

Detto valore minimo viene, allora, sottratto da tutti i numeri che costituiscono l'insieme dei dati digitalizzati relativi a un dato territorio. Questa procedura è nota come rimozione della foschia.

La seconda causa di errore radiometrico è dovuta ai sensori, che non hanno "guadagni" identici costanti nel tempo.

In altre parole la "risposta" dello strumento può essere diversa pur in presenza di uguali livelli di riflettanza.

I dati che essi forniscono devono perciò essere calibrati a ogni ciclo di rilevazione mediante il confronto con una sorgente luminosa a radianza nota interna al sistema.

3 - Elaborazione delle misure, interpretazione e classificazione

I dati registrati su CCT sono pronti per essere elaborati in funzione delle utilizzazioni volute.

Poiché, come si è detto, nonostante la correzione automatica effettuata dalla stazione ricevente, permangono alcune distorsioni, prima di procedere all'elaborazione dei dati può essere eseguita una correzione ulteriore delle distorsioni sfruttando "punti di controllo" a terra (G.C.P. - Ground Control Point).

Detti punti possono essere costituiti ad esempio da incroci fra grandi strade, piccoli specchi d'acqua, o altri punti di riferimento facilmente riconoscibili su carte dettagliate del territorio considerato e sulle immagini satellitari.

Nel processo di correzione vengono individuati numerosi punti di controllo esprimendone le coordinate vere (note) e quelle risultanti dall'immagine satellitari.

Tali valori vengono sottoposti poi a un'elaborazione (analisi di riflessione) su computer in modo che i dati originali distorti siano posti in corrispondenza di quelli geometricamente corretti. Questa elaborazione consente un processo di "ricampionamento" che si esegue trasferendo i valori delle

riflettanze spettrali dalla configurazione “originale” dei pixel (come rilevati dal satellite) alla configurazione geometricamente corretta.

Dopo le correzioni degli errori, le immagini possono essere ulteriormente migliorate agendo sui valori digitalizzati dell’insieme dei dati. Si intuisce la delicatezza di tale operazione che “modifica” i dati ottenuti.

Nell’insieme di tali dati, infatti, alcuni valori si ripetono con elevata frequenza, altri con frequenza bassa o molto bassa.

Tenuto conto di ciò, essi si possono raggruppare in relazione alla loro frequenza, in modo da concentrare l’attenzione su quelli più frequenti e quindi più interessanti, e da scartare quelli poco frequenti e quindi poco rappresentativi.

Questa procedura porta ad esempio a contenere il numero dei “toni di grigio”, in modo che l’immagine risulti più chiara e più leggibile. A tal fine, una volta raccolti, corretti e migliorati i dati, si tratta di “estrarre” le informazioni che essi contengono.

Una procedura molto usata è quella detta di classificazione con apprendimento (supervised classification).

In tale processo si definiscono categorie di siti considerate come significative o interessanti e si determina, poi, la possibilità di distinguere le loro firme spettrali.

In un primo stadio del processo (o training) si registrano i valori numerici delle firme spettrali tipiche delle categorie considerate. In un secondo stadio di classificazione, i valori delle firme spettrali di ciascun pixel vengono confrontati con quelli tipici di ciascuna delle categorie suddette.

Ogni pixel viene assegnato alla categoria i cui valori sono più vicini.

Se i valori di un pixel sono lontani, oltre una certa soglia, da quelli di ogni categoria già definita, il pixel viene considerato come incerto e non assegnato ad alcuna classe.

In un terzo stadio l’attribuzione dei pixel alle varie categorie viene rappresentata su di una carta geografica, ad esempio attribuendo ai pixel certi valori convenzionali.

Si possono così ottenere carte dell’uso del suolo. Ovviamente risulta anche possibile ricavare informazioni statistiche varie sulla base delle attribuzioni fatte.

Potrà essere determinata l’area totale di tutti i pixel che appartengono a una certa classe.

Se la classe, ad esempio è quella dei pixel coperti da neve, si potrà dedurre la superficie totale coperta da neve di un certo bacino (fig. 3.1).

Potrà accadere che alcune delle classi spettrali individuate corrispondano in realtà a più categorie informative.

Se, ad esempio, una classe spettrale talora corrisponda a boschi di conifere e altre volte a boschi cedui, i due tipi di boschi non potranno essere distinti sulla base delle informazioni spettrali.

E’ evidente che se questo accade utilizzando i dati dell’MSS non è detto che accada anche con i dati del TM, che per le caratteristiche delle due bande di rilevazione (particolarmente strette e quindi selettive) e per la sua elevata capacità di risoluzione spaziale (30 m), può cogliere differenze che “sfuggono” a strumenti meno sensibili.

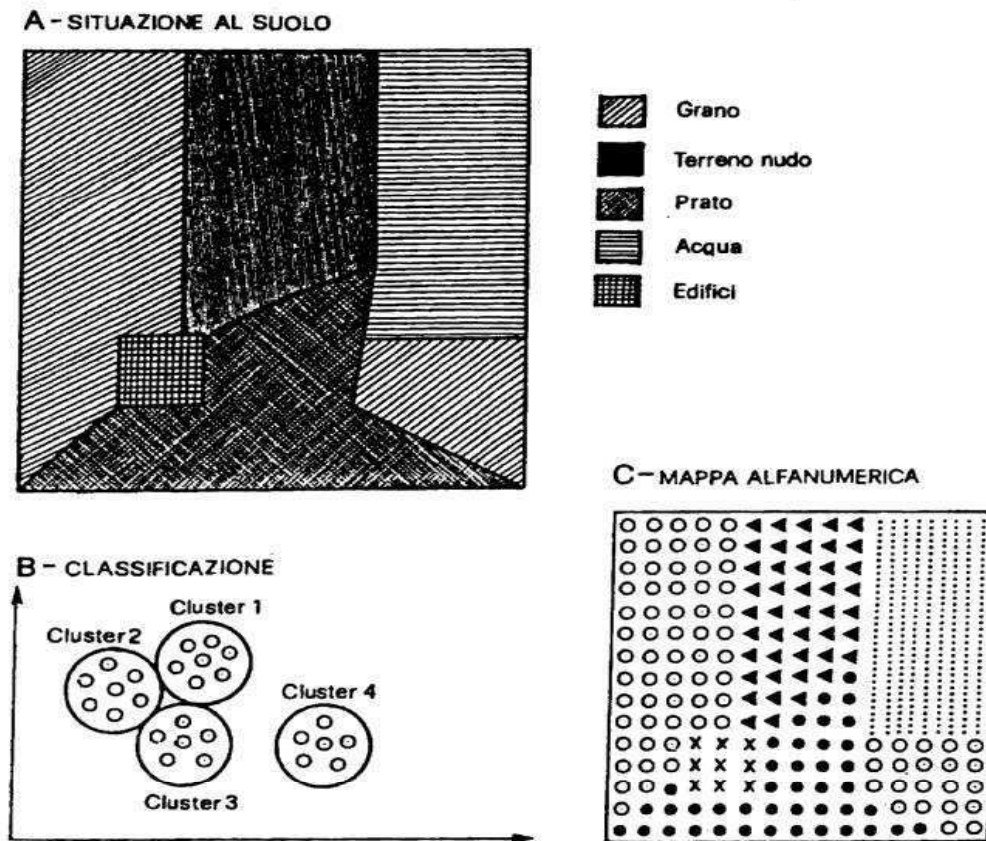


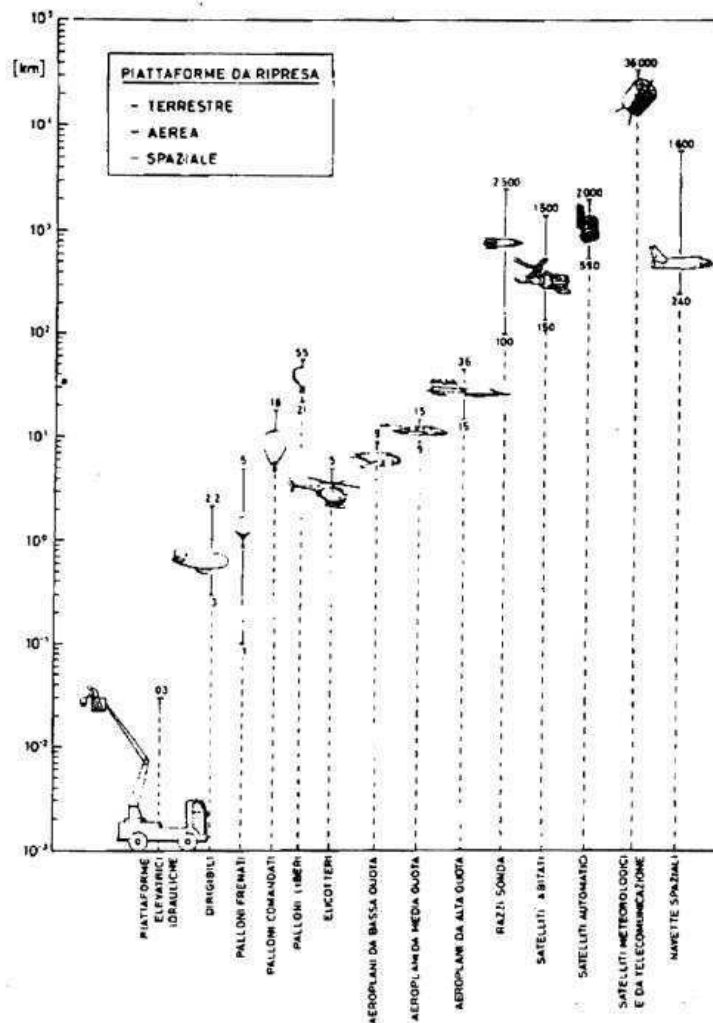
Fig. n. 3.1 - Carte di uso del suolo

4 - Le Piattaforme

Le tecniche di remote sensing possono essere applicate utilizzando strumenti montati su diversi tipi di supporto o, in linguaggio tecnico, piattaforme di osservazione, e ciascuna piattaforma, mobile o stabile, ha le proprie caratteristiche. Da un punto di vista generale tre sono i tipi di base di osservazione di interesse per il remote sensing: le piattaforme per le osservazioni al suolo, da mezzo aereo e da satellite (fig.4.1). L'aeroplano può essere impiegato:

- alle basse e medie altezze (1.500-3.000 m) per osservazioni di interesse locale e riguardanti aree limitate;
- ad alta quota per disporre di informazioni su aree estese.

Particolarmente adatti a una osservazione sinottica di aree vaste sono soprattutto i satelliti.



termico, con risoluzione spaziale minore (lato dei pixel di circa 240 m), venne aggiunta solo a partire dalla missione Landsat 3.

Il secondo sensore su Landsat 1,2 e 3 era il Return Beam Vidicom (RBV).

Su Landsat 1 e 2 il RBV aveva tre bande (la 1 nel blu, la 2 nel rosso e la 3 nell'infrarosso) con la medesima risoluzione spaziale del MSS. Sul Landsat 3 erano alloggiati due RBV identici e affiancati, ciascuno con un'unica banda pancromatica. La loro risoluzione spaziale era di 40 m e ciascun sensore aveva uno swath di 98 km: i due, affiancati con una piccola sovrapposizione, fornivano dunque uno swath di 185 km (fig. 5.1). In ordine crescente rispetto alle bande 1, 2 e 3 del RBV le bande MSS furono chiamate 4, 5, 6, 7 (e 8) nelle prime tre missioni Landsat. A partire da Landsat 4 il RBV venne sostituito dal Thematic Mapper (TM); quindi nella missione fallita del Landsat 6 si passò all'Enhanced Thematic Mapper (ETM) e infine, nel Landsat 7, all' ETM+. Nelle missioni più recenti le bande MSS sono state rinominate da 4, 5, 6, 7 a 1, 2, 3 e 4 rispettivamente.

Le informazioni rilevate dagli strumenti di bordo e codificate in sequenze di cifre binarie da un convertitore analogico-digitale vengono poi trasmesse a terra e ricevute da una stazione al suolo in contatto radio diretto in linea d'aria con il satellite.

Le informazioni relative alle zone non coperte dalle stazioni vengono registrate su nastro a bordo del satellite e trasmesse alle stazioni riceventi all'atto del passaggio del satellite nella loro area di ricezione. Questo per ciò che concerne i primi tre satelliti del programma.

Il Landsat 4 e 5 grazie alla loro più sofisticata strumentazione di rilevamento dispongono di un sistema migliore anche per la trasmissione a terra dei dati; usufruiscono di un satellite appoggio posto su di un'orbita alta e geosincrona cui trasmettono i dati rilevati.

L'altezza del satellite appoggio (36.000 Km) è tale che esso può ritrasmettere direttamente i dati a tutte le stazioni di un intero emisfero.



Fig. n. 5.1 - Fasce esplorate dal satellite durante un ciclo di orbita

Le stazioni in attività sono nove:

- una in Italia nella piana del Fucino;
- una in Svezia;
- due in Canada;
- due in Sud America (Brasile e Argentina);
- due in Asia (Giappone e India);
- una in Australia.

Ciascuna stazione riceve i dati relativi a un territorio circolare di raggio 3.200 Km, con centro nelle stazioni.

Nel loro moto intorno alla terra i satelliti Landsat con appositi strumenti (sensori) tengono costantemente sotto osservazione la superficie terrestre. Tali strumenti, come si è già accennato, non riprendono immagini fotografiche del suolo, bensì misurano momento per momento la riflettanza dei corpi, da cui è possibile risalire alle firme spettrali. Essi in sostanza acquisiscono una serie di dati, di valori, e non di immagine.

L'esplorazione è fatta con il sistema a scansione: mentre il satellite avanza, uno specchio oscillante esplora istante per istante delle strisce perpendicolari alla direzione del moto, da un estremo all'altro della fascia larga 185 Km (fig. 5.2).

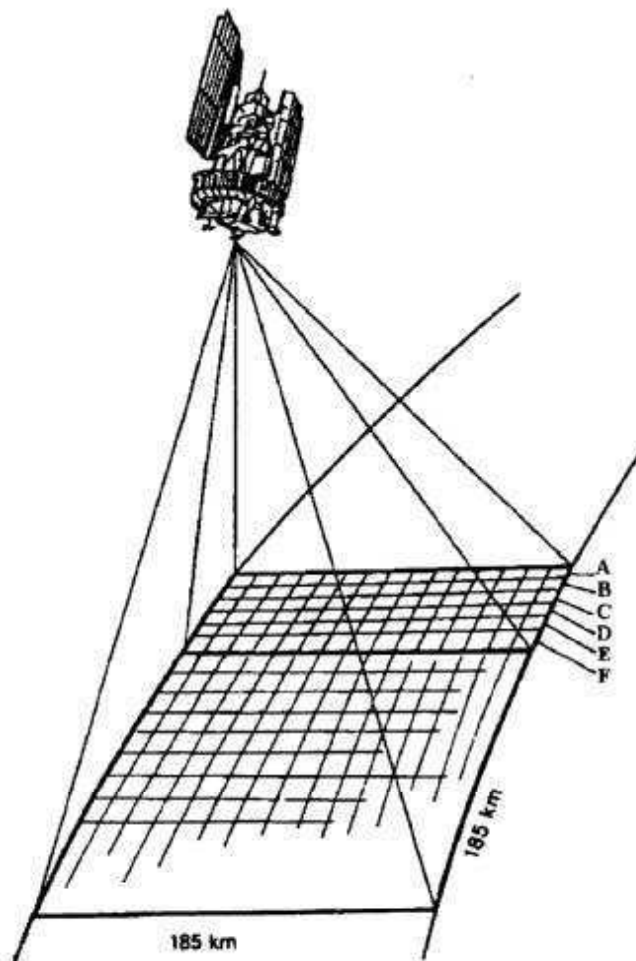


Fig. n. 5.2 - Esplorazione eseguita con il sistema di scansione

Scanner Multispettrale (MSS)

L'MSS (Multi-Spectral Scanner) è caratterizzato da 4 bande spettrali (fig. 5.3), scelte nell'intervallo dello spettro elettromagnetico $0,3 \div 1,3 \mu\text{m}$, in cui è più basso l'assorbimento atmosferico, e massima l'energia irradiata dal sole.

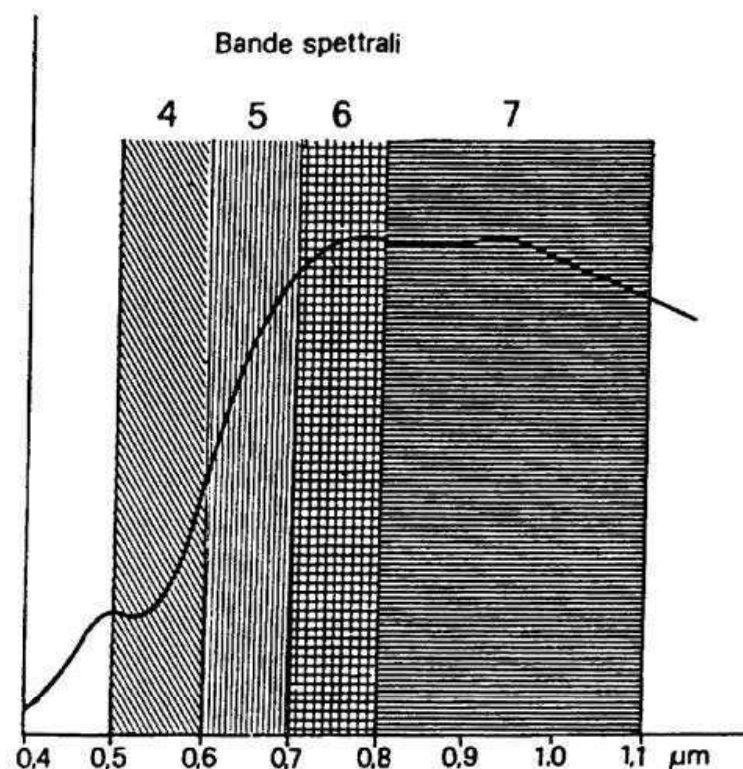


Fig. n. 5.3 - Bande di frequenza nell'intervallo dello spettro elettromagnetico $0.3 \div 1.3 \mu\text{m}$

Le parti fondamentali sono costituite da un sistema 24 rilevatori, 6 per ognuna delle 4 bande spettrali, che permette la ripresa di 6 strisce parallele e contigue a ogni oscillazione dello specchio. I rilevatori sono dei fotomoltiplicatori per le prime tre bande e fotodiodi al silicio per la quarta. La radiazione proveniente dalla scena che colpisce i rilevatori è trasformata in un segnale elettrico che viene digitalizzato a 6 bit rappresentando numeri indice da 0 a 63. La dimensione dei pixel (ossia la più piccola area che i sensori possono distinguere) è di 57 metri (ovest-est) per 79 metri (nord-sud). Le bande spettrali di funzionamento dei rivelatori che permettono la ripresa multispettrale della stessa scena sono così indicate:

MSS 4: 0,5 - 0,6 micron (verde, giallo).

Utilizzata soprattutto per studi su corpi idrici, a motivo della buona penetrazione.

In questa banda si ha però un notevole effetto di diffusione atmosferica.

MSS 5: 0,6 - 0,7 micron (arancio, rosso).

In questa banda, compresa in una delle regioni di assorbimento della clorofilla, si ha una buona separazione fra zone vegetali e terreni scoperti, e fra i diversi tipi di vegetazione. Banda di riferimento per controlli geometrici, poiché la diffusione atmosferica è minore che nel resto del visibile.

MSS 6: 0,7 - 0,8 micron (rosso, infrarosso).

La penetrazione in acqua è molto ridotta, mentre mette in evidenza le caratteristiche morfologiche e la tessitura dei suoli. Per la sua alta correlazione con la banda 7 permette numerosi procedimenti di elaborazione.

MSS 7: 0,8 - 1,1 micron (infrarosso).

Utilizzata per la separazione fra corpi rigidi e terra, per l'alto assorbimento dell'acqua: il reticolo idrografico appare nettissimo in immagini riprese in questa banda. La vegetazione è molto riflettente a queste lunghezze d'onda e le latifoglie maggiormente delle conifere. Le nuvole ed il manto nevoso hanno i valori vicini alla saturazione, anche se la neve bagnata riflette meno di quella asciutta.

MSS 8: 10,4 - 12,6 micron (infrarosso termico).

Questa banda è stata utilizzata solo sul Landsat 3, quale esperimento in vista del suo utilizzo in modo operativo per il sistema TM dei Landsat successivi.

Le informazioni raccolte in questa banda sono legate allo stato termico delle superfici.

Nell'ambito di ogni banda, i sensori possono rilevare 256 valori della riflettanza.

Ciascuno di questi livelli ricevuto come "segnale" elettrico, viene trasformato in una "cifra binaria" (digit), codificata secondo il linguaggio del calcolatore, e come tale appunto può essere gestito dal computer. A tale codificazione provvede un convertitore analogicodigitale prima che queste sequenze di cifre vengono trasmesse a terra. I dati acquisiti, poiché sono affetti da errori di vario tipo (di cui si tratterà più dettagliatamente in seguito) vengono sottoposti ad accurate e complesse procedure automatiche di correzione messe in atto dalla stazione di ricezione, che provvede poi a registrare i dati su nastro magnetico (CCT: computer compatible tape) adatto all'impiego su calcolatore.

In definitiva, dopo che il satellite ha ripreso i dati di riflettanza, e dopo che i segnali rilevati sono stati trasmessi a terra e corretti, l'elaboratore fornisce, per ogni area rilevata, quattro immagini, una per ogni banda (4 - 5 - 6 - 7). In ciascuna di queste immagini o "mappe", per ogni pixel rilevato a terra esiste un'area corrispondente sulla mappa stessa, identificata da un tono di grigio corrispondente al livello di riflettanza di "quel" pixel nella banda considerata. In teoria, poiché i sensori, come si è detto, possono rilevare 256 valori di riflettanza, dovrebbero esserci 256 possibili toni di grigio: in pratica i livelli di riflettanza sono raggruppati in modo da utilizzare meno di una decina di toni di grigio.

Queste mappe consentono un certo e diverso tipo di lettura della realtà al suolo: ogni banda si è visto essere adatta a evidenziare in modo più specifico alcuni aspetti della realtà al suolo.

Le scene riprese dal Landsat si possono anche stampare in "falso colore" sovrapponendo i segnali di tre bande: normalmente, della banda 4 stampati in blu, della banda 5 in verde e della banda 7 in rosso, simulando così l'effetto che si ottiene con le pellicole infrarosse a colori.

Ne deriva, ad esempio, che la vegetazione rigogliosa appare in rosso brillante per l'effetto dell'alta riflettanza in banda 5. Questo tipo di immagini in falso colore è la più comunemente usata.

Per i satelliti Landsat 4 e 5, le bande spettrali 4,5,6,7 vengono indicate MSS 1, MSS 2, MSS 3, MSS 4 (fig. 5.4).

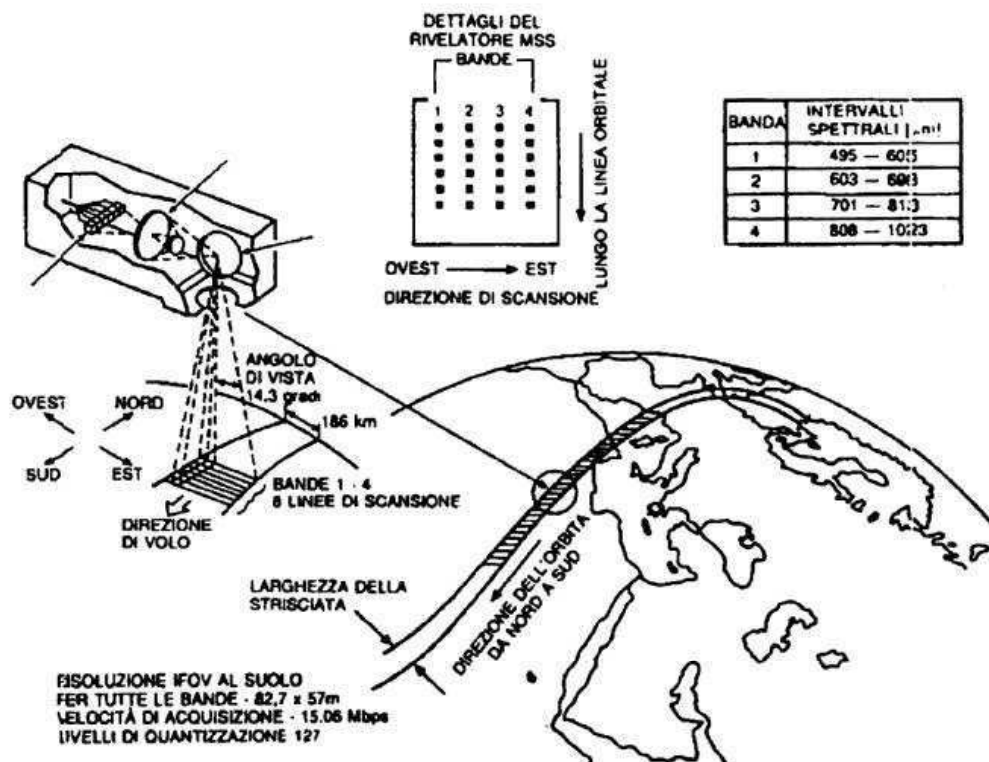


Fig. n.5.4 - Caratteristiche e schema di funzionamento del MultiSpectral Scanner

Return Beam Vidicon (RBV)

Sui satelliti Landsat 1, 2 e 3 come sensore di ripresa delle immagini è stato utilizzato anche l'RBV. Il Return Beam Vidicon è un particolare tipo di tubo Vidicon che sfrutta il fascio di elettroni per creare immagini.

Il sistema RBV non contiene pellicola, ma le immagini vengono riprese attraverso un'ottica con otturatore centrale e immagazzinate su una superficie fotosensibile.

Questa superficie viene quindi analizzata a scansione lenta in forma *raster* mediante un fascio di elettroni o pennello elettronico, per produrre un segnale video come in una telecamera.

Poiché il sistema RBV contiene un reticolo di riferimento nel piano dell'immagine, le immagini finali contengono dei riferimenti in posizioni ben definite.

Per il Landsat 1 e 2 il sistema RBV era di tipo multispettrale costituito da tre Vidicon, con ottica con lunghezza focale di 126 mm, con una risoluzione al suolo di circa 80 m; le bande spettrali erano così definite:

Banda 1 da 0,475 micron a 0,575 micron

Banda 2 da 0,580 micron a 0,680 micron

Banda 3 da 0,680 micron a 0,830 micron

Per questa ragione le bande spettrali utilizzate dal sistema MSS furono denominate Banda 4, 5, 6 e 7.

Sul Landsat 3 il sistema RBV multispettrale venne sostituito da un sistema RBV pancromatico operante da 0,505 micron a 0,750 micron, costituito da due Vidicon che esploravano zone a terra adiacenti, con un'ottica focale doppia (236 mm) e con una risoluzione al suolo di circa 40 metri.

Thematic Mapper (TM)

Nella seconda generazione di satelliti per l'osservazione e lo studio delle risorse terrestri è stato installato a bordo un nuovo tipo di sensore, il TM, caratterizzato da 7 bande spettrali.

Le parti fondamentali che lo costituiscono sono (vedi fig. 5.5):

- uno specchio che oscilla attorno ad un asse parallelo alla direzione del volo, realizzando così una scansione trasversale della scena. La frequenza di oscillazione è di 7 hertz, cioè un periodo di 142 millisecondi. La ripresa viene effettuata sia durante la semioscillazione dello specchio da Ovest verso Est, sia nella fase di **ritorno**, la striscia esplorata al suolo durante ognuna delle due fasi di oscillazione è di 185 km;
- un sistema di rivelatori, 16 per ognuna delle bande spettrali da 1 a 5 e per la 7 e con un campo istantaneo di vista (IFOV) che da una risoluzione a terra, o pixel, di 30 m per 30 metri, mentre per la banda 6 i rivelatori sono 4 con una risoluzione a terra di 120 metri

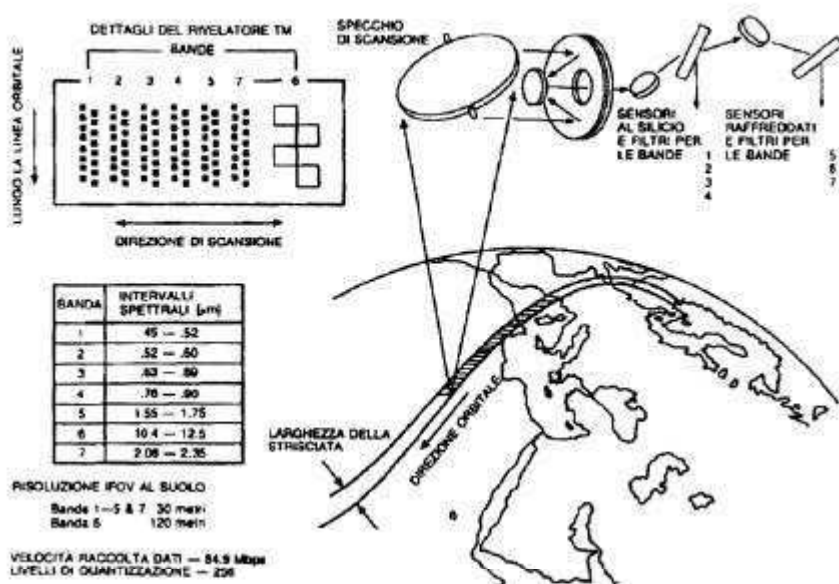


Fig. n. 5.5 - Caratteristiche del sistema di ripresa Thematic Mapper

La radiazione proveniente dalla scena che colpisce i rivelatori è trasformata in un segnale elettrico che viene digitalizzato a 8 bit, rappresentando numeri indice da 0 a 255.

Le bande spettrali di funzionamento dei rivelatori permettono la ripresa multispettrale della stessa scena.

ETM+

L' ETM+, operativo su Landsat 7, ha più o meno le stesse caratteristiche del TM; però la risoluzione spaziale della sua banda 6 è di 60 m (anziché di 120 m) ed è presente anche una banda pancromatica con 15 m di risoluzione. Fra i sensori più ampiamente utilizzati, TM ed ETM+ sono gli unici che coprono i tre colori di base (RGB) nello spettro del visibile: le relative bande possono quindi essere utilizzate per produrre immagini a colori reali.

5.2 - Satellite SPOT

I satelliti SPOT, concepiti e disegnati dal CNES (Centre National d'Etudes Spatiales) francese, sono costruiti in Francia in associazione con alcuni paesi europei; consiste essenzialmente di un sistema per l'osservazione della terra e di stazioni per il ricevimento dei dati. Il lancio, del primo della serie, è avvenuto il 22 Febbraio 1986. Il primo satellite SPOT (fig. 5.2.1), osserva la terra in tre bande nella porzione del visibile e del vicino infrarosso (con una risoluzione al suolo dell'ordine di 20 metri), e in una banda spettrale più larga, pancromatica (con una risoluzione al suolo dell'ordine di 10 metri).

Una delle caratteristiche chiave degli strumenti SPOT è la possibilità di visione al di fuori del Nadir; gli strumenti possono cioè vedere lateralmente da ambo le parti rispetto alla traccia. Le possibilità introdotte da questa caratteristica sono di poter incrementare le coperture di una certa area ad intervalli compresi tra 1 e 26 giorni. Questo permetterà il monitoraggio localizzato di fenomeni che evolvono in un tempo molto rapido. Un'altra possibilità è la raccolta, durante passaggi successivi del satellite, di coppie stereoscopiche di immagini di una certa area.

Le principali applicazioni ed utilizzazioni delle immagini continuamente inviate dal sistema SPOT sono:

- studi sull'uso del suolo (uso agricolo, pianificazione, ecc.)
- accertamento e studio delle risorse rinnovabili (agricoltura, boschi, biomassa, ecc.)
- lavori cartografici a media scala (1:100.000÷1:50.000)
- sviluppo di nuovi tipi e più frequenti aggiornamenti delle mappe tematiche a scale intorno a 1 : 50.000 e minori.

nadirale al suolo.

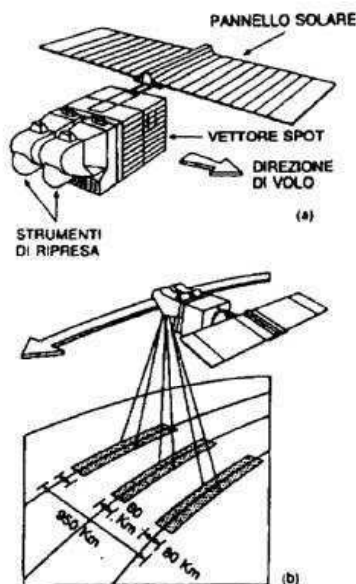


Fig. n. 5.2.1 - (a) Il satellite francese SPOT e (b) schema della visione nadirale dello strumento HRV pancromatico

Il carico utile del primo satellite SPOT è costituito da due strumenti identici HRV (Haute Résolution Visible) e da un apparato di registrazione dei dati su nastro magnetico e di trasmissione

a terra. Lo strumento HRV, che è composto da un telescopio con focale 1082 mm e $f/3.5$, può funzionare in due modi diversi nello spettro visibile e nell'infrarosso vicino:

- in modo “pancromatico” che corrisponde ad un’osservazione in una banda spettrale larga
- in modo “multibanda” che corrisponde ad una osservazione in tre bande spettrali più ristrette.

Il passo di campionamento scelto per lo strumento corrisponde, in presa nadirale, ad un elemento osservato al livello del suolo (pixel) di 10 m di lato nel primo caso, e di 20 m di lato nel secondo caso. Questa scelta è adatta all’osservazione di quelle particelle agricole di piccole dimensioni che sono comuni in molti Paesi e risponde nello stesso tempo alle esigenze cartografiche attuali.

Uno specchio piano orientabile mediante telecomando è disposto a valle dello strumento di presa vero e proprio; esso consente di spostare l’asse di presa in un piano perpendicolare all’orbita (fig. 5.2.2).

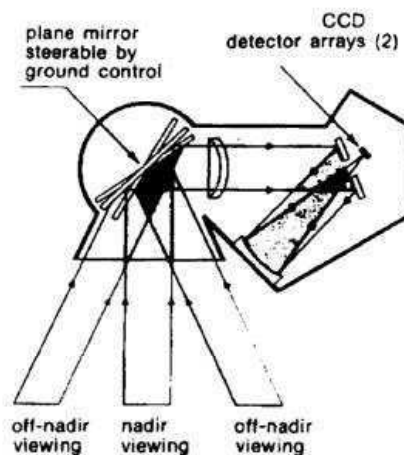


Fig. n. 5.2.2 - Specchio orientabile mediante telecomando

Il satellite SPOT 5, lanciato con successo il 4 maggio 2002 dalla Guiana Space Centre, Kourou, è operativo dal 16 luglio 2002. Esso continua la missione dei precedenti satelliti SPOT 1 – 4 che, lanciati a partire dal 1986, sono tuttora in orbita. Il satellite Spot 5 monta a bordo strumenti tecnologicamente innovativi:

- HRG (High Resolution Geometric)
- HRS (High Resolution Stereoscopic)

La risoluzione geometrica dei dati del satellite Spot 5 è fino a 4 volte maggiore di quella dei dati Spot 1 → 4:

- 5 metri e 2.5 metri del pancromatico, invece di 10 metri;
- 10 metri del multispettrale, invece di 20 metri.

La dimensione delle immagini è la stessa dei satelliti Spot 1-4: 60 km x 60 km; oppure 60 km x 120 km quando sono usati i due strumenti HRG di cui il satellite dispone.

L'accuratezza di localizzazione, senza punti di controllo, è inferiore ai 50 metri, contro i 350 metri dei satelliti Spot 1-4.

Questi miglioramenti risultano particolarmente vantaggiosi per le applicazioni che richiedano scale di riferimento da 1:25.000 a 1:10.000.

Grazie al nuovo strumento HRS a bordo dello Spot 5, sarebbero inoltre possibili, solamente per la produzione di modelli digitali del terreno, acquisizioni simultanee di stereocoppie con una larghezza (swath) eccezionalmente ampia: difatti le stereocoppie sono di dimensione 600 km x 120 km. Questi prodotti offrono un'eccellente accuratezza senza punti di controllo, indicata in:

- accuratezza in altezza (z): migliore di 10 metri;
- accuratezza di localizzazione: migliore di 15 metri.

Caratteristiche specifiche del modo pancromatico PA:

- risoluzione a terra:
5 metri/2.5 metri in modalità SUPERMODE;
- radiazione misurata:
banda unica dello spettro elettromagnetico corrispondente al visibile, con esclusione del blu (0.51- 0.73 micron).

Il processamento denominato SUPERMODE offre un miglioramento della risoluzione geometrica, in modalità pancromatica, da 5 metri a 2.5 metri, e permette anche di produrre, attraverso la combinazione con le bande multispettrali a 10 metri di risoluzione, immagini a colori a 2.5 metri di risoluzione. Ideato dal CNES, l'Agenzia Spaziale Francese, produce un'immagine a 2.5 metri di risoluzione attraverso due immagini pancromatiche, acquisite simultaneamente, a 5 metri di risoluzione. Il processo SUPERMODE viene effettuato in parte a bordo del satellite, in parte a terra: l'immagine viene acquisita, a 5 metri, da due sensori, con uno sfalsamento tra di loro (offset) nel piano della focale di mezzo pixel, sia orizzontalmente che verticalmente. Le due immagini a 5 metri di risoluzione, attraverso successivi processi di interpolazione, producono un'immagine con 2.5 metri di risoluzione effettiva.

Caratteristiche specifiche del modo multispettrale XI:

- risoluzione a terra: 10 metri (la banda 4 ha una risoluzione effettiva di 20 metri, ma viene fornita ricampionata a 10 metri, per uniformità con le rimanenti tre bande);
- spettro di radiazione osservata: 4 bande spettrali (corrispondenti al verde, rosso, infrarosso vicino e infrarosso medio dello spettro elettromagnetico).

5.3 - EROS

La missione EROS, composta da 6 satelliti pancromatici dedicati all'osservazione del territorio ad alta ed altissima risoluzione, rappresenta un vero e proprio atout nel campo del Remote Sensing di ultima generazione. La missione di proprietà di ImageSat International Ltd vede attualmente in orbita il primo dei 6 satelliti, ossia EROS A1, lanciato il 5 Dicembre 2000 dal poligono di Svobodny in Siberia. Il satellite posto in orbita circolare eliosincrona ad una altitudine di 480 km orbita intorno alla terra circa 15 volte al giorno trasmettendo i dati delle immagini in tempo reale a 16 stazioni riceventi nel mondo che rappresentano la più folta schiera di GRS dedicate ad una singola missione. Il satellite monta a bordo un sensore pancromatico CCD (*Charged Couplet Device*) di tipo *pushbroom* con risoluzione radiometrica a 11 bit con 2.048 livelli di grigio (risoluzione geometrica 1,8/1,0 m). Le immagini con la risoluzione a 1 m vengono spedite direttamente dal satellite a terra, e non sono oggetto di una elaborazione successiva (si ottengono

tramite un uso più ampio dell'*integration time*): non sono quindi da considerarsi prodotti a valore aggiunto, ma una diversa risoluzione dell'immagine. Un particolare estremamente significativo della missione EROS, che con le sue 16 stazioni di ricezione copre il 75% delle terre emerse, è rappresentato dal registratore a stato solido montato a bordo dei satelliti ed in grado di garantire una copertura completa di tutta la Terra.

IPT informatica per il territorio S.r.l. è l'azienda responsabile in esclusiva delle acquisizioni del satellite EROS A1 nel bacino del Mediterraneo con la sua Stazione Satellitare Multimissione ubicata in Sardegna.

I satelliti EROS di classe B monteranno a bordo sia sensori pancromatici che multispettrali con risoluzione radiometrica a 10 bit con 2.048 livelli di grigio e risoluzione geometrica a 0,69 m nel pancromatico e 2,76 m nel multispettrale.

L'intera missione prevede 6 satelliti operativi in orbita entro il 2005 con un tempo di rivisitazione dello stesso sito di due satelliti al giorno.

I satelliti EROS sono estremamente manovrabili a fronte del ridotto peso, circa 250 kg al momento del lancio, e possono essere puntati e stabilizzati in breve tempo sul sito di interesse del cliente a partire da una ripresa al nadir, perpendicolarmente alla superficie, fino a raggiungere un'inclinazione massima di ripresa pari a 45° con azimuth selezionabile nei 360°. La capacità di osservazione inclinata del satellite permette già solo con il satellite EROS A1 in orbita di osservare qualsiasi sito sulla terra da tre a quattro volte per settimana. La presenza di 16 stazioni di ricezione, dotate di una catena di processamento completa e di un catalogo locale per l'archiviazione e la consultazione via interfacce Web delle immagini ricevute, permette di ridurre i tempi di consegna dal momento dell'acquisizione a massimo tre giorni solari per qualunque scena acquisita in qualunque parte del mondo. Il sistema comprende inoltre un archivio centrale di backup presso la ImageSat International per garantire la continuità nel servizio di accesso e distribuzione delle immagini.

Infine, è disponibile un'abbondante offerta di immagini d'archivio a prezzi "contenuti".

5.4 - IKONOS

Il satellite IKONOS è in orbita dal settembre 1999, ed è operativo dall'inizio del 2000. Sul satellite sono montati due sensori, un sensore pancromatico ed uno multispettrale. Il primo ha una risoluzione geometrica al suolo di 1 m ad 11 bit (2.048 livelli di grigio) e acquisisce nella banda spettrale dei 0.45-0.90 μm .

Il sensore multispettrale ha invece una risoluzione geometrica al suolo di 4 m, e 4 bande ad 11 bit (2048 livelli).

I prodotti IKONOS sono distribuiti in tre diverse modalità: GEO, STEREO ed ORTORETTIFICATI. I prodotti GEO sono geometricamente corretti ed ortorettificati. Il livello standard di accuratezza è di 25 m (RMSE), escludendo gli effetti provocati da spostamenti del terreno. Possono essere distribuiti in 3 diverse opzioni:

- PAN: Dato Pancromatico (Bianco e Nero) con risoluzione geometrica al suolo di 1 metro.
- MSI: Dato Multispettrale (4 bande) con risoluzione geometrica al suolo di 4 metri.
- PSM: Dato Pan Sharpened (3 bande fuse con il pancromatico) con risoluzione geometrica al suolo di 1 metro.

I prodotti STEREO sono ottenuti da stereo-coppie acquisite nello stesso passaggio orbitale, minimizzando così le variazioni di luminosità; le immagini sono riproiettate e ricampionate a 1 metro in modo da permetterne una più facile visualizzazione. Sono distribuiti in due differenti livelli di accuratezza:

- Standard Stereo: 12 metri di accuratezza orizzontale (RMSE) e 13 metri di accuratezza verticale (RMSE);
- Precision Stereo: 1 metro di accuratezza orizzontale (RMSE) e 2 metri di accuratezza verticale (RMSE).

Il processo di ortorettifica rimuove le distorsioni delle immagini, causate dalla geometria dell'acquisizione e dalla variabilità del suolo, utilizzando modelli digitali del terreno e punti di controllo al suolo. I prodotti ORTORETTIFICATI vengono distribuiti secondo diversi livelli di accuratezza e nelle tipologie (PAN, MSI, PSM).

- REF (Reference): 12 metri (RMSE) di accuratezza orizzontale; cartografia di riferimento: 50.000 NMAS (National Map Accuracy Standards).
- PRO: 5 metri (RMSE) di accuratezza orizzontale; cartografia di riferimento: 1: 10.000 NMAS.
- PRE (Precision): 2 metri (RMSE) di accuratezza orizzontale; cartografia di riferimento: 1: 5.000 NMAS.
- PRE+ (Precision Plus): 1 metro (RMSE) di accuratezza orizzontale; cartografia di riferimento: 1: 2.500 NMAS.

•

I prodotti PRE, PRE PLUS e PRE STEREO richiedono la fornitura, da parte del cliente, dei punti di controllo al suolo (GCP) che possono essere anche forniti dalla Space Imaging Eurasia pagando un prezzo addizionale.

I prodotti STEREO non possono essere mosaicati. Il prezzo dei prodotti GEO, al contrario dei prodotti ortorettificati, non comprende la mosaicatura. Space Imaging Eurasia è in grado di effettuare, su specifica richiesta e pagando una tariffa addizionale, una rapida acquisizione e consegna dei prodotti GEO. Normalmente occorrono circa 10 giorni per 2 tentativi di acquisizione e circa 3 giorni per la spedizione e consegna del dato.

I prodotti possono anche essere forniti con una risoluzione geometrica di 0.82 metri. Ciò è possibile soltanto con una acquisizione nadirale (o quasi nadirale) ma che farebbe aumentare i tempi di consegna dei prodotti. Essendo IKONOS operativo già da tre anni, è possibile disporre di un consistente archivio di immagini.

5.5 - QUICKBIRD

QuickBird è stato lanciato il 18 Ottobre del 2001, ed è in fase operativa dalla primavera del 2002, su un'orbita polare eliosincrona, con 97,2 gradi di inclinazione, e con una velocità al suolo di 7.1 km/secondo. È in grado di acquisire sia in modalità multispettrale (tre bande del visibile + un infrarosso vicino) che pancromatica, con risoluzione tra 61 e 66 centimetri per angoli di acquisizione standard, cioè compresi tra 0 e 15 gradi. Il satellite ha capacità stereoscopiche *intrack*, cioè è in grado di acquisire coppie stereo lungo la stessa orbita; tale caratteristica, tuttavia, non viene ancora sfruttata dal punto di vista commerciale. I dati QuickBird sono disponibili sostanzialmente secondo due tipologie di prodotto: Basic e Standard. Il prodotto *Basic*, al quale sono applicate solo correzioni radiometriche e di sensore, è basato sulla singola scena di circa 16.5 x 16.5 km. La dimensione del pixel è variabile, e dipende dall'angolo di acquisizione della scena.

Nel *packaging* (circa 1.6 Gb per scene pancromatiche) sono forniti sia file immagine non georiferito che file ausiliari relativi a metadati, *Rational Polynomial Coefficients* (RPC), effemeridi, calibrazione geometrica, altitudine. Il prodotto Basic può essere processato geometricamente tramite un modello rigoroso oppure utilizzando software basato sull'uso dei *Rational Polynomial Coefficients*. Per un processamento ottimale, le informazioni fornite a corredo dell'immagine possono essere integrate con un DEM con punti di controllo a terra, dalla cui qualità dipende la precisione del risultato finale. Il prodotto Basic è rivolto ad utenti in grado di effettuare un

processamento avanzato dell'immagine dal punto di vista geometrico, in modo da ottenere la massima precisione nella georeferenziazione del dato. Il prodotto *Standard* differisce dal prodotto *Basic* in quanto ad esso vengono applicate anche delle correzioni geometriche, per cui il prodotto risulta inquadrato in un sistema di riferimento (WGS84) ed il pixel viene ricampionato ad una dimensione di 60 o 70 cm. Il *packaging* comprende: il file immagine, i metadati, il file degli RPC. Questo tipo di prodotto può essere acquistato anche su una superficie complessiva minore della singola scena. La qualità della georeferenziazione, che è basata unicamente sull'impiego dei dati orbitali post-processati ed integrati da un DEM a bassa risoluzione, può essere ulteriormente migliorata mediante processamento basato sugli RPC. Tuttavia, per ottenere precisioni più spinte, si consiglia di partire dal dato *Basic*. La quotazione di ciascun prodotto *Standard* viene effettuata sulla base della superficie (espressa in kmq) della particolare area di interesse del cliente, il quale potrà fornire le coordinate geografiche degli estremi e pagherà il prodotto solo per l'area richiesta. Per ottenere le immagini ortorettificate (georiferite secondo un sistema di proiezione, ellissoide e datum di riferimento), vengono usati GCP (ground control points), ed il DEM per correggere le distorsioni causate dall'altitudine; l'accuratezza planimetrica dipende dalla qualità dei GPC e del DEM, ma anche dall'angolazione con cui è stata effettuata la ripresa satellitare. I dati QuickBird sono distribuiti in esclusiva per tutta Europa da Eurimage, con l'eccezione dell'Italia ove la distribuzione è effettuata dalla Telespazio.

5.6 - I satelliti SAR ERS-1 e 2

Il principio fondamentale sul quale si basano i sistemi SAR è quello di emettere la radiazione elettromagnetica (nella regione delle microonde, in particolare, per ERS, nella banda C con una frequenza di 5,3 GHz) in direzione della superficie della Terra e di registrare la quantità ed il tempo di ritorno dell'energia di diffusione (backscattering). Questi sensori consentono di acquisire immagini indipendentemente dalla illuminazione solare e dalla presenza di nubi. I due satelliti SAR ERS-1 e 2, il cui lancio è avvenuto rispettivamente nel luglio 1991 e nell'aprile 1995 da parte dell'E.S.A. (Agenzia Spaziale Europea) si cui fa parte l'ASI (Agenzia Spaziale Italiana), presentano le stesse caratteristiche:

orbita elio-sincrona, circolare, con inclinazione di 98,5°, rotazione attorno alla Terra ogni 100 minuti ad un'altezza di 785 km. I due satelliti sono in grado di acquisire immagini in ogni punto della superficie della Terra ogni 17 giorni in orbita ascendente e discendente, cosicché l'intera copertura della Terra è ottenuta in 35 giorni, con una risoluzione al suolo pari all'incirca a 25 metri.

5.7 - Il satellite J-ERS

Il satellite J-ERS è un progetto della NASDA giapponese. Esso utilizza un sensore SAR che opera nel campo delle microonde, in particolare nella banda L ad una lunghezza d'onda di 24 cm. Il satellite ruota attorno alla Terra ad un'altezza di 570 km. Esso è in grado di fornire dati utili soprattutto nel campo della geologia, mentre il suo uso non è particolarmente indicato per studi nel campo della oceanografia, ciò in relazione alle lunghezze d'onda a cui opera.

5.8 - Il satellite RADARSAT

Il satellite RADARSAT è un progetto dell'Agenzia Spaziale Canadese. Esso utilizza un tipo SAR che opera nella regione delle microonde in particolare nella banda C ad una frequenza di 5,3 GHz. Le variazioni nel segnale di ritorno dovute alla diffusione sono il risultato delle variazioni della rugosità e della topografia della superficie osservata, così come delle proprietà fisiche quali, ad esempio, il contenuto di umidità del terreno.

Il satellite presenta un'orbita circolare, elio-sincrona, con un'inclinazione compresa tra 200 e 600, ed una rotazione attorno alla Terra ad un'altezza di 798 km. L'intera copertura della Terra è ottenuta, sia in orbita ascendente che in orbita discendente, ogni 24 giorni, con una risoluzione al suolo pari all'incirca a 25 metri.

6 - La Geodesia Spaziale

Sino dai primi anni '60 è apparso chiaro che la geodesia poteva trarre grandi benefici dalle acquisizioni della tecnologia spaziale, potendo disporre nei satelliti artificiali di punti di appoggio o di stazione per reti di vari tipi di osservazioni posti a grande altezza sull'orizzonte e collegati a punti a terra distanti tra loro al contrario dei classici schemi delle reti geodetiche tradizionali. Questa favorevole circostanza geometrica, unita ai progressi compiuti nel campo degli strumenti ad onde elettromagnetiche, sia nello spettro radio che in quello ottico, alla capacità di misurare il tempo con estrema accuratezza ed al grande sviluppo di potenti strumenti di calcolo, ha aperto la strada ad una nuova disciplina che si può chiamare geodesia spaziale, "satellite geodesy" in inglese.

Le osservazioni principali di tipo geometrico che si possono oggi ottenere nell'ambito di questa disciplina sono (fig.6.1):

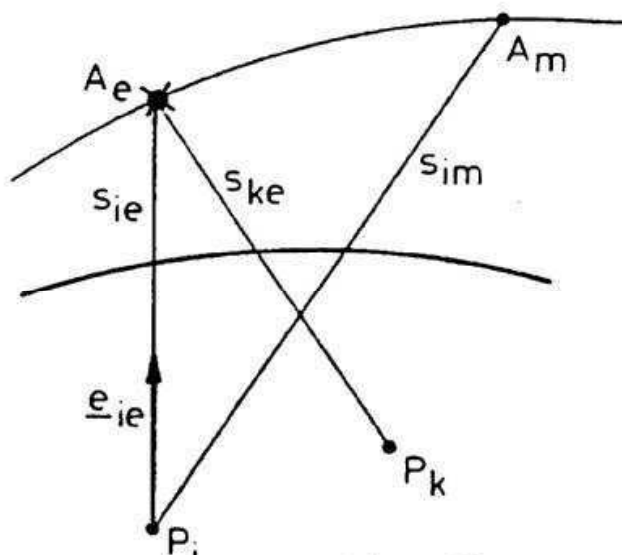


Fig. n. 6.1 - Osservazioni di tipo geometrico

- $S_{ie} = [A_e - P_i]$ ovvero la distanza o "range" tra stazione e satellite. Questa grandezza può essere misurata o tramite un distanziometro a laser (satellite laser ranging) o indirettamente mediante la misura del tempo τ_{ie} impiegato da una microonda per andare dal satellite A_e

alla stazione P_i in tal caso si parla di “pseudorange” $S_{ie} = c\tau_{ie}$ che ove c è la velocità della luce (metodo GPS navigazionale).

- $S_{im} - S_{ie} = [A_m - P_i] - [A_e - P_i]$ cioè una variazione di range: questa grandezza può essere ottenuta integrando sull'intervallo di tempo $t_m - t_e$, impiegato dal satellite per percorrere l'orbita da A_e ad A_m , la derivata S_{im} o “range rate” la quale a sua volta può essere misurata tramite l'effetto Doppler, ovvero la variazione di frequenza di una onda campione emessa dal satellite. Si osservi subito che la quantità $S_{im} - S_{ie}$ ottenuta tramite l'effetto Doppler va considerata come una vera osservazione autonoma e non va confusa con una semplice combinazione di due misure di distanza S_{im} ed S_{ie} che ottenuta ad esempio da due osservazioni laser.
- $S_{ie} - S_{ke} = [A_e - P_i] - [A_e - P_k]$ cioè la differenza di range tra una posizione A_e di un satellite e due stazioni P_i, P_k che quindi stanno osservando contemporaneamente A_e . Questa osservazione può essere effettuata in modo interferometrico, misurando la differenza di fase di un'onda emessa da A_e e ricevuta, ovviamente in tempi diversi, da P_i e da P_k . Questa misura potrebbe essere considerata come una semplice combinazione di due osservazioni di pseudorange, se non fosse che essa può essere effettuata anche quando non si abbia una buona conoscenza a priori della particolare forma d'onda emessa dal satellite, quale invece è necessaria per la misura dello pseudorange. Questo tipo di tecnica può essere assimilata ad una nota tecnica radioastronomica, Very Long Base-Line Interferometry (VLBI)¹

E' possibile utilizzare modelli puramente geometrici. Tale può essere considerato ad esempio il posizionamento di un singolo punto di stazione a terra con metodo Doppler o GPS in cui la posizione del o dei satelliti tracciati è considerata come fissa e nota (questo metodo è comunemente usato in navigazione).

Ancora puramente geometrici sono quei modelli che sfruttano osservazioni simultanee tra stazioni a terra e singole posizioni del satellite, non legate tra loro da una legge dinamica: ad esempio un gruppo di 4 o più stazioni che osservino contemporaneamente il range o facciano misure interferometriche per molti passaggi dello stesso o di più satelliti possono determinare la loro posizione relativa in modo completamente indipendente dall'orbita. Ciò può essere facilmente intuito pensando ad esempio che 4 stazioni che facciano 4 misure contemporanee di range (fig. 6.2) danno luogo ad una relazione di condizione tra le coordinate delle stazioni stesse, dopo aver eliminato le coordinate di A.

¹ una delle stazioni mondiali VLBI è a Matera. Con questa tecnica si raggiungono accuratèzze dell'ordine di $1\text{cm}/1000\text{km}(10^{-8})$

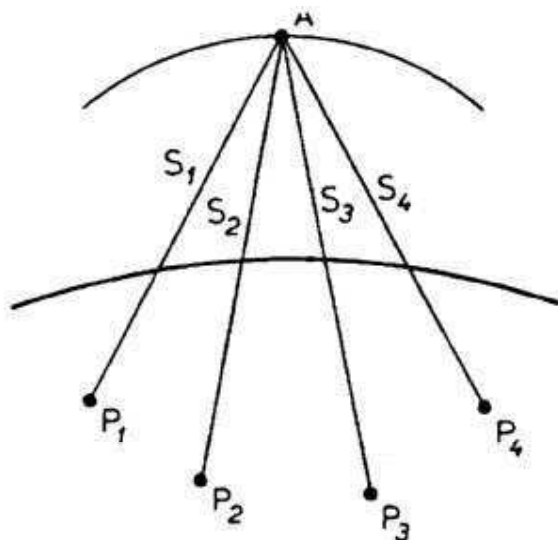


Fig. n. 6.2 - Osservazioni contemporanee di range da 4 stazioni a terra

Vi sono però altri metodi che, senza essere completamente rigorosi, utilizzano il fatto che una piccola porzione di traiettoria, ad esempio, un arco di una frazione inferiore ad un terzo d'orbita, appare con molta buona approssimazione come un arco di ellisse, si segue cioè la traiettoria di un satellite in un campo gravitazionale puramente sferico. Tali metodi vengono detti di arco corto ("short arc") e nella loro applicazione per ogni arco lungo per il quale sono effettuate delle misure compare un gruppo di 6 parametri accessori che descrivono appunto un segmento di traiettoria ellittica o kepleriana. In ogni caso, che si tenga conto o no della dinamica del satellite, poiché tutti i metodi descritti si basano sulla trasmissione di onde elettromagnetiche tra stazione e satellite, onde che ovviamente devono attraversare mezzi diffrattivi come l'atmosfera e la ionosfera, è necessario che le equazioni di osservazione includano dei parametri che descrivono gli effetti fisici di questi mezzi sul segnale.

7 - Satellite Laser Ranging (SLR)

Il principio di osservazione del range stazione-satellite mediante laser è ben noto fin dall'introduzione dei distanziometri elettrolitici (EDM). Nel caso particolare del SRL la stazione è sita a terra ed il satellite funge da bersaglio. Da un punto di vista puramente geometrico si tratta di risolvere una trilaterazione nello spazio in cui le posizioni dei satelliti tracciati sono o considerate fisse (prese ad esempio dalle "precise ephemeris") o intervengono con un certo numero di incognite (ad esempio i 6 parametri kepleriani per un arco corto).

Le particolarità del SLR sono legate alla necessità di sparare impulsi laser a grandi distanze (impulsi ad alta energia) e con un fascio abbastanza ristretto perché si abbia sufficiente energia per ottenere una risposta misurabile: a tale riguardo torna utile osservare che con i moderni fotomoltiplicatori è possibile rilevare un raggio di ritorno fotone per fotone.

La precisione intrinseca dell'osservazione di distanza varia oggi attorno ai 5 cm rispetto alla distanza satellite stazione. La ristrettezza del fascio emesso fa sì che il laser debba essere diretto verso il satellite con grande precisione. Ciò comporta che il cannone laser sia montato su supporto mobile in ogni direzione (simile a quello di un teodolite) comandato nei suoi movimenti da un minicomputer. Il calcolatore, che già ha un programma per il tracciamento dell'orbita in osservazione, compie anche delle correzioni in tempo reale per ottimizzare la ricezione del segnale di ritorno.

Il bersaglio, quando si tratta di un satellite dedicato al SLR come il Lageos o lo Starlette, deve a sua volta avere caratteristiche precise: in particolare deve essere il più massiccio possibile ed avere la forma la più regolare per poter minimizzare il drag atmosferico e l'effetto della pressione di radiazione. Ciò ha determinato la caratteristica forma di sfera coperta di prismi retroriflettori che si può osservare in fig.7.1

Date queste caratteristiche tecniche l'apparato della stazione non è facilmente miniaturizzabile; ciò nonostante esistono già oggi i primi apparati laser mobili montati su camion. Una delle stazioni della rete SLR è stata istituita a MATERA e funziona dal 1985.

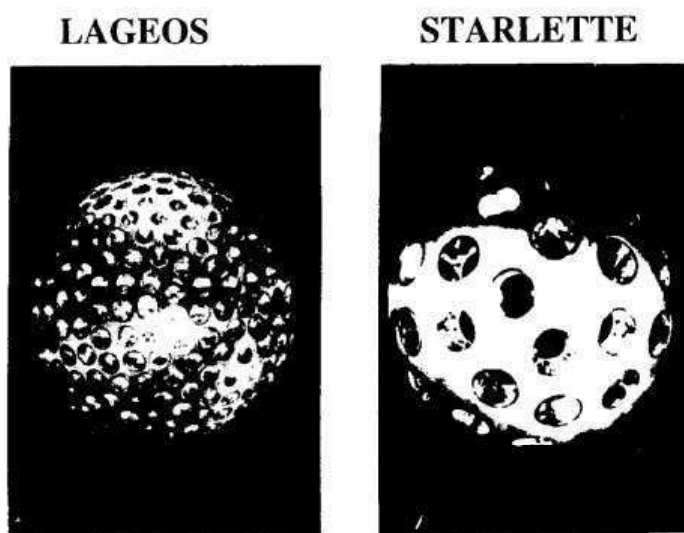


Fig. n. 7.1 - (a) Raggio 30 cm, peso 410 kg, 426 riflettori, altezza media 5900 km, inclinazione 109,9° ; (b) Raggio 12 cm, peso 47,3 Kg, 60 riflettori, altezza media 950 km, inclinazione 49,8°

8 - Sistema Inerziale

La novità maggiore, che rappresenta una vera e propria rivoluzione nel campo del posizionamento, è rappresentata senza dubbio dalla messa in funzione dei sistemi inerziali, che permettono di determinare automaticamente la posizione plano-altimetrica dei punti, nonché la deviazione della verticale ed anche le differenze di gravità.

Si può di fatto pensare ad una “scatola nera” che, montata su di un veicolo, fornisce le coordinate dei punti stabiliti lungo il percorso.

Ciò che rende il sistema inerziale l'apparato del futuro è, oltre alla velocità di lavoro ed alla precisione, che può essere ancora migliorata, la sua grande versatilità e varietà di campi di impiego.

Il ben noto effetto di persistenza della direzione dell'asse di un giroscopio in un sistema di riferimento inerziale, garantisce il parallelismo della piattaforma a se stessa quando questa è montata su un veicolo in movimento: ciò naturalmente a meno delle imprecisioni dello strumento (deriva dei giroscopi).

Anziché restare sempre parallela a se stessa la piattaforma stabilizzata può essere ruotata da servomotori in modo, ad esempio, che essa si mantenga sempre orizzontale e che gli assi di due dei tre giroscopi siano disposti nord-sud ed est-ovest: questa è, ad esempio, la soluzione adottata dalla Ferranti per il suo FILS (Ferranti Inertial Land Surveyor), di cui riportiamo lo schema in fig. 8.1.

In questo caso la velocità di rotazione attorno ad ogni asse, sempre rispetto ad un sistema inerziale, è misurabile tramite le coppie che questa provoca sugli assi dei giroscopi stessi: pertanto, integrando, diventa possibile in ogni istante sapere di quanto la piattaforma è stata ruotata.

A parte l'uso corrente con autovetture ed elicotteri, il sistema è stato montato su slitte, su fuoristrada, su navi e sottomarini, su aeroplani e su treni.

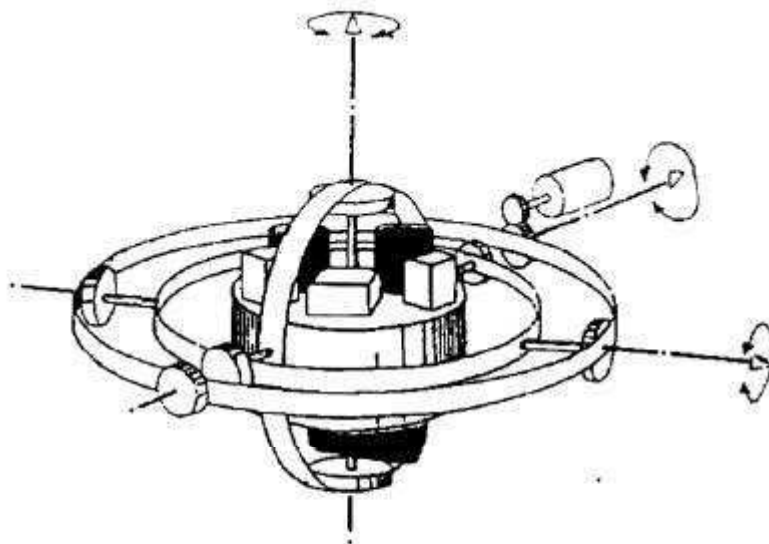


Fig. n. 8.1 - Schema del sistema inerziale

9 – Conclusioni

In definitiva, è auspicabile che “il risultato ultimo della geodesia sia quello di rendere accessibile a tutti un accurato posizionamento con una facilità pari a quella dell'acquisizione del tempo”.

Attualmente è in atto la diffusione del *Mobile Mapping* ossia una pluralità di tecniche per la conoscenza del territorio aventi come elemento cardine l'uso di un veicolo in movimento. La tecnica più diffusa è quella in cui su di un furgone vengono installati una serie di sensori necessari all'orientamento istantaneo del mezzo (GPS, INS, Odometri) e alcune telecamere digitali, o anche tradizionali, che riprendono in continuo, lungo direzioni differenti, il territorio nel quale il furgone si muove.

L'epoca attuale è senz'altro ricca di novità che consentono l'innovazione di molteplici soluzioni per cui, in un non lontano futuro, prevediamo che la gente andrà in giro con un localizzatore da polso, al prezzo di una decina di dollari, che gli permetterà un posizionamento istantaneo con una precisione millimetrica (esiste già in commercio un GPS da polso che il Prof. fa vedere durante le lezioni).

Capitolo 4 – GPS

Capitolo 4 – GPS	89
1 - Introduzione	91
2 - Principio di base del posizionamento GPS	92
3 - Organizzazione del sistema GPS	93
3.1 - La sezione spaziale	93
3.2 - La sezione di controllo	95
3.3 - La sezione utenza	96
4 - Modalità di misura del sistema GPS	97
4.1 - Metodo per misura di pseudoranges	97
4.2 - Metodo per misura di fase	99
4.2.1 - Equazioni indifferenziate	100
4.2.2 - Equazioni alle differenze singole	101
4.2.3 - Equazioni alle differenze doppie	103
4.2.4 - Equazioni alle differenze triple	103
4.2.5 - Precisione del posizionamento differenziale per misura di fase	105
5 - Uso geodetico e topografico del sistema GPS	106
5.1 - Modalità di esecuzione delle misure	106
5.1.1- Utilizzo del GPS in modo statico	106
5.1.2 - Utilizzo del GPS in modo cinematica, statico rapido e “realtime”	107
5.2 - Reti di baselines	108
5.3 - Elaborazione dei dati acquisiti	109
6 - Vantaggi e svantaggi del Sistema GPS rispetto alle tecniche topografiche classiche	110
7 - La Rete IGM95	111

1 - Introduzione

Il sistema in oggetto, la cui denominazione completa è “NAVSTAR GPS”, abbreviazione di “Navigation Satellite Timing And Ranging Global Positioning System” (sistema di posizionamento globale mediante misure di tempo e distanza rispetto a satelliti per navigazione) è stato realizzato a partire dal 1975 circa, a cura dello U.S. Department of Defence; il sistema, concepito per un posizionamento ad alte prestazioni per scopi militari in particolare per la missilistica, è impiegabile anche per usi civili quali la navigazione l'utilizzo topografico, ecc.

La costellazione di satelliti cui il sistema fa riferimento è costituita da 24 satelliti, progettata in modo da renderne visibili fino a 12 contemporaneamente, da ogni punto della superficie terrestre, in ogni ora del giorno e della notte (si vedrà come questa sia condizione indispensabile per l'applicabilità del metodo).

Il sistema permette di eseguire il posizionamento di punti appartenenti alla superficie terrestre, ovvero la determinazione della posizione dei punti interessati in un assegnato sistema di riferimento.

Il posizionamento che si ottiene con il GPS è tridimensionale, quindi contemporaneamente planimetrico e altimetrico.

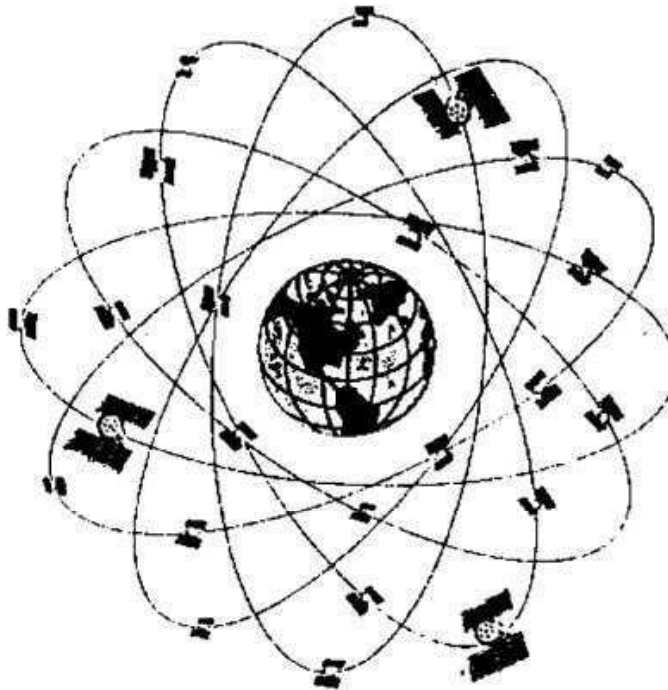


Fig. n. 1- La costellazione GPS

Tale operazione può essere eseguita utilizzando due distinte possibilità operative, diverse sia concettualmente che per il grado di precisione conseguibile:

A) La tecnica di “POINT POSITIONING” (posizionamento assoluto di un punto singolo nel sistema di riferimento assegnato), con la quale allo stato attuale si determina la posizione con un'incertezza dell'ordine di decine di metri; si tratta quindi di una tecnica di precisione

sufficiente per la navigazione o per tracciamenti speditivi (con i cosiddetti apparati “pathfinder”), ma non per le applicazioni topografiche.

B) La tecnica del “DIFFERENTIAL POSITIONING” o “RELATIVE POSITIONING”, con la quale si determina la posizione di un punto rispetto a un altro punto considerato noto. In pratica viene determinato il vettore posizione relativa della “linea base” o base (in inglese “BASELINE”) tra i due punti nelle sue tre componenti rispetto a una terna cartesiana assegnata.

La precisione ottenibile su queste tre coordinate relative è dell’ordine di qualche milionesimo della distanza, pari quindi, o superiore, a quella di operazioni geodetiche classiche eseguite con la massima accuratezza.

Per eseguire il posizionamento GPS relativo è necessario disporre di due ricevitori, che devono essere posti sui due estremi della base e funzionare contemporaneamente per l’intera durata della sessione di misura, che va da qualche decina di minuti a qualche ora, a seconda della distanza e della precisione richiesta.

Con apparati in grado di ricevere entrambe le frequenze (più complessi e costosi) si possono determinare baselines di centinaia di chilometri.

Va subito detto che non è necessario che i due estremi di una base siano tra loro intervisibili: è questo il fondamentale vantaggio del metodo GPS rispetto alle tecniche topografiche classiche.

2 - Principio di base del posizionamento GPS

Il sistema si basa su un principio concettualmente molto semplice (anche se poi l’applicazione concreta risulta più complessa).

Il metodo di posizionamento si fonda essenzialmente sulle osservazioni distanziometriche tra i satelliti, di cui si suppone nota la posizione, e i punti a terra da determinare.

Si supponga infatti di conoscere le effemeridi orbitali di un satellite S, ovvero le tre coordinate $X_s(t)$, $Y_s(t)$, $Z_s(t)$ del centro del satellite in funzione del tempo, in un assegnato riferimento cartesiano.

E’ stato stabilito un sistema cartesiano geocentrico(WGS-‘84), avente origine nel centro di massa della Terra, asse Z diretto secondo l’asse di rotazione terrestre ed assi X e Y giacenti sul piano equatoriale.

Si supponga inoltre di poter misurare la distanza, in un dato istante t, fra un punto P a terra (di coordinate incognite X_p , Y_p , Z_p , nel sistema precedentemente definito) e il satellite S.

Sotto queste ipotesi, è possibile ottenere una equazione di osservazione (ovvero una equazione che lega le grandezze incognite a quelle note, da utilizzare in un procedimento di compensazione di osservazioni indirette) uguagliando la distanza cartesiana alla distanza misurata, con riferimento all’istante di misura:

$$\overline{PS} = \sqrt{(X_s(t) - X_p)^2 + (Y_s(t) - Y_p)^2 + (Z_s(t) - Z_p)^2} \quad (2.1)$$

Con tre equazioni del tipo (2.1), ottenibili misurando la distanza dal punto P a tre diversi satelliti, il problema sembrerebbe risolto¹.

¹ Il satellite comunica al ricevitore le effemeridi usando un suo orologio interno. Il ricevitore, al momento delle misure ricerca tra le effemeridi ricevute la posizione del satellite in base alla conoscenza del tempo con il suo

Prof. MAURO CAPRIOLI
Capitolo 4 – GPS

Occorre però precisare la definizione dell'istante temporale nel quale avviene la misura: il tempo misurato a terra (da un orologio incorporato nel ricevitore) differisce da quello di riferimento del sistema GPS (orologi a bordo dei satelliti) di una quantità non trascurabile.

Nell'equazione scritta compaiono quindi in realtà 4 incognite: X_p , Y_p , Z_p , e uno sfasamento temporale $\delta t = t_p - t_s$. E' pertanto necessario che dal punto P risultino visibili almeno 4 satelliti, in modo da poter scrivere 4 equazioni e determinare le 4 incognite.

Le equazioni sono in realtà più complesse della (2.1), nella quale sono state omesse alcune incognite "secondarie" o "parametri di disturbo" (in inglese "nuisance parameters"), che rappresentano errori sistematici piccoli ma non trascurabili.

In pratica tutto ciò che è stato sin qui assunto come "noto", in realtà non lo è, ovvero risulta affetto da incertezze. Ad esempio si è ipotizzato:

- che lo sfasamento temporale δt tra l'orologio a terra e i 4 orologi dei satelliti sia unico, ovvero che gli orologi dei satelliti siano sincronizzati tra loro (in realtà non lo sono, al livello di precisione con cui tale sincronismo sarebbe richiesto);
- si è considerata nota l'orbita, che è invece, necessariamente, un'orbita di previsione e quindi approssimata;
- sono state trascurate le incertezze nella misura della distanza terra-satellite dovute all'imperfetta conoscenza della velocità di propagazione dell'onda elettromagnetica impiegata per la distanziometria (nella ionosfera e nella troposfera l'onda subisce dei rallentamenti di diversa natura).

Nella pratica, per conseguire risultati di precisione adeguata alle esigenze topografiche, occorre tener conto di molti dei fattori sopra citati e modellarli opportunamente; in alternativa alla modellazione, che risulta spesso difficoltosa, è possibile applicare particolari procedimenti operativi (fra i quali sono essenziali le tecniche differenziali usate per il posizionamento relativo) che riducono di molto l'influenza di quei parametri dalla stima incerta, eliminandoli dalle equazioni mediante opportune combinazioni lineari delle stesse.

3 - Organizzazione del sistema GPS

Il GPS è un sistema abbastanza complesso; per descriverne l'organizzazione, conviene esaminare distintamente le tre sezioni in cui si articola (sezione spaziale, di controllo e dell'utenza).

3.1 - La sezione spaziale

E' costituita, come anticipato, dalla costellazione di satelliti NAVSTAR; il programma della NASA consta di 24 satelliti, di cui 18 operativi, 6 di riserva (i satelliti operativi in orbita possono essere sino a 32). Il numero massimo di satelliti è legato alla posizione degli stessi nelle orbite e al tipo di codifica di trasmissione dei dati. Sono disposti su 6 orbite, inclinate di 55° rispetto all'equatore. Su ogni orbita sono quindi disposti 3 satelliti operativi e 1 di riserva.

Oltre al suddetto sistema GPS americano esiste il sistema satellitare russo GLONASS, il quale è costituito da una costellazione di 14 satelliti contemporaneamente visibili sull'orizzonte. I segnali GLONASS non sono alterati come quelli GPS consentendo, quindi, con il loro utilizzo

orologio, che però differirà di δt da quello del satellite. Ogni satellite trasmette l'almanacco ogni 12 minuti e mezzo.

posizionamenti con errori limitati (dell'ordine di alcuni metri in point positioning). Le caratteristiche dei GLONASS e dei GPS sono molto simili anche se in questi ultimi l'inclinazione dell'orbita sull'asse polare è maggiore, ciò garantisce una migliore copertura delle regioni a latitudine elevata. I due sistemi sono inoltre elettronicamente compatibili, infatti è possibile usare un ricevitore misto con un'unica antenna.

I satelliti GPS orbitano ad una distanza dalla Terra di 20.200 km circa, alla quale non risentono molto delle anomalie del campo gravitazionale. Sono posizionati su 6 piani orbitali intervallati di circa 60 gradi sessagesimali.

Il periodo di rivoluzione attorno alla Terra è di 12 ore siderali, corrispondenti a $11^h 56'$, per cui i satelliti si presentano ogni giorno con 4 minuti di anticipo rispetto al giorno precedente.

I satelliti trasmettono a terra dati contenenti un segnale di tempo, le proprie effemeridi orbitali (ovvero le coordinate $X_s(t)$, $Y_s(t)$, $Z_s(t)$, di cui si è detto), e informazioni varie quali il proprio numero di identificazione, lo stato di "salute", ecc.

Tutti i dati sono trasmessi (fig.3.1) mediante un codice binario modulato su due frequenze portanti radio (i cui valori sono, rispettivamente, $f_1=1575$ MHz e $f_2=1228$ MHz, corrispondenti alle lunghezze d'onda $\lambda_1 \cong 19\text{cm}$ e $\lambda_2 \cong 24\text{cm}$).

I codici usati nella trasmissione sono tre: il codice C/A ("coarse acquisition", ovvero "acquisizione grossolana") e il P ("precision", consente una precisione molto maggiore), con frequenza di modulazione rispettivamente pari a 1 e 10 MHz, sono accessibili a tutti gli utenti; in alcuni casi il codice P può essere cifrato per essere reso non accessibile per uso non militare e viene denominato codice Y.

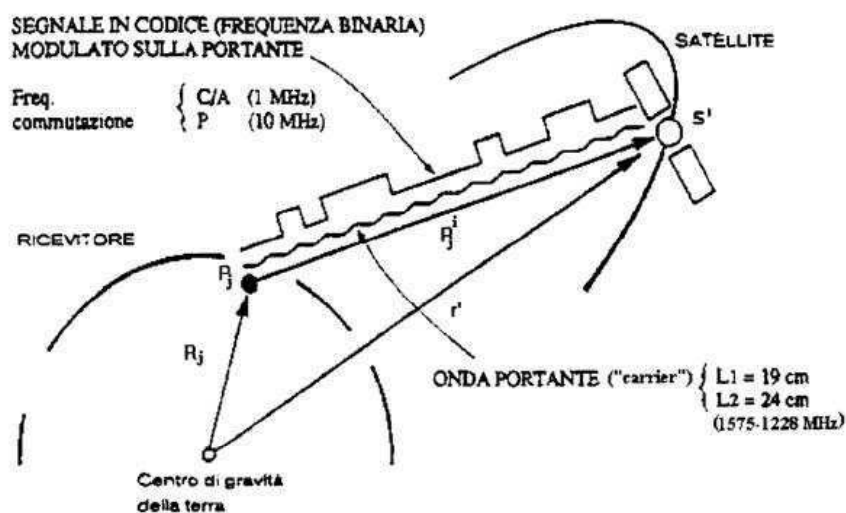


Fig. n. 3.1 - Modalità di emissione dei segnali dai satelliti

Le "finestre" temporali di osservabilità per una data zona e in un certo giorno, possono essere determinate preventivamente ("planning") utilizzando appositi programmi predisposti dalle ditte produttrici di ricevitori GPS, consentendo così la programmazione delle sessioni di misura. Per poter operare è necessaria, come si è detto, la contemporanea presenza di almeno 4 satelliti nel cielo sopra il punto di stazione. Vanno considerati "utili" solo i satelliti visibili con angolo di elevazione sull'orizzonte superiore a una certa soglia (di solito 15-20 gradi): quelli con elevazione inferiore comportano un segnale disturbato sia dall'eccessivo spessore di

atmosfera attraversata che, eventualmente, dalla presenza di alberi od edifici ed altre ostruzioni che impediscono la buona ricezione dei segnali.

Il software di preparazione dà informazioni anche sulla bontà della configurazione geometrica con cui i satelliti si presenteranno (i satelliti fungono in pratica da “trigonometrici” e devono essere il più possibile uniformemente distribuiti nel cielo), quantificandola mediante il parametro PDOP (Position Dilution Of Precision, ovvero “calo di precisione nel posizionamento tridimensionale”). Questo parametro (unitamente ad altri dal significato leggermente diverso, come l’HDOP, legato alle sole componenti planimetriche o il VDOP legato alla sola quota) rappresenta il contributo del fattore configurazione geometrica dei satelliti all’errore quadratico medio nel posizionamento; perché risulti soddisfacente, il suo valore deve essere limitato a poche unità. Nella definizione delle “finestre” di osservabilità occorre quindi tener presente, oltre al numero di satelliti (v. fig. 3), anche il valore del PDOP. Nell’impiego geodetico-topografico del GPS, basato sulle tecniche differenziali, sembra sufficiente che il PDOP sia basso all’inizio e alla fine di una sessione di misura, e che siano tollerabili valori più alti nel periodo intermedio; si tratta comunque di un aspetto da approfondire sia a livello teorico che sperimentale.

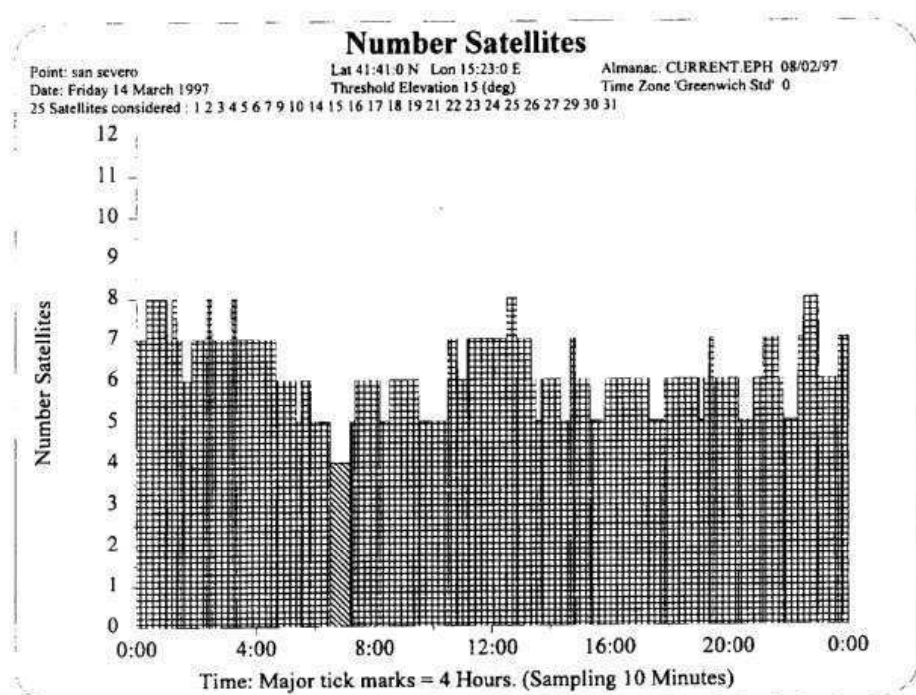


Fig. n. 3.2 - Periodi di tempo con relativi numeri di satelliti, previsione del giorno 14 marzo 1997

3.2 - La sezione di controllo

Si tratta del sistema di controllo da terra dei satelliti e delle informazioni da questi trasmesse. È composto essenzialmente da:

- cinque stazioni fisse di “tracking”, situate in punti di posizione geografica nota, disposti grosso modo nella fascia equatoriale della Terra, uniformemente intervallati in longitudine. Le stazioni di tracking, dotate di apparati di ricezione estremamente precisi, seguono con continuità i satelliti, rilevandone l’orbita; le posizioni note di questi 5 punti definiscono in pratica il sistema di riferimento nel quale le orbite vengono calcolate, cioè il sistema cartesiano geocentrico di cui si è parlato all’inizio,

detto anche “sistema convenzionale terrestre”, attualmente nella definizione “World Geodetic System 1984” o “WGS 84”, sigla che indica un ben definito ellissoide.

- una stazione principale di controllo (“master”), situata a Colorado Springs, negli Stati Uniti. La stazione “master” riceve i dati rilevati dalle stazioni di tracking, i quali vengono elaborati per calcolare le effemeridi orbitali di previsione dei vari satelliti, e per valutare le correzioni da apportare periodicamente sia agli orologi dei satelliti stessi sia, eventualmente, anche alle orbite (azionando a distanza piccoli motori a razzo correttori di rotta di cui sono i satelliti sono dotati). La stazione master è dotata di sofisticate apparecchiature di elaborazione e di un orologio atomico di riferimento all'idrogeno, più preciso di quelli (al cesio o al rubidio) presenti a bordo dei satelliti.
- tre stazioni di trasmissione che ricevono i dati corretti dalla stazione master e li trasmettono periodicamente ai satelliti, effettuando il cosiddetto “upload” (aggiornamento) dei satelliti stessi.

I dati orbitali e di tempo trasmessi dai satelliti a terra sono quindi più affidabili subito dopo l'effettuazione di un “upload”, e si degradano via via, fino al successivo. Con le tecniche di misura in uso (fra le quali, si è detto, sono essenziali le tecniche differenziali) l'effetto di tale degrado sui risultati è tuttavia modesto.

I blocchi di satelliti delle prossime generazioni beneficeranno di sistemi di rilevamento e comunicazione più efficaci e precisi, con i quali si potranno ottenere effemeridi molto meglio approssimate e rendere più continuo l'aggiornamento dei satelliti.

3.3 - La sezione utenza

L'utenza civile del sistema GPS si va sempre più estendendo, sia per numero di utilizzatori che per varietà di impieghi. I ricevitori e il software per le principali applicazioni del sistema sono in commercio ormai da più di dieci anni.

I ricevitori GPS sono ricevitori radio sintonizzati sulle frequenze usate dal sistema, dotati di sistemi di decodifica ed elaborazione dei segnali ricevuti e di una memoria per l'immagazzinamento dei dati. Sono composti da un'antenna che viene disposta sul punto da determinare (su un comune treppiede o pilastrino, essendo dotata di un basamento compatibile con quelli dei più diffusi strumenti topografici). L'antenna è collegata mediante un cavo schermato al ricevitore propriamente detto, ovvero al gruppo di sintonizzazione e acquisizione dati. Quest'ultimo ha le dimensioni di un'autoradio e comprende oltre ai sistemi di decodifica del segnale e al microprocessore di elaborazione, un orologio di precisione (normalmente un oscillatore al quarzo di elevata qualità), una memoria fisica interna o su supporto magnetico mobile, un software interno per il controllo del processo di acquisizione dati; fondamentale è la batteria di alimentazione per assicurare il funzionamento per almeno otto ore di acquisizione. Le caratteristiche dei ricevitori variano a seconda della ditta produttrice e del modello; le differenze più significative fra i vari tipi di ricevitori attualmente in commercio sono le seguenti:

- possibilità di ricezione dei segnali su una sola frequenza, (apparati “monofrequenza”) o su entrambe le frequenze disponibili;
- numero di canali di ricezione (ne occorre uno per satellite), da un minimo di 4 a un massimo di 12;
- misura con il metodo “pseudorange” o anche con il metodo per “differenza di fase”, dei quali si dirà al successivo punto; i ricevitori per applicazioni geodetico-topografiche devono disporre di entrambe le modalità di misura.

Il software per gli utenti GPS comprende programmi per la preparazione delle sessioni di misura (essenzialmente, come detto, per stabilire preventivamente gli orari migliori per

l'acquisizione "planning"), e ulteriori programmi per il "postprocessing" ovvero per l'elaborazione dei dati registrati dai ricevitori, che vengono allo scopo riversati, con opportune interfacce, su personal computer.

Il software di postprocessing esiste nelle versioni "commerciali", adatte all'uso pratico predisposte dalle ditte produttrici e vendute assieme ai ricevitori, e in versioni "scientifiche" che permettono una maggiore libertà nell'elaborazione (consentendo ad esempio l'uso di diversi modelli per la propagazione atmosferica delle onde radio, o la possibilità di ricalcolare le orbite), ma presuppongono un'approfondita conoscenza del sistema a livello teorico.

4 - Modalità di misura del sistema GPS

Le misure con cui si perviene al posizionamento GPS possono essere essenzialmente raggruppate in due categorie: misure di PSEUDORANGES (pseudodistanze) e misure di FASE sulla portante ("carrier phase"). Si tratta di due modalità di misura completamente diverse con le quali si ottengono due livelli di precisione ben distinti, adatti ad impieghi differenti.

4.1 - Metodo per misura di pseudoranges

E' il metodo concettualmente e praticamente più semplice, ma anche il meno preciso. E' basato sul semplice principio esposto in precedenza:

effettuare misure di distanza dal punto incognito a terra a satelliti di posizione nota; si tratta in pratica di eseguire una "intersezione all'indietro" in tre dimensioni con sole misure di distanza con i satelliti che fungono da "trigonometrici".

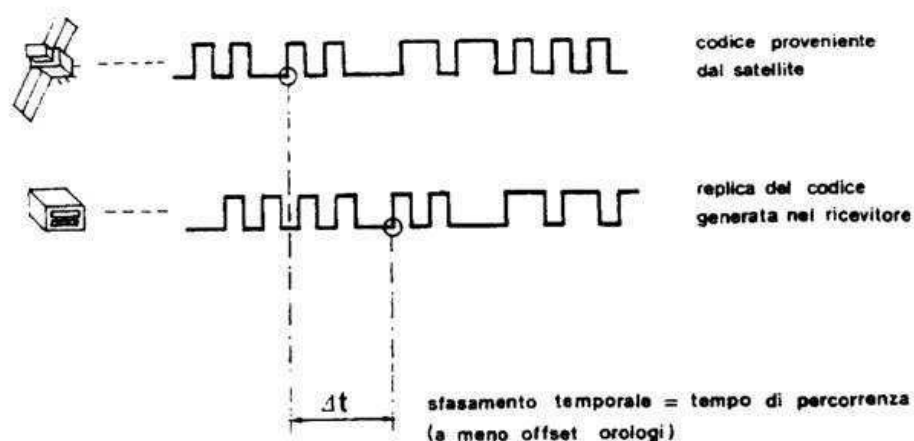


Fig. n. 4.1 - Principio della misurazione di pseudorange

La misura di distanza viene eseguita misurando il tempo di propagazione di un impulso (una data fase del segnale in codice modulato sulla portante) trasmesso dal satellite verso terra. In pratica il segnale in codice (successione di numeri binari) che arriva dal satellite viene confrontato con una copia dello stesso codice generata nel ricevitore (i ricevitori più recenti sono in possesso sia del codice C/A che del codice P) e viene determinato lo sfasamento temporale fra i due codici, ovvero lo "scorrimento" in termini di tempo necessario per portare i

due segnali a sovrapporsi perfettamente (fig. 4.1), che coincide con il tempo di propagazione cercato.

Il tempo di percorrenza moltiplicato per la velocità di propagazione dell'onda elettromagnetica fornisce la distanza satellite-ricevitore, detta "pseudodistanza" (PSEUDORANGE) perché affetta da numerose incertezze e quindi poco precisa.

Ogni misura di "pseudodistanza" genera un'equazione del tipo:

$$\sqrt{(X_s(t_0) - X_p)^2 + (Y_s(t_0) - Y_p)^2 + (Z_s(t_0) - Z_p)^2} = \int_{t_0}^t v \cdot dt \quad (4.1)$$

nella quale è stato riportato l'integrale della velocità nel tempo anziché il prodotto velocità per tempo, considerando che, come si è detto, la velocità non è costante lungo il percorso, a causa delle diverse caratteristiche fisiche degli strati atmosferici attraversati.

Il tempo scandito dall'orologio del singolo satellite risulta diverso dal tempo relativo al singolo ricevitore, di uno sfasamento temporale δt :

$$t_r = t_s + \delta t \quad (4.2)$$

Le incognite presenti nella (4.1) sono quindi quattro: X_p , Y_p , Z_p , δt ; sono pertanto necessari quattro satelliti per ottenere quattro equazioni e determinare la posizione del punto.

Una soluzione "di emergenza" può essere ottenuta anche con tre satelliti quando sia nota a priori la quota del punto P (ad es. nella navigazione in mare).

Il metodo per pseudoranges è caratterizzato da una serie di approssimazioni che lo rendono poco preciso. Le principali incertezze riguardano le effemeridi orbitali dei satelliti ($X_s(t)$, $Y_s(t)$, $Z_s(t)$) e la velocità di propagazione del segnale nell'atmosfera (il comportamento è sensibilmente diverso nella ionosfera e nella troposfera).

Vi è inoltre un limite "intrinseco" di precisione del metodo dovuto alla elevata lunghezza d'onda del segnale modulato su cui si va a operare.

A causa di tutte queste incertezze, la precisione nel posizionamento per misura di pseudoranges è dell'ordine di alcune decine di metri sulle coordinate planimetriche e sulla quota, del tutto insufficiente quindi ai fini topografici e geodetici.

Il metodo presenta, peraltro, numerosi vantaggi:

- la posizione viene ottenuta in tempo reale (dopo due-tre minuti dall'accensione si ha il primo posizionamento, automaticamente visualizzato in coordinate geografiche e quota, e l'aggiornamento è poi praticamente continuo);
- si opera con un solo ricevitore. Si tratta quindi di una tecnica particolarmente adatta alla navigazione; allo scopo esistono in commercio numerosi tipi di ricevitori con questa sola possibilità di misura, molto più semplici ed economici rispetto ai ricevitori "geodetici", che operano anche per misura di fase.

La praticità ed economicità del metodo per pseudoranges fa sì che siano allo studio perfezionamenti per migliorarne la risoluzione. Con tecniche particolari (ad es. lo "pseudorange differenziale" oppure la tecnica "phase-smoothed pseudoranges", la quale ultima presuppone però la misura di fase) la precisione di posizionamento può essere migliorata; queste tecniche tuttavia non vengono molto impiegate in topografia dato che, come detto, per il GPS topografico si usano gli apparati a misura di fase.

4.2 - Metodo per misura di fase

E' una tecnica di misura concettualmente diversa da quella per pseudorange, anche se nella interpretazione geometrica si riconduce al principio base esposto all'inizio.

Si agisce in questo caso sull'onda radio portante, depurata dal segnale in codice ad essa "sovrainpresso".

La misura consiste nel confrontare la fase della portante ("carrier phase") trasmessa dal satellite con quella di un segnale di pari frequenza generato nel ricevitore.

Si intuisce subito come la precisione possa essere maggiore rispetto a quella della tecnica per pseudorange: si va ad operare su un segnale di lunghezza d'onda molto inferiore (19 o 24 cm anziché 30 o 300 m).

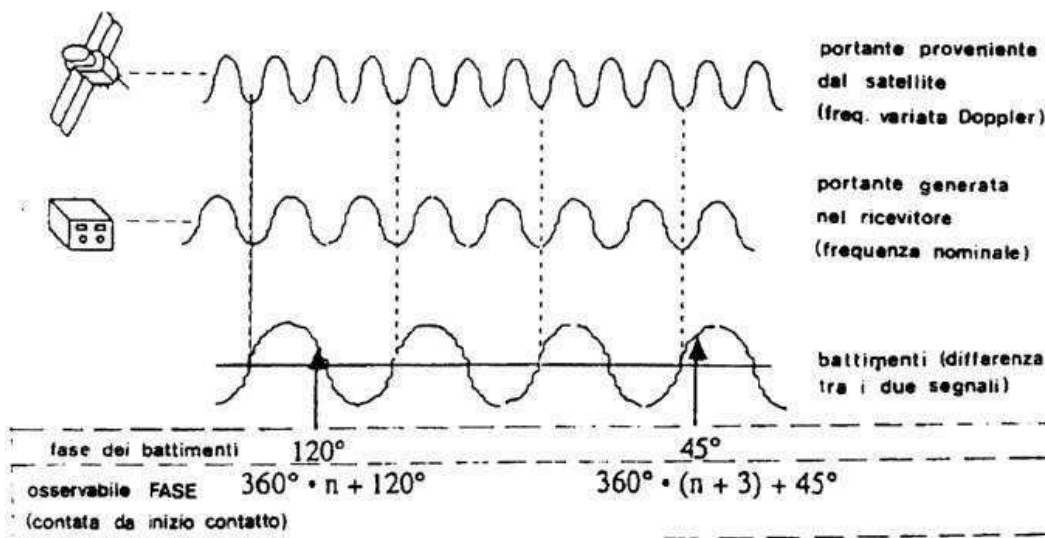


Fig. n. 4.2 - Principio della misurazione di fase

Per comprendere il principio della misura di fase va tenuto presente che l'onda proveniente dal satellite viene ricevuta a terra con una frequenza variata rispetto a quella di emissione, per l'effetto Doppler dovuto alla velocità relativa fra satellite e ricevitore a terra.

La "replica" del segnale generata dal ricevitore è invece stabile sulla frequenza nominale.

Dal confronto fra le due onde, di frequenza poco diversa, risulta un'onda di battimenti ("beat signal", fig. 4.2).

La "osservabile" (ovvero la grandezza misurata) FASE è definita come la differenza tra la fase del segnale proveniente dal satellite e quella della replica generata nel ricevitore; si può dimostrare che tale differenza coincide con la fase di battimenti di cui si è detto, definita a meno di un numero intero di cicli N (detto "ambiguità"), che è incognito.

La misura viene eseguita seguendo il satellite per un certo tempo senza perdere il contatto con esso.

A partire dall'istante del primo contatto, la grandezza FASE, come sopra definita, viene misurata con continuità, e l'indeterminazione N resta costante sul valore iniziale.

Per ogni misura di FASE (relativa a un dato istante di tempo t) si ottiene un'equazione di osservazione del tipo:

$$\varphi_s(t) - \varphi_p(t) = \phi(t) + N + f \cdot \delta t \quad (4.3)$$

nella quale:

- $\phi(t)$ è la fase dei battimenti in cicli

- N è l'ambiguità (numero intero iniziale di cicli, incognito)
- $f \cdot \delta t$ (con f = frequenza) è il termine relativo allo sfasamento tra l'orologio del satellite e quello del ricevitore (anche qui presente, e incognito)
- $\varphi_s(t)$ e $\varphi_p(t)$ sono rispettivamente le fasi (in cicli) del segnale proveniente dal satellite e di quello generato dal ricevitore, nell'istante t di misura.

Moltiplicando l'equazione per la lunghezza d'onda λ della portante si perviene ad una interpretazione geometrica della misura di fase:

$$\lambda(\varphi_s - \varphi_p) = \overline{PS}(t) = \lambda[\phi(t) + N + f \cdot \delta t] \quad (4.4)$$

nella quale $\overline{PS}(t)$ è la distanza fra satellite e ricevitore, che può essere espressa in funzione delle coordinate cartesiane come nella (3.1), introducendo così le incognite di posizione del ricevitore e le effemeridi orbitali note:

$$\frac{\sqrt{(X_s(t) - X_p)^2 + (Y_s(t) - Y_p)^2 + (Z_s(t) - Z_p)^2}}{\lambda} = [\phi(t) + N + f \cdot \delta t] \quad (4.5)$$

In questa equazione compaiono 5 incognite: X_p , Y_p , Z_p , δt e N .

Anche qui, però, permangono i fattori di incertezza visti a proposito della misura di pseudorange:

- la lunghezza d'onda λ è data dal rapporto fra la velocità di propagazione (affetta da incertezze per la sua variabilità nell'attraversamento di ionosfera e troposfera) e la frequenza (che si può invece considerare nota);
- le effemeridi orbitali $X_s(t)$, $Y_s(t)$ e $Z_s(t)$ come già più volte detto, non sono precise come si vorrebbe.

Grazie alle tecniche differenziali si riesce a ridurre gli effetti di queste cause di incertezza, e si risolve anche il problema dovuto alla presenza dell'ambiguità N .

Vediamo appunto come con queste tecniche, applicabili nel posizionamento relativo, si possa pervenire a una soluzione del problema con le equazioni di fase.

Il procedimento più frequentemente adottato consiste nell'eseguire successive combinazioni lineari delle equazioni di osservazione di fase, eliminando via via i parametri incogniti; lo si espone di seguito, in sintesi, nei suoi "passi" logici successivi.

4.2.1 - Equazioni indifferenziate

Si faccia riferimento alla misura di fase eseguita da un ricevitore r a terra, in un dato istante t , sul segnale trasmesso da un dato satellite j (fig. 4.3); l'equazione di osservazione, come visto sopra, è la:

$$\phi_{r-j}(t) = \frac{RJ(t)}{\lambda} - N_{r-j} - f \cdot [\delta t_r + \delta t_j] \quad (4.6)$$

nella quale si è ora posto $\delta t = \delta t_r + \delta t_j$, ovvero lo sfasamento temporale è stato scomposto in due sfasamenti, rispettivamente dell'orologio del satellite e di quello del ricevitore, rispetto al tempo "assoluto" (come riferimento si assume, nel caso del GPS, il tempo del Naval

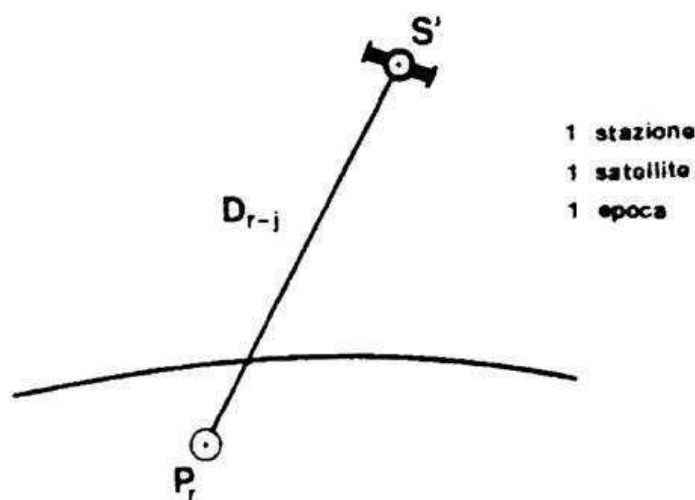


Fig. n. 4.3 - Misura singola di fase

Observatory di Washington), e $\phi_{r-j}(t)$ indica sempre la FASE come precedentemente definita, misurata senza soluzione di continuità dall'istante del contatto iniziale. Con riferimento alla fig. 4.2, la fase misurata può anche essere espressa come:

$$\phi = M + k$$

dove M è il numero di cicli interi contati a partire dall'istante iniziale, e k è una frazione di ciclo.

Se viene perso, anche per un istante, il contatto con il satellite (ad es. per l'interporsi di un qualsiasi ostacolo), si determina un cosiddetto "cycle slip", ovvero una interruzione nel conteggio del numero M di cicli, che costringe a "spezzare" la sessione di misura, ovvero a considerare un nuovo valore incognito dell'ambiguità N da quel momento in poi. Con tecniche particolari, di vario tipo, il "cycle slip" può essere "recuperato" consentendo di riunificare la sessione, ritornando a un'unica ambiguità iniziale incognita.

Le incognite che compaiono nella (6) sono:

- le tre coordinate X_p, Y_p, Z_p , (contenute entro RJ);
- gli "offset" degli orologi δt_r e δt_j (in questo caso, volendo effettuare un posizionamento di precisione, è necessario considerare un offset diverso per ciascun satellite);
- e l'ambiguità iniziale N_{r-j}

4.2.2 - Equazioni alle differenze singole

Si supponga di avere disposto un secondo ricevitore su un altro punto a terra, operante contemporaneamente al primo (è il posizionamento differenziale o relativo di cui si è detto), e si considerino due misure di fase eseguite nello stesso istante dai due ricevitori rispetto a un medesimo satellite (fig. 4.4); si possono allora scrivere due equazioni indipendenti del tipo (4.6).

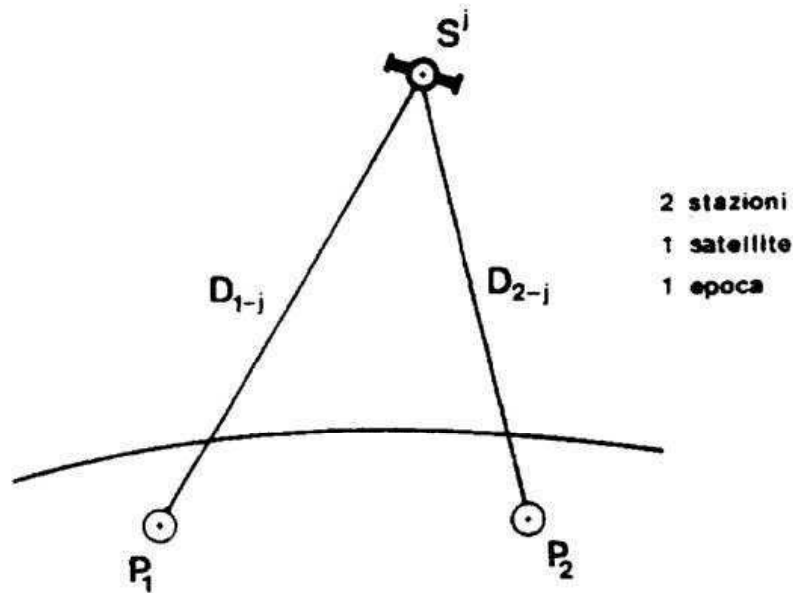


Fig. n. 4.4 - Differenza singola

Si definisce DIFFERENZA SINGOLA (“single difference”, SD) la differenza tra le fasi misurate su uno stesso satellite nella medesima “epoca” (istante) dai due ricevitori (indicati con 1 e 2):

$$SD_{12-j}(t) = \phi - j(t) - \phi_{1-j}(t)$$

L’equazione di osservazione “alla differenza singola” è una combinazione lineare (differenza) di due equazioni del tipo (4.6):

$$\begin{aligned} SD_{12-j}(t) &= [P_2 J - P_1 J(t) / \lambda] - N_{2-j} + N_{1-j} - f \cdot [\delta t_2 + \delta t_j - \delta t_1 - \delta t_j] = \\ &= 1 / \lambda \cdot [P_2 J(t) - P_1 J(t)] + N_{1-j} - N_{2-j} + f \cdot [\delta t_1 - \delta t_2] \end{aligned} \quad (4.7)$$

Si può notare come da questa equazione sia scomparso lo sfasamento temporale dell’orologio del satellite, che è il medesimo nelle due equazioni di cui si è eseguita la differenza; si ha così un’incognita in meno rispetto alla (4.6).

Rispetto all’equazione indifferenziata si ha inoltre un altro fondamentale vantaggio: dato che la distanza tra i due punti a terra è piccola rispetto alla distanza terra-satellite, la differenza eseguita attenua in maniera sostanziale gli errori sistematici dovuti all’imperfetta conoscenza dell’orbita e della velocità di propagazione del segnale nell’atmosfera.

4.2.3 - Equazioni alle differenze doppie

Se dalle due stazioni a terra vengono ricevuti due satelliti j e k (gli stessi due) contemporaneamente è possibile scrivere due equazioni del tipo (4.7) con riferimento ai 2 satelliti (fig. 4.5)

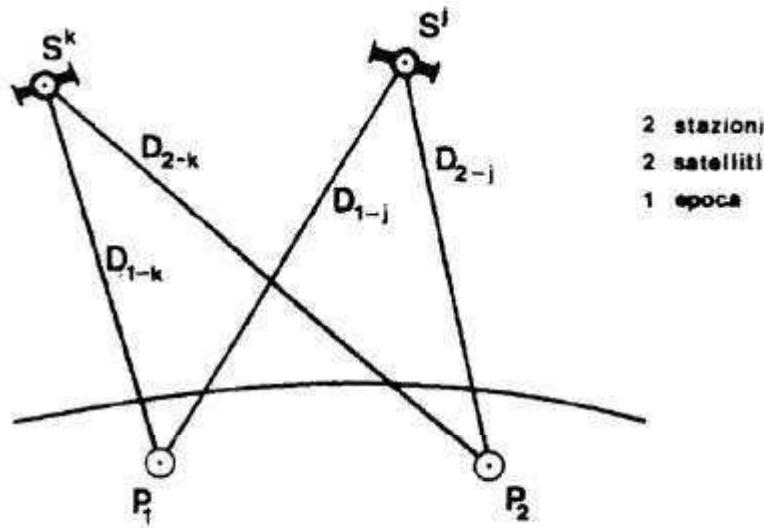


Fig. n. 4.5 - Differenza doppia

Se si esegue la differenza tra le due differenze singole (definita DIFFERENZA DOPPIA o “double difference”, DD) si ottiene una equazione del tipo:

$$\begin{aligned} DD_{12-jk}(t) &= SD_{12-k}(t) - SD_{12-j}(t) = \\ &= 1/\lambda \cdot [P_2K(t) - P_1K(t) - P_2J(t) + P_1J(t)] + N_{1-k} - N_{2-k} - N_{1-j} + N_{2-j} + \\ &+ f \cdot [\delta t_1 - \delta t_2 - \delta t_1 + \delta t_2] \end{aligned}$$

da cui

$$DD_{12-jk(t)} = 1/\lambda \cdot [P_2K(t) - P_1K(t) - P_2J(t) + P_1J(t)] + N_{12-jk} \quad (4.8)$$

nella quale sono scomparsi anche gli sfasamenti temporali degli orologi di entrambi i ricevitori; le 4 ambiguità (tutti numeri interi) sono state conglobate in un unico numero intero incognito N_{12-jk} .

4.2.4 - Equazioni alle differenze triple

Si considerino, infine, due osservazioni di differenza doppia eseguite in due differenti “epoche” t_1 e t_2 , senza perdere contatto con i satelliti fra l’una e l’altra, in modo che l’ambiguità rimanga costante sul valore iniziale (fig. 4.6).

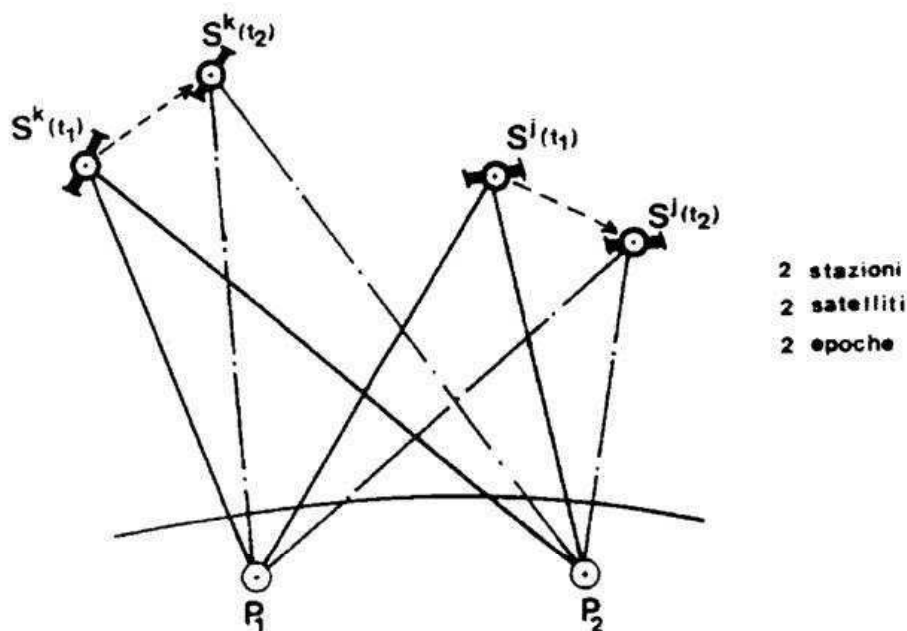


Fig. n. 4.6 - Differenza tripla

Se eseguiamo la differenza tra le due differenze doppie nei due istanti t_1 e t_2 (definita DIFFERENZA TRIPLA, o “triple difference”, TD) otteniamo un’equazione di nuovo tipo, combinazione lineare di 2 equazioni del tipo (4.8):

$$\begin{aligned}
 TD_{12-jk}(t) &= DD_{12-jk}(t_2) - DD_{12-jk}(t_1) = \\
 &= 1/\lambda \cdot [P_2K(t_2) - P_1K(t_2) - P_2J(t_2) + P_1J(t_2) - P_2K(t_1) + P_1K(t_1) + P_2J(t_1) - P_1J(t_1)] - \\
 &- N_{12-jk} + N_{12-jk}
 \end{aligned}$$

da cui:

$$TD_{12-jk}(t_1, t_2) = 1/\lambda \cdot \quad (4.9)$$

$$[P_2K(t_2) - P_2K(t_1) + P_1K(t_1) - P_1K(t_2) + P_2J(t_1) - P_1J(t_1) - P_2J(t_2) + P_1J(t_2) - P_1J(t_1)]$$

Si elimina così anche la “fastidiosa” incognita costituita dall’ambiguità iniziale. La condizione necessaria è, come detto, che non ci siano perdite di contatto con i satelliti fra gli istanti t_1 e t_2 . Gli istanti t_1 e t_2 devono essere abbastanza distanziati, altrimenti al secondo membro della (4.9) si ottengono differenze fra quantità quasi uguali (ad es. $P_2K(t_2) - P_2K(t_1)$), ovvero termini molto piccoli che determinano una “instabilità numerica” della soluzione. La sessione di misura deve avere quindi una certa durata (qualche decina di minuti: si parla in questo caso di GPS statico, v. oltre).

Nelle equazioni tipo (4.9), se si considera noto uno dei due punti a terra, risultano incognite solo le coordinate dell’altro punto. In pratica, ad uno dei due punti si assegnano coordinate arbitrarie (ad es., i valori approssimati ottenuti dalla soluzione in “pseudorange”), e si determinano le 3 componenti del vettore “baseline” $P_1 P_2$:

$$\Delta X = X_2 - X_1$$

$$\Delta Y = Y_2 - Y_1$$

$$\Delta Z = Z_2 - Z_1$$

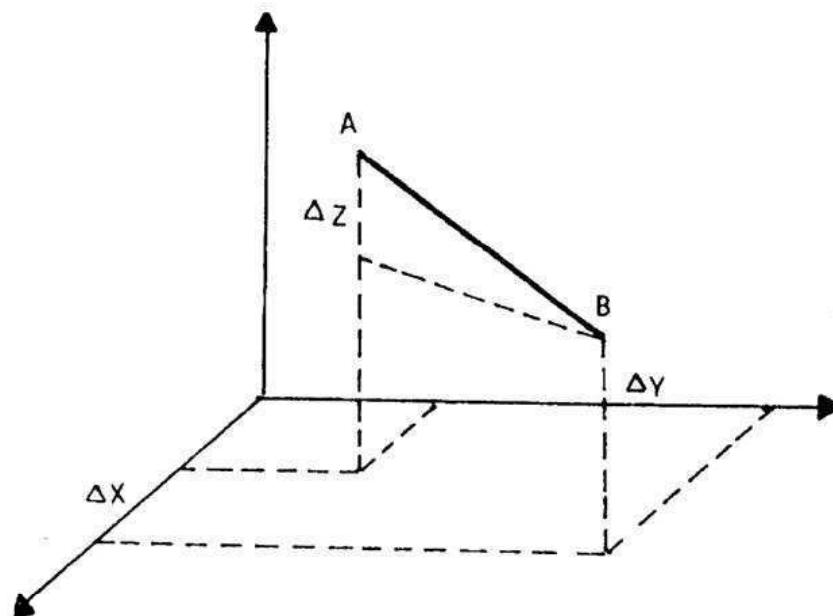


Fig. n. 4.7 - Misura della base A - B

ovvero le differenze di coordinate tra i due punti, che consentono il calcolo della lunghezza del vettore nello spazio e del suo orientamento.

Per determinare le tre incognite sono necessarie almeno 3 equazioni del tipo (4.9), che si ottengono, anche in questo caso, con un numero minimo di 4 satelliti (per avere almeno 3 combinazioni indipendenti dei satelliti presi due a due). Ottenuta la soluzione dalle equazioni alle differenze triple, la si può introdurre come valore di prima approssimazione nelle equazioni alle differenze doppie, determinando così il valore di N , risultando da una stima numerica, non risulta in genere uguale ad un numero intero (come vorrebbe la teoria), anche se normalmente si avvicina alla cifra “tonda”. La soluzione così ottenuta viene detta “FLOATING” o “FLOT”.

Una terza soluzione viene poi ottenuta fissando il valore N sull'intero più vicino e rideterminando le coordinate relative (soluzione “FIXED”). Esistono anche altre “strategie” di soluzione per migliorare la precisione dei risultati.

Nel caso in cui si utilizzano i sistemi misti GPS+GLONASS esiste la necessità di aumentare a 5 il numero di satelliti necessari per ottenere il posizionamento tridimensionale, ciò è dovuto alla presenza di un'ulteriore incognita che è lo scostamento temporale fra gli orologi della costellazione GPS e quelli della costellazione GLONASS.

4.2.5 - Precisione del posizionamento differenziale per misura di fase

Con la misura di fase si può pervenire, come anticipato, ad una precisione nel posizionamento relativo dell'ordine di alcune parti per milione della distanza tra i due punti sulle componenti planimetriche (pari a $\pm 1 \text{ mm}/(\text{km h})^2$); leggermente inferiore, in genere, la precisione sulla

² È evidente la dipendenza dall'accuratezza dalla maggiore o minore quantità di segnale ricevuto

componente altimetrica, a causa della configurazione geometrica del problema, che è più sfavorevole per le quote (i satelliti ricevibili sono tutti necessariamente sopra l'orizzonte, quindi tutti da uno stesso lato rispetto al punto, mentre ai fini della planimetria sono distribuiti in tutte le direzioni attorno al punto).

La precisione indicata è considerata lo “standard” ideale del procedimento; allo stato attuale, a volte le configurazioni non sono ottimali (PDOP abbastanza elevato) e nella pratica si ottengono a volte precisioni inferiori anche di un ordine di grandezza. In alcune campagne di misura eseguite per finalità scientifiche, ricorrendo a particolari “raffinatezze” nel postprocessing (ad es. considerando orbite ricalcolate a posteriori invece di usare le effemeridi di previsione trasmesse dal satellite) sono state ottenute precisioni anche dell'ordine di qualche decimilionesimo della distanza, il che renderebbe il metodo particolarmente adatto a misure geofisiche, o al controllo di precisione di opere di ingegneria, o alla determinazione dei movimenti del suolo. Nel valutare la precisione dei risultati occorre, comunque, essere cauti:

gli errori quadratici medi delle componenti di una base o delle coordinate relative dei vertici di una rete di sole misure GPS, risultano spesso formalmente molto piccoli, ma possono nascondere errori sistematici di maggiore entità, dovuti in genere all'errata modellazione della velocità di propagazione del segnale. Per valutare la effettiva precisione è quindi sempre consigliabile integrare le misure GPS con alcune osservazioni di tipo classico, o inserire nel rilievo GPS alcuni punti noti per altra via.

5 - Uso geodetico e topografico del sistema GPS

5.1 - Modalità di esecuzione delle misure

Nell'impiego geodetico-topografico del GPS possono essere adottate diverse modalità operative; quelle maggiormente utilizzate sono tre: il modo statico, il modo cinematico lo “stop and go”.

5.1.1- Utilizzo del GPS in modo statico

Il GPS statico è il metodo più “sicuro” e più frequentemente adottato, specie quando le esigenze di precisione sono spinte ed i punti da determinare sono più importanti. Il metodo consiste nell'eseguire sessioni di misura di una certa durata (da un minimo di 30 min. ad alcune ore) con i ricevitori fissi sugli estremi delle “baselines” da determinare. Le stesse vengono determinate mediante la soluzione “alle differenze triple” esposta al precedente punto 4; come visto, è necessario ricevere almeno quattro satelliti e non perdere il contatto con essi durante la sessione.

La modalità statica è anche quella caratterizzata dalla massima precisione dei risultati: come anticipato, la precisione “standard” raggiungibile nel posizionamento relativo è dell'ordine di $10^{-6}D$ spinta anche a $10^{-7}D$ in alcune esperienze (D = distanza fra gli estremi della “baseline”), grazie ad elaborazioni particolari e a configurazioni di satelliti particolarmente favorevoli. Gli accorgimenti da adottare nella misura in modalità statica variano a seconda della lunghezza delle basi: per quelle sino a qualche decina di chilometri sono in genere sufficienti ricevitori monofrequenza e sessioni di durata dell'ordine di 1 ora; salvo che per casi

particolari in cui le esigenze di precisione siano estremamente spinte, si possono usare tranquillamente i software “commerciali” forniti con gli apparati GPS. In pratica la misura è alla portata di chiunque disponga di una coppia di ricevitori.

Per basi più lunghe si richiedono ricevitori in doppia frequenza che permettono di ottenere una migliore modellazione della velocità di propagazione del segnale nell’atmosfera; le sessioni di misura devono essere più lunghe (da una ad alcune ore). Per “baselines” di estrema lunghezza (è possibile raggiungere e superare i 1000 km), oltre a quanto sopra, si richiede anche una modellazione “a posteriori” delle orbite, ovvero l’introduzione nel calcolo di effemeridi più precise rispetto a quelle di previsione trasmesse dal satellite; si tratta però, evidentemente, di applicazioni a livello scientifico.

In ogni stazione in questo caso si rilevano i dati atmosferici di pressione, temperatura ed umidità, che vengono poi introdotti nel programma di calcolo della base, nel quale intervengono nella modellazione della propagazione del segnale. Non tutti gli sperimentatori sono d’accordo sulla messa in conto di queste correzioni atmosferiche, che hanno effettivamente un valore locale poco significativo nei riguardi di ciò che accade nel lunghissimo percorso satellite-terra; si tratta comunque di una “condizione al contorno” che è sempre possibile escludere a posteriori.

Con il GPS statico si effettuano le misure geodetiche propriamente dette (determinazione di punti isolati o blocchi di punti), per le quali il metodo è notevolmente vantaggioso rispetto alle tecniche topografiche classiche, sia in termini di precisione che di tempo e di costo. Le “baselines” misurate vengono a formare reti analoghe a quelle trigonometriche, che possono essere compensate in blocco. Le reti GPS possono essere a maglie triangolari come le reti trigonometriche classiche, ma non necessariamente. E’ ancora oggetto di studio quale sia la loro conformazione ottimale e se tale conformazione della rete in-fluisca effettivamente sulla bontà dei risultati: nel caso del GPS la “rete” dovrebbe, infatti, includere anche i satelliti.

5.1.2 - Utilizzo del GPS in modo cinematica, statico rapido e “realtime”

Il GPS cinematico consiste nel tenere fisso uno dei due ricevitori e nel trasportare il secondo (detto “rover” o “roving antenna”) sui punti da determinare, seguendo un certo percorso con continuità di movimento ed eseguendo determinazioni di posizione a intervalli di tempo regolari.

Il “rover” può essere portato a mano da un operatore, o montato su un veicolo (quale un’autovettura o un aereo).

Il metodo statico rapido è, sostanzialmente, una variante operativa del metodo cinematico propriamente detto, con la quale i punti da determinare vengono occupati successivamente dal “rover”, tenendolo fermo per qualche istante (con una palina telescopica munita di livella sferica); il metodo si presta in modo particolare al rilievo di dettaglio. Nel caso del GPS cinematico (o statico rapido: le due varianti sono identiche dal punto di vista del calcolo) non è possibile ottenere la soluzione con le equazioni alle differenze triple, perché i ricevitori non sono fissi e cadono le ipotesi fatte al punto 4.2.4. E’ quindi necessario determinare l’ambiguità iniziale N (non potendola eliminare dalle equazioni). A tale scopo è possibile, in alternativa:

- a) eseguire una sessione GPS statica preliminare su due punti noti (anche molto vicini tra loro), con la quale è possibile ricavare il valore di N, e quindi partire con il “rover” per il percorso stabilito;
- b) applicare il cosiddetto metodo dell’ “antenna swap “scambio di antenne”, con il quale le antenne vengono tenute per alcuni minuti sugli estremi non noti, di una corta base quindi scambiate di posizione e mantenute così per qualche minuto ancora, dopo di che

si parte con il “rover” (si utilizzano equazioni di forma simile a quelle alle triple differenze, ma con alcuni segni cambiati, dalle quali risulta possibile determinare l’ambiguità anche non conoscendo la posizione relativa dei due punti iniziali, e con una sessione di durata brevissima.

Nel GPS cinematico (o statico rapido) è comunque necessario non perdere mai il contatto con i satelliti (se ciò avviene, occorre interrompere il rilevamento e rideterminare un nuovo valore dell’ambiguità N). Ciò costituisce purtroppo una limitazione operativa abbastanza sensibile, che rende il metodo difficilmente applicabile in zone urbane o con presenza di vegetazione di alto fusto.

Il GPS cinematico trova svariate applicazioni anche al di fuori della topografia, per la sua capacità, ad esempio, di fornire il percorso di un mezzo in movimento. Un’applicazione molto promettente in fase di studio è quella aerofotogrammetrica, per la determinazione delle coordinate del centro di presa dei fotogrammi.

Nel settore delle misure topografiche, il metodo cinematico è impiegato in particolare per rilievi di dettaglio. A differenza dal caso statico, non si ottiene in genere una rete compensabile, ma un insieme di corte baselines che fanno tutte capo al punto traslocante (occupato dal ricevitore fisso). La precisione è normalmente inferiore a quella del GPS statico, ma comunque sufficiente per quasi tutte le applicazioni topografiche.

Oltre ai tre principali modi d’impiego anzidetti, esistono altre possibili modalità operative quale ad esempio il “real time”.

Questa procedura di rilevamento è comunemente designata con l’acronimo RTKS (Real Time Kinematic Survey cioè rilevamento cinematico in tempo reale). o, più semplicemente, RTK (Real Time Kinematic, cioè cinematico in tempo reale).

Questa tecnica di rilevamento, relativamente recente, è frutto della costante evoluzione della tecnologia GPS, sia nel settore della strumentazione che dei programmi applicativi.

Mentre le procedure di rilevamento cinematico prima accennate prevedono che la elaborazione dei dati per la determinazione delle posizioni dei punti rilevati sia effettuata in ufficio, dopo le operazioni sul terreno, in questo caso la determinazione della posizione di ciascun punto rilevato viene effettuata direttamente in campagna in tempo reale, cioè al momento stesso in cui il punto da determinare viene occupato dal ricevitore mobile.

Il metodo prevede l’utilizzazione di ricevitori a doppia frequenza, collegati fra loro via radio.

Il ricevitore fisso (di norma collocato su un punto di posizione nota) comunica la sua posizione ed i dati satellitari al ricevitore mobile, che in base ai dati suddetti, calcola in tempo reale la sua posizione rispetto al ricevitore fisso.

Lo scambio dei dati fra la stazione fissa ed il ricevitore mobile viene effettuato in un appropriato formato.

Questa tecnica di rilevamento, risulta particolarmente interessante per la esecuzione di operazioni topografiche di picchettamento, tracciamento e simili, per le quali i tempi di esecuzione vengono abbreviati in maniera considerevole.

5.2 - Reti di baselines

Effettuando più misure elementari che collegano a due a due i punti interessati dal rilievo, si ottiene una rete di “baselines”. Ogni punto deve essere collegato al resto della rete da almeno una base. Se le misure sono eseguite con ridondanza rispetto allo stretto indispensabile (il minimo è una base in meno rispetto al numero dei punti), la rete può essere compensata, in modo analogo a quanto si fa per le reti topografiche classiche; le equazioni di osservazione sono ovviamente di tipo diverso, poiché, come si è detto, rappresentano misure di differenze di coordinate.

In ogni caso, sia che la rete venga compensata (se esiste ridondanza di misure), sia che le misure rimangano allo stato “grezzo”, il risultato che si ottiene è la determinazione delle differenze di coordinate cartesiane tra i punti della rete nel sistema di riferimento WGS 84.

Se si assegnano le coordinate di uno dei punti della rete (sempre nel WGS 84) risultano allora determinate quelle di tutti gli altri. Il punto assunto come fisso viene denominato “punto traslocante” perché la sua posizione determina quella di tutto il blocco di punti ad esso collegati.

Nei casi più favorevoli, ad esempio, quando nella zona oggetto del rilievo esistono punti già determinati in tale sistema (vedi IGM 95), assumendo uno di tali punti come traslocanti, verranno allora determinati nello stesso sistema anche tutti gli altri punti della rete misurata.

Dalle coordinate cartesiane X, Y, Z , inadatte agli impieghi pratici, è possibile calcolare le coordinate geografiche (φ, λ) e la quota “ellissoidica” dei punti della rete con riferimento all’ellissoide geocentrico WGS-‘84, avente cioè il centro coincidente con l’origine della terna cartesiana X, Y, Z e ben definiti parametri (semiassi, eccentricità, ecc...). Tale trasformazione, pur non immediata, è effettuata con relativa semplicità dai programmi commerciali di corredo degli strumenti.

5.3 - Elaborazione dei dati acquisiti

Per utilizzare in maniera più completa i risultati delle misure GPS è necessario eseguire il cambiamento di sistema di riferimento per passare a uno dei “datum” locali della cartografia esistente: IGM (GAUSS - BOAGA) o UTE (CASSINI - SOLDNER). A tale scopo sono stati approntati i relativi programmi di trasformazione che consistono nella seguente procedura:

- a) dalle coordinate cartesiane geocentriche dell’ellissoide WGS 84 si passa, con una rototraslazione nello spazio, alle coordinate cartesiane riferite a una terna, avente l’origine nel centro dell’ellissoide locale prescelto (Hayford, Bessel), orientata secondo gli assi dell’ellissoide considerato. Sono perciò necessari sette parametri, tre di traslazione, tre di rotazione, più un fattore di scala che viene in genere inserito per meglio adattare le nuove misure alla rete esistente essendo le reti caratterizzate da incertezze diverse. I sette parametri vengono stimati ai minimi quadrati in base alla conoscenza delle coordinate cartesiane in entrambi i sistemi, sia nel WGS 84 che nel sistema locale, per un congruo numero di punti (ne occorrono minimo tre).
- b) da queste coordinate cartesiane “ellissocentriche”¹ si passa alle coordinate geografiche (φ, λ) e alla quota ellissoidica rispetto all’ellissoide locale, con un procedimento di trasformazione da coordinate cartesiane geografiche, cui si è accennato nel capitolo della geodesia (equazioni parametriche dell’ellissoide). La parte planimetrica di un rilievo GPS viene così resa congruente con le reti trigonometriche locali e con la cartografia esistente. Per l’altimetria, è necessario conoscere, nei punti rilevati, le ondulazioni del geoide rispetto all’ellissoide locale per poter ricavare le quote ortometriche dalla relazione: $Q = h - N$

¹ Per coordinate cartesiane “ellissocentriche” si intendono quelle relative al sistema avente origine nel centro dell’ellissoide (Hayford) da non confondersi con quelle geocentriche cioè aventi origine nel centro della Terra.

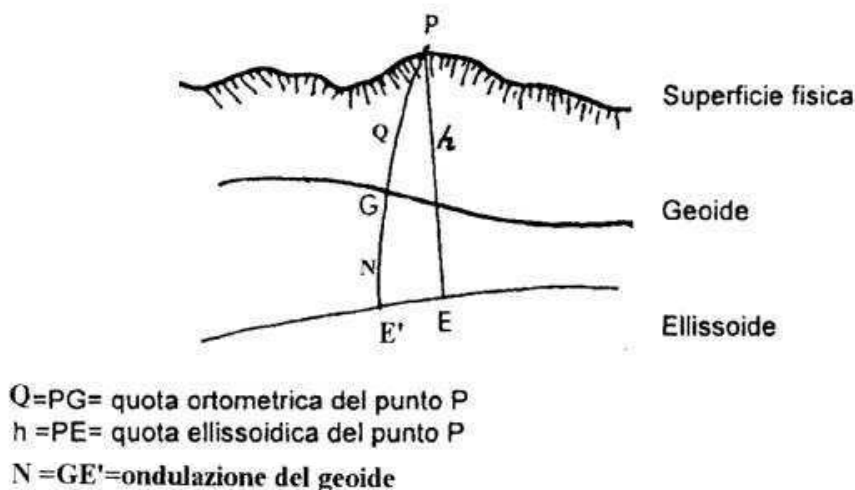


Fig. n. 11 - Le superfici di riferimento

Il problema della determinazione della ondulazione del geoide è quindi di primaria importanza per la pratica utilizzabilità della tecnica GPS, soprattutto per distanze notevoli.

6 - Vantaggi e svantaggi del Sistema GPS rispetto alle tecniche topografiche classiche

Il GPS presenta numerosi aspetti che lo rendono particolarmente “allettante” agli occhi degli operatori della Topografia:

- la semplicità d'uso è estrema: a differenza dei metodi classici, non si richiede personale particolarmente esperto per le operazioni di campagna; anche la post-elaborazione dei dati è sempre più automatizzata con i programmi commerciali, che hanno raggiunto un'ottima affidabilità;
- non è necessaria l'intervisibilità fra i punti collegati da una “baseline”
- non si risente delle condizioni atmosferiche e di visibilità: si può operare con pioggia, nebbia e anche di notte;
- il posizionamento relativo che si ottiene è molto preciso, almeno pari a quello ottenibile con misure classiche di precisione spinta;
- il tempo per l'effettuazione dei rilievi è sensibilmente inferiore a quello dei metodi classici.

Per contro, esistono anche alcuni problemi pratici, che limitano di fatto l'applicazione del metodo. Alcuni di essi sono dovuti a caratteristiche intrinseche del sistema, e difficilmente saranno superati anche in futuro:

- per ogni baseline misurata occorre la disponibilità contemporanea di due ricevitori e di due squadre di operatori (che peraltro possono essere molto ridotte, viste le semplici operazioni da compiere), da disporre sui due punti;
- una stazione GPS richiede la visibilità dei satelliti, ovvero l'assenza di ostacoli fisici al di sopra dell'antenna in un cono visuale ad asse verticale con un angolo al vertice di 150° circa (15° ÷ 20° al di sopra dell'orizzonte);

Altri problemi, presenti allo stato attuale, potranno essere superati nei prossimi anni con il progredire della tecnica e delle sperimentazioni:

il GPS, per il suo principio di misura puramente geometrico, non legato a fenomeni fisici, fornisce la quota ellissoidica dei punti determinati e, quindi, le differenze di quote e non

dislivelli ortometrici. La quota ellissoidica, infatti, non ha significato dal punto di vista fisico ed ingegneristico; ciò che nella pratica interessa è la quota geoidica (o ortometrica).

Il metodo GPS richiede, pertanto, una conoscenza delle ondulazioni geoidiche di precisione adeguata al tipo di livellazione richiesta.

In mancanza di tale conoscenza è necessario il collegamento ad un CAPOSALDO della rete di livellazione IGM.

Il costo della strumentazione è ancora piuttosto elevato, anche se la concorrenza in atto sta favorendo una certa riduzione dei prezzi.

E' comunque evidente che i vantaggi del sistema sono veramente notevoli, e il rapporto tra questi e gli svantaggi non potrà che migliorare con il tempo, consentendo, già dai prossimi anni, un impiego sempre più esteso del GPS.

Il sistema tuttavia non potrà soppiantare del tutto i metodi di rilevamento classici, sempre insostituibili quando la ricezione dei satelliti sia impedita da ostacoli fisici (si pensi alle città, alle zone ricche di vegetazione, alle strette valli montane, ecc.).

Il futuro è, quindi, sicuramente nella integrazione fra GPS e topografia classica; integrazione perfettamente realizzabile, anche se richiederà un'evoluzione del grado di preparazione teorica dei tecnici, che dovranno tra l'altro essere in possesso di conoscenze geodetiche più approfondite che in passato.

7 - La Rete IGM95

Grazie al lavoro svolto recentemente dall'IGMI, l'Italia è fra le prime nazioni europee che si è dotata di una rete geodetica (denominata IGM95) basata sul sistema GPS e realizzata con le procedure proprie di tale sistema. La rete è ormai completata per tutto il territorio italiano ed è certo che la sua disponibilità contribuirà in misura considerevole al rapido affermarsi delle tecniche GPS in Italia.

E' attualmente in via di formazione una banca dati ad essa associata, per la gestione della grande massa di informazioni, sia numeriche che grafiche, prodotte durante la realizzazione della rete.

La rete IGM95 include circa 1150 vertici e costituisce una rete di inquadramento generale del territorio italiano. Essa è l'equivalente del I ordine della rete classica ma ha una densità tre volte superiore rispetto ad essa (che è costituita infatti da circa 360 punti.).

L'innovazione più importante consiste nel fatto che l'IGM95, sfruttando le eccezionali capacità del GPS, non necessita di reti di raffittimento. La rete classica, al contrario, per raggiungere una densità sufficiente al suo utilizzo, ha avuto bisogno di essere raffittita con circa 30.000 vertici (un punto ogni 10 Km quadri circa). E' a tutti evidente la difficoltà di mantenere efficiente una rete costituita da un numero talmente elevato di punti.

L'utilizzo della rete GPS è molto semplice. La densità dei vertici (uno ogni 300 Km quadri circa, con una distanza massima fra i punti di circa 20 Km) è infatti tale che un vertice IGM95 è disponibile, in ogni zona del territorio nazionale, ad una distanza non superiore a 10 km.

In questa condizione, utilizzando due ricevitori, è possibile misurare una base che collega direttamente il punto da determinare ad un vertice IGM95, senza necessità di passaggi intermedi.

Ciò consente, per esempio, di determinare direttamente i punti (PUNTI D'APPOGGIO per la cartografia) necessari all'inquadramento geometrico della zona da cartografare, con vantaggi evidenti dal punto di vista dei tempi di esecuzione e dei costi. I tempi di osservazione saranno, ovviamente, più o meno lunghi a seconda del tipo di strumento utilizzato, della procedura di rilevamento adottata e della precisione desiderata.

I punti della rete IGM95 sono stati tutti scelti in modo che possa essere realizzata comodamente una stazione “in centro” con la strumentazione GPS. Essi non presentano ostruzioni tali da impedire o rendere difficoltosa l’acquisizione dei segnali provenienti dai satelliti. Inoltre, nella maggior parte dei casi, i vertici sono posizionati in prossimità di strade e sono quindi raggiungibili con autovettura. Ciò li rende quindi pienamente sfruttabili in ogni stagione.

Per la materializzazione dei punti sono stati utilizzati centrini in acciaio inox, con possibilità di centramento millimetrico che garantisce il pieno sfruttamento delle precisioni raggiungibili con il GPS.

La realizzazione della rete IGM 95 ha consentito, in ogni zona del territorio italiano, la determinazione dei parametri necessari alla trasformazione dei risultati dal sistema WGS84 (cioè dal sistema di riferimento del GPS) al sistema Gauss-Boaga. Tali parametri, validi per un’area avente un raggio di circa 10 Km intorno al punto, fanno parte integrante della monografia di ogni vertice IGM95, permettendo di passare facilmente dai valori WGS84 al sistema di riferimento nazionale, facilitando in tal modo il pieno sfruttamento delle metodologie satellitari.

Per assicurare un buon collegamento della rete IGM95 con le reti europee e mondiali, 10 punti sono stati posizionati su vertici della rete EUREF ed altrettanti sono stati fatti coincidere con punti di equivalenti reti geodetiche satellitari appartenenti alle nazioni confinanti con l’Italia.

Tutti i vertici sono stati collegati altimetricamente a capisaldi della rete di livellazione di alta precisione e possiedono quindi la quota geoidica. La disponibilità di questo tipo di dati è estremamente importante, come già visto, per consentire la determinazione delle quote ortometriche dei punti determinati mediante GPS.

Per consentire una maggiore precisione del GPS, anche nelle determinazioni altimetriche, il progetto IGM95 prevede, inoltre, l’esecuzione di misure astronomiche e di gravità che contribuiranno a migliorare la conoscenza del geoide nazionale già noto, determinato con sufficiente precisione (± 10 cm) in quasi tutto il territorio nazionale.

CAPITOLO V
TEORIA DELLE MISURE

Capitolo V

TEORIA DELLE MISURE

1- Concetto di Misura

La misura é un procedimento semplice o complesso che permette di quantificare, assegnando dei numeri, le proprietà di grandezze geometriche; nel nostro caso:

- lunghezze, secondo particolari direzioni, quali *distanze* e *dislivelli*;
- angoli, secondo particolari piani, quali *angoli azimutali* ed *angoli zenitali*.

Definita, ad esempio, la distanza tra due punti A e B (cioè il segmento di retta compresa tra i punti A e B) misurare tale grandezza significa determinare un numero da associare a questa entità fisica. Scrivere

$$L_{AB} = 254.87 \text{ m}$$

vuoi dire esprimere, comunque, un concetto complesso visto che:

- è difficile individuare esattamente i due punti A e B, i quali, nella loro materializzazione, possono assumere forme diverse che influiscono sulla loro dimensione minima (picchetto in legno e/o in ferro, chiodino, forellino, ecc.);
- è difficile stabilire il numero di cifre da indicare dopo la virgola, che può assumere anche un valore elevato (infinito?), specie nel caso delle misure indirette, delle quali parleremo più innanzi (intese come risultato di un calcolo).

Il segno di uguaglianza che lega le due entità, quella geometrica e quella numerica diviene, pertanto, molto vago non potendosi paragonare tra loro elementi di natura diversa.

Si vedrà, in seguito, come al concetto di indefinibilità della misura verrà associato il termine di incertezza, che individua, invece con estrema facilità, il criterio di approssimazione del numero indicato a destra del segno di uguale, alla realtà fisica o geometrica da misurare. Per eseguire una misura si utilizzano opportuni strumenti (ottici, elettrici, elettronici, analogici o digitali); qualunque sia la strumentazione utilizzata, l'esecuzione corretta di

una misura richiede sempre la conoscenza dell'unità di misura¹. Le misure si possono distinguere: misure *dirette* e misure *indirette*.

1. Una grandezza è misurata direttamente quando può essere confrontata con una determinata quantità campione che materializza la grandezza geometrica.

In particolare, si parla di *misure dirette* quando si può paragonare direttamente la grandezza incognita con l'unità di misura scelta, mediante un determinato strumento.

La misura diretta di una grandezza è, quindi, definita nella seguente maniera:

- si definisce una *unità campione*²;

¹ Eratostene, nel terzo secolo avanti Cristo, misurando l'angolo del Sole allo zenit tra Alessandria e Syene e conoscendo la distanza fra queste città, ha ricavato la lunghezza della circonferenza terrestre pari a 250.000 stadi. L'esecuzione di misure topografiche ha richiesto, quindi, l'adozione di opportune unità di misura, specie per grandezze quali lunghezza, tempo, massa e angolo. Il *Sistema Metrico decimale* viene istituito in Francia solo nell'aprile del 1795 fissando quattro grandezze fondamentali: *metro, kilogrammo, litro, secondo*. Nel 1875 nasceva la Conferenza generale dei Pesi e delle Misure (CGPM); in questa sede fu definito il concetto di "misure" intendendo come tali le lunghezze e le loro grandezze geometriche derivate.

La nascita del sistema MKSA è stata ufficialmente sancita dal CIPM (Comité International des Poids et Mesures) nel 1946, dopo la seconda guerra mondiale. Il *SISTEMA INTERNAZIONALE* (SI.) è stato legalmente adottato in Italia, con la legge del 14/4/1978, n.122 e con il D.P.R. 12/8/1982, n. 802. Tale sistema metrico è fondato sulle sei fondamentali unità di misura: *metro, kilogrammo, secondo, ampere, Kelvin, candela*.

² Le unità fondamentali del S.I. sono:

- il *metro* (m): lunghezza del tragitto compiuto nel vuoto dalla luce in un intervallo di tempo pari a $1/299.792.458$ di secondo;
- il *kilogrammo* (kg) : massa di un prototipo internazionale;
- il *secondo* (s);
- l' *ampere* (A) : intensità di corrente elettrica;
- il *kelvin* (K): frazione pari a $1/273,16$ della temperatura termodinamica del punto triplo dell'acqua;
- la *candela* (cd): intensità luminosa;
- la *mole* (mol).

Vi sono inoltre due unità supplementari:

- il *radiante* (rad): angolo piano fra due raggi di un cerchio che sottende sulla circonferenza un arco di lunghezza pari al raggio;
- lo *steradiano* angolo solido; e due unità non S.I. ammesse:
- angolo sessagesimale;
- ettaro.

- si sceglie un idoneo strumento che contenga tale unità campione;
- si confrontano le unità dello strumento con la grandezza da misurare;
- si contano il numero di unità intere e quelle frazionate del campione contenute nella grandezza considerata;

La nostra misura sarà data da un numero pari a “n” volte la misura campione più un certo valore “p”, sottomultiplo del campione stesso che individua la parte decimale del numero.

Non avendo a disposizione il *campione vero* ma una sua copia sarà opportuno valutare la *taratura* del campione, ossia controllare se il nostro campione corrisponde all’entità fisica di riferimento.

Tale controllo preliminare si inserisce in una serie di procedure ed accorgimenti relativi alle modalità di effettuazione della misura, di cui parleremo di seguito.

Le misure eseguite in maniera diretta, inoltre, devono soddisfare ad una o più relazioni geometriche che si possono esprimere sotto forma di equazioni, denominate *equazioni di condizione*, ad esempio la somma degli angoli interni di un triangolo piano deve essere uguale a π .

Tale condizione costituirà l’elemento di controllo dei valori attribuiti ai tre angoli misurati.

Le equazioni di condizione, che a volte possono assumere anche espressioni molto complesse, sono utili per verificare l’entità delle incertezze delle misure, delle quali si parlerà successivamente, permettendo di valutare globalmente il grado di attendibilità delle misure eseguite.

L’applicazione di tali procedure di controllo porterà alla metodologia delle *misure dirette condizionate*.

1. Si presenta spesso il caso di dover determinare il valore di grandezze che non sono direttamente misurabili (quote, coordinate di punti, ecc.).

Nel caso planimetrico è possibile ricavare anche misure di angoli e/o di distanze, utilizzando le relazioni matematiche (fig. 1.1)

$$\alpha = \arctg \frac{x_3 - x_1}{y_3 - y_1} - \arctg \frac{x_2 - x_1}{y_2 - y_1}$$

$$d_{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

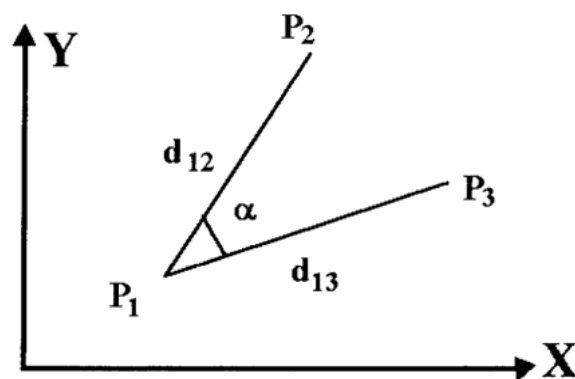


Fig. n. 1.1 - Misura “indiretta” di angoli e distanze

Tali legami funzionali (*equazioni generatrici*) costituiranno l’elemento di controllo delle misure eseguite e permetteranno, come nel caso delle dirette condizionate, di stabilire, nota l’incertezza delle misure direttamente eseguite, quella delle grandezze ricavate indirettamente (ellissi di errore delle coordinate).

La presenza di controlli consente, quindi, di poter valutare la validità della uguaglianza tra ente geometrico ed espressione numerica. Tale validità viene evidenziata dal valore dell’incertezza attribuita al numero che esprime la misura.

L’assenza della incertezza fa ritenere priva di senso l’equazione citata, per cui l’indicazione numerica di una misura priva del parametro “incertezza” non fornisce il risultato della stessa misura, anzi trae in inganno l’utilizzatore che attribuirà tale valore al numero di cifre indicate dopo la virgola.

2 - Gli errori di misura

Nell’esecuzione pratica della misura di una qualsiasi grandezza o entità geometrica, inevitabilmente si presentano degli “inconvenienti” nelle operazioni di misura: essi vengono denominati, per consuetudine, “*errori di misura*”

Si deve sottolineare che l’errore non va inteso come una disattenzione commessa da qualcuno, bensì come la difficoltà del misurare connessa ai problemi della misura stessa; al posto di errore, quindi, sarebbe più opportuno parlare di incertezza nel risultato di una misura, tuttavia la terminologia adottata è più idonea per evidenziarne il carattere incisivo e ineluttabile.

L'errore di misura può essere definito in maniera rigorosa come differenza fra il valore della grandezza osservata nella esecuzione di una generica misura (“osservazione”) ed il valore “vero” della grandezza stessa.

Tale definizione presuppone la conoscenza del valore esatto della misura, risultato impossibile nella sua determinazione pratica per le ragioni citate.

Gli errori che si possono incontrare nella esecuzione di una misura sono di tipo e di natura diversa.

Essi si possono distinguere in: *grossolani, sistematici, accidentali*.

2.1 - Errori Grossolani

Sono quelli attribuibili alla mancanza di accuratezza (in inglese “accuracy”) in sede di esecuzione delle misure, ovvero nella registrazione dei risultati delle stesse.

Bisogna star attenti a non confondere l'errore grossolano con l'errore di grande entità; infatti quest'ultimo è facilmente rilevabile con semplici controlli (ripetizioni delle misure), cosa non sempre immediata quando il numero di misure da eseguire è rilevante; in questo caso diviene necessario il confronto degli “scarti” delle stesse misure con un parametro aleatorio, definito come “tolleranza”, di cui saranno fornite in seguito le giustificazioni concettuali.

Per questa ragione nella tecnica del rilievo si segue sempre il principio del “controllo” delle misure, ovvero si adottano schemi operativi che consentano il controllo e la verifica delle misure.

Si consideri, ad esempio, la misura degli angoli di un triangolo misurandone due si può ricavare il valore del terzo, ma l'operazione di misura non è controllata.

Misurando, invece, tutti e tre gli angoli e verificando lo scarto rispetto a π della somma delle misure ottenute si realizza il controllo delle misure effettuate.

2.2 - Errori Sistematici

Sono prodotti da cause che agiscono in maniera nota, secondo una legge ben determinata, sul risultato della misura.

Esistono due tipi di errori sistematici, diversi tra loro a seconda delle cause caratteristiche che li producono:

- quelli che al ripetersi della misura conservano valore e segno costante;

- quelli che al ripetersi della misura mantengono invariato il segno mentre il valore assoluto varia da una misura all'altra.

Un tipico esempio del primo caso, nella misura di una distanza, è quello dovuto ad una imperfetta taratura dello strumento di misura, denominata *srettifica dello strumento*. Se è nota l'entità dell'errore di taratura dello strumento è evidente che si può apportare "a posteriori" la correzione più opportuna per eliminare la sua influenza.

Per quel che riguarda il secondo tipo, si può citare ad esempio, quello di "non allineamento" nella stesura del campione sulla congiungente i punti A e B.

Ciò comporta un risultato della misura sistematicamente più alta del reale valore della distanza.

Ogni ripetizione della misura comporterà valori diversi per cui il valore assoluto dell'errore sarà diverso pur conservando sistematicamente un segno positivo.

Per loro eliminazione, o la eventuale riduzione al di sotto di certi limiti si adottano diversi metodi:

1. *taratura dei campioni di misura o rettifica degli strumenti con altri caratterizzati da una precisione superiore a quella delle misure a cui sono destinati;*
2. *procedimenti operativi*, un esempio tipico è quello della regola di Bessel per la misura degli angoli orizzontali;
3. *condizioni operative che rendano variabile il loro segno, in modo tale che, mediando i risultati, se ne compensino parzialmente gli effetti, come nel caso della "reiterazione" per la misura degli angoli.*

Va sottolineato che l'esecuzione di una qualunque misura dovrà essere depurata, preventivamente, da tutte le fonti di sistematismi presenti e possibili, per cui l'esame dei vari tipi e metodi di misura dovrà sempre essere preceduto dalla individuazione ed eliminazione degli inconvenienti di cui si conoscono le cause.

2.3 - Fluttuazioni Accidentali

In questa categoria si pongono tutti gli errori non previsti nei primi due casi, ossia quelli prodotti da elementi che sembrano sfuggire ad ogni controllo.

La ripetizione di una misura, depurata dagli errori sistematici ed effettuata in condizioni di massima accuratezza, mostra che:

- i risultati cambiano ad ogni ripetizione in maniera imprevedibile;
- gli scarti tra le diverse misure sono piccoli con riferimento alla precisione

dello strumento e al metodo di misura utilizzato; lo scarto massimo, inoltre, è in genere inferiore ad un determinato limite (tanto più piccolo quanto più accurata è la misura);

- se le ripetizioni sono sufficientemente numerose la frequenza relativa (ossia il rapporto tra il numero di volte in cui ogni valore si è presentato ed il numero totale di misure) tende a stabilizzarsi.

Si può, pertanto, supporre che la fluttuazione accidentale dei risultati di una misura ripetuta sia un *fenomeno aleatorio*. Quanto detto è evidente solo quando le operazioni di misura sono caratterizzate da una elevata precisione; in mancanza di essa la grossolanità stessa degli strumenti e/o dei metodi impiegati per la misura sono tali da mascherare l'influenza delle fluttuazioni accidentali.

2.4 - Metodi per la riduzione degli errori

Dalla classificazione appena fatta possiamo dedurre alcune regole per poter ridurre e/o eliminare la loro influenza:

- *errori grossolani*: porre la massima attenzione nell'esecuzione delle misure per evitare, per quanto possibile, l'errore umano ma, soprattutto, organizzare le misure per poter disporre di elementi di controllo (ad esempio: ripetizione delle misure e verifica degli scarti rispetto alla tolleranza);
 - *errori sistematici*:
 - *studiare la causa che ha prodotto l'errore*, e dedurre le espressioni matematiche che permettono la sua valutazione e/o eliminazione;
 - *inventare un "metodo operativo"*, ad esempio la regola di Bessel o la livellazione "dal mezzo", per causare l'inversione di segno dell'errore e, quindi, la sua eliminazione con la ripetizione in condizioni diverse.
- errori accidentali* (quelli caratterizzati dalla variabilità del segno):
ripetere più volte la misura ed applicare ai risultati i criteri che governano la teoria delle variabili casuali.

Sarà necessario approfondire alcuni argomenti di discipline più vaste (Statistica, Calcolo delle probabilità, ecc.) di cui nel seguito si forniscono le parti indispensabili.

3 - Variabili Statistiche ad una dimensione

La variabile statistica ad una dimensione descrive la modalità con cui un determinato "attributo" (altezza, peso, età, ecc.) si distribuisce fra gli

individui di una “popolazione”.

Una popolazione è in insieme di individui ben definito (una classe, un comune, una regione, ecc.) al cui interno è possibile indicare o scegliere un ben preciso attributo X che può assumere diversi valori argomentali x_1, x_2, \dots, x_s ($s \geq 2$), determinati attraverso il censimento diretto degli N individui della popolazione (indagine, questionario, ecc.).

Dopo aver censito la popolazione si effettua:

- una suddivisione, mediante n intervalli di ugual ampiezza Δx , della differenza fra il valore argomentale più alto e quello più basso;
- l'attribuzione ad ogni intervallo così individuato di un determinato valore argomentale (x_i);
- l'attribuzione a ciascun individuo, compreso nell'intervallo, del valore argomentale scelto;
- la determinazione del numero di individui F_i compresi in ciascun intervallo: si ottiene così una variabile statistica ad una dimensione

$$X \begin{cases} x_1 & x_2 & x_3 & \dots & x_n \\ F_1 & F_2 & F_3 & \dots & F_n \end{cases} \quad \sum_{i=1}^n F_i = N \quad (3.1)$$

La condizione (3.1) assicura che tutti gli individui sono stati presi in considerazione.

Se consideriamo, ad esempio, una classe di studenti aventi diverse altezze, dividiamo l'intervallo di altezza in diverse fasce; ed attribuiamo ad ogni fascia il numero di individui della popolazione (classe) che hanno altezza compresa in quell'intervallo:

$$X \begin{cases} 130 \div 140 & 141 \div 150 & 151 \div 160 & \dots & 201 \div 210 \\ 3 & 5 & 8 & \dots & 1 \end{cases}$$

Dividendo i valori F_i per il numero N di individui della popolazione si ottengono le frequenze relative $f_i = F_i / N$, pertanto la variabile statistica può assumere la seguente forma

$$X \begin{cases} x_1 & x_2 & x_3 & \dots & x_n \\ f_1 & f_2 & f_3 & \dots & f_n \end{cases} \quad \sum_{i=1}^n f_i = 1 \quad (3.2)$$

la variabile statistica può anche essere rappresentata graficamente mediante l'*istogramma*.

La costruzione dell'istogramma viene eseguita riportando lungo l'asse delle ascisse i valori corrispondenti a ciascuna classe ed a cavallo di questi si costruisce un rettangolo, di base pari al valore Δx e di altezza pari ad h_i , tale che $h_i \times \Delta x = F_i$ (l'area del rettangolo rappresenta il numero di individui aventi valore argomentale x_i).

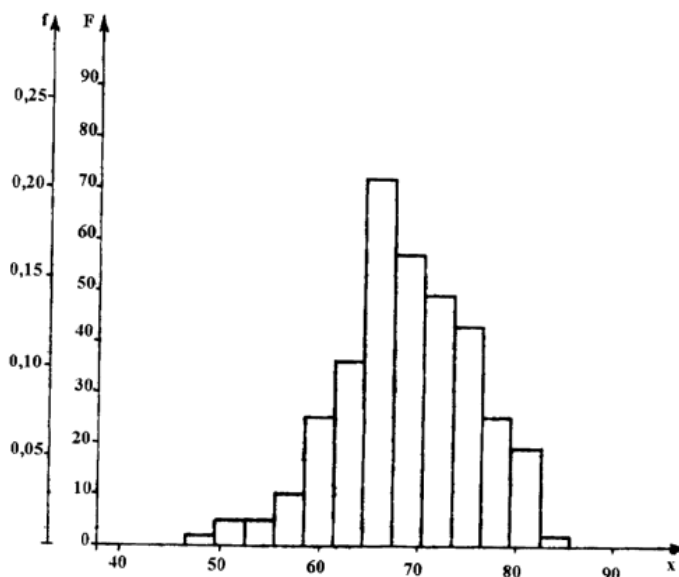


Fig. n. 3.1 - Istogramma

L'istogramma può anche essere costruito facendo riferimento alle frequenze relative.

Si possono calcolare dei parametri che pur non rappresentando compiutamente la variabile, danno delle informazioni globali sulle caratteristiche della distribuzione: questi parametri sono i momenti. Si definisce *momento k.mo* o *momento di k.mo grado* il valore:

$$m_k(x) = \sum_{i=1}^n x_i^k f_i \quad (3.3)$$

Un sufficiente numero di momenti può descrivere completamente la variabile, ovvero può permettere di dedurre tutte le caratteristiche della distribuzione.

I momenti che più interessano sono il *momento di primo grado*, o “*MEDIA*” e quello di *secondo grado*, o *valore quadratico medio* “*V.Q.M.*”

$$m_1(x) = \sum_{i=1}^n x_i f_i \quad \text{momento di I}^\circ \text{ (media)} \quad (3.4)$$

$$m_2(x) = \sum_{i=1}^n x_i^2 f_i \quad \text{momento di II}^\circ \text{ (valore quadratico medio)} \quad (3.5)$$

Da una variabile statistica x si può dedurre la variabile statistica *scarto* v , avente la stessa distribuzione della variabile e valori argomentali pari a $v_i = x_i - m_1(x)$.

Graficamente la variabile scarto è rappresentata dallo stesso istogramma della variabile x , ma con l'origine dell'asse delle ascisse in corrispondenza del valore della media m_1

Un parametro caratteristico della distribuzione è la *varianza* σ^2 ovvero il *momento di secondo grado della variabile scarto*:

$$\sigma^2 = \sum_{i=1}^n v_i^2 f_i = \sum_{i=1}^n (x_i - m_1)^2 f_i \quad (3.6)$$

momento di II° della variabile scarto
 (“*VARIANZA*” o “*DISPERSIONE*”)

con semplici considerazioni matematiche si ha:

$$\sigma^2 = \sum_{i=1}^n (x_i - m_1)^2 f_i = \sum_{i=1}^n x_i^2 f_i + m_1^2 \sum_{i=1}^n f_i - 2m_1 \sum_{i=1}^n x_i f_i = \quad (3.6')$$

$$= m_2 + m_1^2 - 2m_1^2 = m_2(x) - m_1^2(x)$$

che consente di determinare la varianza attraverso la conoscenza dei momenti di primo e secondo grado.

Si può introdurre anche lo *SCARTO QUADRATICO MEDIO* (*S.Q.M.*)

$$\sigma = \pm \sqrt{\sigma^2}$$

L'operatore *media* (m_1) gode di alcune proprietà:

- rende nullo il valore della media della variabile scarto, ossia il momento del primo ordine della variabile scarto

$$\sum_{i=1}^n v_i f_i = \sum_{i=1}^n (x_i - m_1) f_i = \sum_{i=1}^n x_i f_i - M \sum_{i=1}^n f_i = 0$$

- rende minimo il quadrato degli scarti

$$\sum_{i=1}^n v_i^2 f_i = \sum_{i=1}^n (x_i - m_1)^2 f_i = \min$$

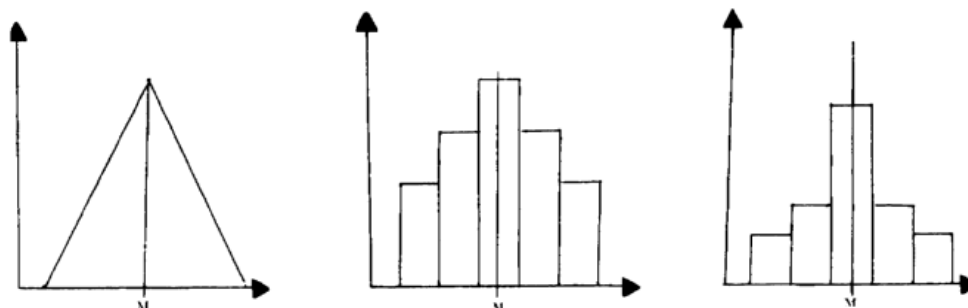
Ciò vuol dire che la derivata della funzione nella variabile x è uguale a zero:

$$2 \sum_{i=1}^n (x_i - m_1) f_i = 0$$

$$\sum_{i=1}^n x_i f_i - M \sum_{i=1}^n f_i = 0$$

Osserviamo che pur conoscendo il valore medio non è possibile descrivere come si distribuiscono i valori intorno alla media.

Ad esempio, supponiamo di avere i seguenti tre casi:



questi pur avendo lo stesso valore medio hanno distribuzioni differenti; per poterle valutare è necessario conoscere le loro varianze. *La varianza, infatti, è l'elemento che permette di valutare "come" i valori argomentali si distribuiscono intorno alla media:*

- per σ^2 alto tali valori risultano “dispersi” (per cui la varianza può anche denominarsi “DISPERSIONE”);
 - per σ^2 basso i valori risultano addensati attorno alla media.
- La media non individua il fenomeno in modo completo, è necessario perciò determinare ogni volta anche il valore della varianza.

3.1- Disuguaglianza di Tchebycheff

La varianza è un indice di dispersione molto significativo in quanto si può dimostrare, in base alla disuguaglianza di TCHEBYCHEFF, che la maggior parte dei valori argomentali di una variabile statistica è contenuta entro i limiti $m_1(x) \pm 3\sigma(x)$ qualunque sia la forma della distribuzione.

Scrivendo infatti per esteso la (6) abbiamo:

$$\sigma^2(x) = \nu_1^2 f_1 + \nu_2^2 f_2 + \dots + \nu_n^2 f_n$$

fissiamo un valore dello scarto ν_m compreso tra ν_1 e ν_n , e poniamo uguali a zero tutti i valori degli scarti inferiori o uguali a ν_m , ed uguali a ν_m , tutti gli scarti superiori; ovviamente si ha che:

$$\sigma^2(x) \geq \nu_m^2 (f_{m+1} + f_{m+2} + \dots + f_n)$$

La somma dei termini tra parentesi definisce la frequenza relativa f_* degli scarti che hanno un valore assoluto superiore a ν_m ; poiché la somma di tutte le frequenze relative è uguale ad uno, la frequenza relativa degli scarti inferiori o uguali a ν_m è

$$f^* = 1 - f_*$$

pertanto

$$\sigma^2(x) \geq \nu_m^2 (1 - f^*)$$

esplicitando f^* otteniamo:

$$f^* \geq 1 - \frac{\sigma^2(x)}{\nu_m^2}$$

ponendo $\nu_m = \lambda\sigma(x)$ si ha la disuguaglianza di TCHEBYCHEFF, significativa per $\lambda > 1$,

$$f^* \geq 1 - \frac{1}{\lambda^2}$$

Per $\lambda=2$, assumendo cioè $\nu_m = 2\sigma(x)$, risulta $f^* = 0,75$ ovvero in qualsiasi distribuzione di una V.S. almeno il 75 % dei valori argomentali è distribuito nell'intervallo $m_1(x) \pm 2\sigma(x)$; per $\lambda=3$ si deduce che almeno l'89 % è distribuito nell'intervallo $m_1(x) \pm 3\sigma(x)$.

Ciò conferma il significato ed il valore della varianza intesa come *dispersione*.

4 - La Variabile Casuale ad una dimensione

Si definisce “aleatorio” o “casuale” un evento di cui non si può prevedere la modalità con cui si verifica (ad esempio il numero ottenuto dal lancio di un dado) pur non essendo stabilita la ragione dell'imprevedibilità; affinché un evento possa definirsi aleatorio è necessario che le modalità con cui esso si verifica siano effettivamente imprevedibili e ripetibili.

Un evento aleatorio che può realizzarsi facilmente e che permette di esemplificare il problema è l'estrazione a caso di un individuo da una popolazione avente una distribuzione, nota ovvero definita da una variabile statistica.

Supponiamo di avere una popolazione, i cui individui sono caratterizzati da un valore argomentale x , e si supponga di aver rilevato il numero degli individui caratterizzati da ciascun valore argomentale:

è possibile a questo punto definire la variabile statistica:

$$X \begin{cases} x_1 & x_2 & x_3 & \dots\dots & x_n \\ f_1 & f_2 & f_3 & \dots\dots & f_n \end{cases} \quad \sum_{i=1}^n f_i = 1$$

Se si estrae a caso un individuo dalla popolazione il valore argomentale x che si ottiene è il risultato di un evento aleatorio che ha tante modalità quanti sono i valori argomentali degli individui della popolazione.

Affinché si abbia un'effettiva estrazione a caso occorre, con riferimento all'atto dell'estrazione, che tutti gli individui siano perfettamente identici, e che la modalità dell'atto che determina l'estrazione sia imprevedibile; ad

esempio si può pensare alla popolazione delle sedi in cui la pallina della roulette può fermarsi, ognuna delle quali è contraddistinta da un numero dallo 0 al 36 (la frequenza di ogni numero è uguale a $1/37$); la sede in cui la pallina andrà a fermarsi è allo stato dei fatti realmente imprevedibile e quindi l'uscita del numero è un evento aleatorio, ottenuto per estrazione a caso.

Supponiamo di eseguire la estrazione a caso N volte (con N grande a piacere) e di rimettere ogni volta l'individuo estratto nella popolazione. Si possono organizzare statisticamente i risultati e definire una variabile statistica che definisce quante volte su N si è presentato ciascun valore argomentale

$$X_c \begin{cases} x_1 & x_2 & x_3 & \dots & x_n \\ f'_1 & f'_2 & f'_3 & \dots & f'_n \end{cases} \quad \sum_{i=1}^n f'_i = 1 \quad \text{variabile statistica campione}$$

All'aumentare del numero N di estrazioni constatiamo che:

1. tutti i valori argomentali sono presenti fra i risultati dell'estrazione
2. le frequenze relative della variabile statistica campione tendono a stabilizzarsi ed a convergere verso il valore della variabile statistica che definisce la distribuzione dell'argomento tra gli individui della popolazione.

Tale convergenza non è rigorosa (nel senso dell'analisi matematica) cioè non è possibile dimostrare che per N crescente esiste un solo valore limite f_i , si può solo dire che sperimentalmente f'_i tende ad f_i .

Tale tendenza è soggetta a fluttuazioni ed è aleatoria anche la perfetta coincidenza di f'_i con f_i .

Quanto espresso nei punti 1) e 2) costituisce la legge empirica del caso.

Se si produce artificialmente un fenomeno aleatorio mediante estrazione a caso da una popolazione avente una distribuzione statistica ben definita, la X_c che viene a costruirsi tende ad assomigliare, al crescere del numero delle estrazioni, alla distribuzione statistica propria della popolazione.

L 'evento statistico ha una stabilità aleatoria nel senso che la distribuzione X_c tende a stabilizzarsi e ad assomigliare a X (ciò vale anche per gli altri eventi aleatori quindi non solo per l'estrazione a caso).

Nel caso di fenomeni aleatori non è mai prevedibile la modalità del verificarsi di un singolo evento mentre si può quasi sempre ottenere una buona previsione di come si *distribuiscono* i risultati di un grande numero di eventi. In altre parole si può dire che un fenomeno aleatorio è conosciuto, se è conosciuta la distribuzione.

La distribuzione che definisce un fenomeno aleatorio viene chiamata **variabile casuale**.

La *variabile casuale* è formalmente identica ad una variabile statistica consta cioè di una serie di valori argomentali a cui sono però associate le *probabilità*, al posto delle frequenze (ben note) della variabile statistica

$$X \begin{cases} x_1 & x_2 & x_3 & \dots & x_n \\ f_1 & f_2 & f_3 & \dots & f_n \end{cases} \quad \sum_{i=1}^n f_i = 1 \quad \text{variabile casuale}$$

ogni valore di probabilità rappresenta il valore a cui tende la frequenza relativa quando, con riferimento alle prove o all'osservazione del fenomeno aleatorio, il numero degli eventi cresce oltre ogni limite.

Le variabili casuali sinora definite sono discrete, constano cioè di un numero finito di valori argomentali a cui sono associati dei valori finiti della probabilità; se costruiamo l'istogramma di una variabile casuale rappresentando l'intervallo Δx piccolo e se supponiamo che i valori delle probabilità varino di poco, è evidente che l'andamento della distribuzione può essere rappresentato da una curva continua:

- *in tal caso la variabile casuale è definita continua, capace, cioè, di assumere tutti i valori compresi in un intervallo a-b, dove a e b rappresentano il minimo e il massimo valore argomentale.*

L'ordinata y della curva, moltiplicata per un intervallo infinitesimo dx fornisce, in accordo con la regola di costruzione degli istogrammi, la probabilità infinitesimale dp associata ai valori argomentali compresi fra x e $x+dx$ (fig. 4.1).

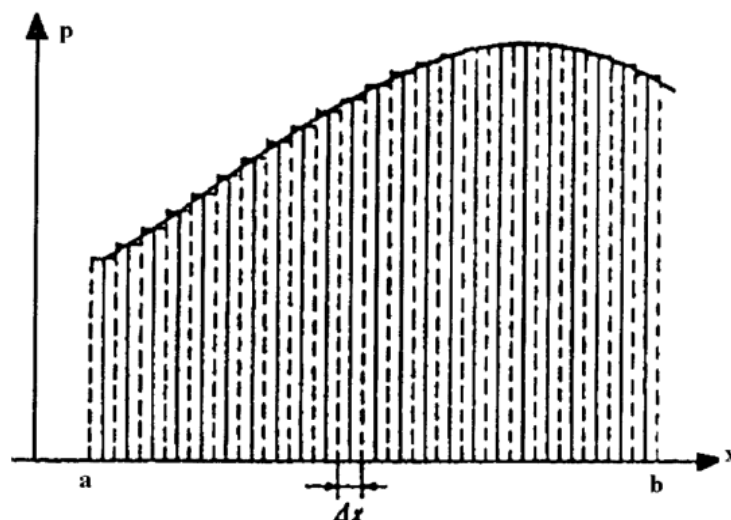


Fig. n. 4.1 - Continuità di una distribuzione di probabilità.

La suddetta curva è definita dalla equazione:

$$y = f(x)$$

si ha pertanto:

$$dp = y dx = f(x) dx$$

La condizione $\sum_{i=1}^n p_i = 1$ si esprime attraverso:

$$\int_a^b f(x) dx = 1$$

La funzione $f(x)$ si chiama *densità di probabilità*.

Per una variabile casuale continua il momento di k.mo grado è fornito dalla:

$$m_k(x) = \int_{-\infty}^{+\infty} x^k dp = \int_{-\infty}^{+\infty} x^k f(x) dx$$

in particolare la varianza è data da

$$\sigma^2 = \int_{-\infty}^{+\infty} (x_i - m_1)^2 f(x) dx$$

Valgono gli stessi significati messi in evidenza per le distribuzioni statistiche e per le distribuzioni discrete di probabilità.

I limiti degli integrali dovrebbero essere uguali ai limiti a e b dell'intervallo di definizione, ma possono, come è stato fatto, essere estesi; infatti se $f(x)$ risulta positiva per ogni valore di x è necessario usare questi limiti, mentre non si introduce nessun errore se la $f(x)$ è nulla in qualche parte dell'asse delle x .

La funzione:

$$F(x) = \int_a^x dp = \int_a^x f(x)dx = \int_{-\infty}^x f(x)dx$$

si chiama funzione di distribuzione rappresenta la probabilità cumulativa di tutti i valori argomentali compresi fra a ed x , ovvero la frequenza relativa limite di tutti i risultati delle prove che risultano compresi fra a e x ; ovviamente $F(b) = 1$.

Una funzione $\eta = \eta(x)$ della variabile casuale x è una variabile casuale.

4.1 - Operazione di media.

Si consideri la variabile casuale discreta

$$X \begin{cases} x_1 & x_2 & x_3 & \dots & x_n \\ p_1 & p_2 & p_3 & \dots & p_n \end{cases} \quad \sum_{i=1}^n p_i = 1$$

la sua media

$$m_1(x) = \sum_{i=1}^n x_i p_i$$

ed il valore quadratico medio

$$m_2(x) = \sum_{i=1}^n x_i^2 p_i$$

il valore quadratico medio può essere interpretato come la media della funzione $\eta = x^2$; infatti la distribuzione dei valori x_i^2 è la stessa dei valori x_i . Ciò vale anche per le variabili casuali continue per qualsiasi funzione $\eta = \eta(x)$ di queste, in quanto il valore medio della η può essere calcolato utilizzando la stessa densità di probabilità della x .

Si indichi con $f(x)$ la densità di probabilità della x e con $g(\eta)$ quella della $\eta(x)$; a rigore la media della $\eta(x)$ andrebbe calcolata con:

$$m(\eta) = \int_{-\infty}^{+\infty} \eta g(\eta) d\eta$$

ma tenendo conto che $g(\eta) d\eta = f(x) dx$ si ha:

$$m_1(\eta) = \int_{-\infty}^{+\infty} \eta f(x) dx \quad (4.1)$$

L'importanza di tale equazione risiede nel fatto che la media di una funzione η della variabile casuale continua x può essere calcolata senza la preventiva determinazione della densità di probabilità $g(\eta)$.

Se nella (4.1) si pone $\eta(x) = x^2$ si ottiene anche per le variabili continue il risultato che la media dei valori x^2 coincide con il momento di secondo grado della x stessa.

In particolare assumendo $\eta(x) = (x - m_1(x))^2 = v^2$ si ha:

$$m_1(v^2) = \int_{-\infty}^{+\infty} (x^2 - 2xm_1(x) + m_1^2(x))f(x)dx = m_2(x) - m_1^2(x)$$

La media assume così il significato di un operatore che genera momenti e come tale può essere denotato con il simbolo $M [\]$, mettendo fra la parentesi quadra la funzione su cui l'operazione di media viene eseguita.

4.2 - Probabilità composta e totale

Un evento aleatorio può risultare da una combinazione di due o più eventi aleatori, ad esempio il lancio di due dadi; tali eventi aleatori, possono anche essere considerati separatamente; in tal caso l'evento è definito da una variabile casuale a due o più dimensioni.

Se invece gli eventi possono essere assoggettati ad una operazione che dia luogo ad un unico risultato, l'evento aleatorio è definito da una variabile casuale ad una dimensione (lanciati i due dadi si fa la somma dei numeri usciti: ad ogni prova si ottiene così un unico risultato).

In tali fenomeni è necessario assumere come ipotesi di base che i fenomeni aleatori che compongono l'evento aleatorio combinato siano indipendenti; (ad esempio se due dadi vengono lanciati in modo che sul movimento di ciascuno di essi non vi sia una sistematica influenza del moto dell'altro si può ritenere che i due fenomeni aleatori siano indipendenti volendo comunque escludere ogni possibilità di interazione si potrebbe lanciare un dado due volte successivamente). Occorre, quindi, valutare la probabilità e quindi la variabile casuale "somma dei numeri che si ottengono lanciando due dadi". Invece di sommare gli argomenti se ne può eseguire la differenza e, in questo caso, si avrà una variabile casuale differenza delle due, così pure dicasi per il prodotto. Si può quindi parlare di variabile casuale somma, differenza o prodotto di due variabili casuali.

Si abbia la variabile casuale

$$X \begin{cases} x_1 & x_2 & x_3 & \dots\dots\dots & x_n \\ p_1 & p_2 & p_3 & \dots\dots\dots & p_n \end{cases} \quad (4.2)$$

e la variabile casuale

$$Y \begin{cases} y_1 & y_2 & y_3 & \dots\dots\dots & y_n \\ g_1 & g_2 & g_3 & \dots\dots\dots & g_n \end{cases} \quad (4.3)$$

assumiamo quale schema di fenomeno aleatorio l'estrazione a caso:

ogni prova consiste cioè nell'estrarre a caso un individuo dalla popolazione possibile definita dalla (4.2) e un individuo dalla (4.3); tutti i valori x si

combineranno con i valori y , dando così luogo ad una serie di valori diversi della somma s

$$s_1 \quad s_2 \quad s_3 \dots \dots \dots s_n$$

con probabilità

$$f_1 \quad f_2 \quad f_3 \dots \dots \dots f_n$$

Se consideriamo un numero N grande ogni limite di doppie estrazioni ed una coppia di valori X_r , di probabilità p_r e y_s di probabilità g_s , essi danno luogo ad un valore somma s_k avente probabilità:

$$p_{rs} = p_r g_s$$

Questo risultato esprime la ***proprietà della probabilità composta***: se un evento E risulta dal concorso di due eventi aleatori indipendenti, di probabilità p_1 e p_2 , la probabilità p_E dell'evento E è data dal prodotto delle probabilità:

$$p_E = p_1 p_2$$

Nel lancio di due dadi, ad esempio, il valore 2 è ottenuto dall'uscita in entrambi i dadi del $n^0 1$; tale risultato ha uguale probabilità, cioè pari ad $1/6$, pertanto la probabilità composta relativa al valore due è pari a $1/36$.

Si deve tener conto dell'eventualità che il valore s_k della somma possa ottenersi anche con un'altra coppia di valori x_u e di y_v di probabilità p_u e g_v , per cui la probabilità del valore s_k composta con i precedenti valori è

$$p_{uv} = p_u g_v$$

Se non vi sono altre modalità per ottenere la somma s_k si avrà infine:

$$f_k = p_{rs} + p_{uv} \quad (4.4)$$

è infatti evidente che computando sulle N estrazioni tutti i casi in cui si è verificata la somma s_k si dovranno sommare tutti quelli provenienti dalla diverse coppie (ad esempio: il risultato del valore 3 è ottenuto nel lancio dei dadi dall'uscita sul primo dado del n° 2 o del n° 1 e rispettivamente sul secondo dado dall'uscita del n° 1 o del n° 2; la probabilità per ciascun risultato è di $1/36$, pertanto quella totale è pari a $1/18$). La (4.4) esprime la **proprietà della probabilità totale**: se un evento aleatorio E può presentarsi con modalità diverse ognuna avente probabilità p_1, p_2 la probabilità di E è la somma delle probabilità

$$p_E = p_1 + p_2$$

Si tratta ora di trovare la media, il valore quadratico medio e la varianza della variabile casuale combinata, note che siano le medie, i valori quadratici medi e le varianze delle variabili casuali componenti:

Se indichiamo con $m_1(x)$, $m_2(x)$, $\sigma^2(x)$, $m_1(y)$, $m_2(y)$, $\sigma^2(y)$, le medie, i valori quadratici medi e le varianze delle variabili (4.2) e (4.3) e con

$$S \begin{cases} s_1 & s_2 & s_3 & \dots & s_n \\ f_1 & f_2 & f_3 & \dots & f_n \end{cases}$$

la variabile casuale somma $s = x + y$, si ottiene:

$$m_1(s) = \sum_{i=1}^n s_i f_i$$

e per quanto precedentemente detto si ha:

$$\begin{aligned} m_1(s) &= \sum_{j=1}^n \sum_{k=1}^m (x_j + y_k) p_{jk} = \sum_{j=1}^n (x_j p_j \sum_{k=1}^m g_k + p_j \sum_{k=1}^m y_k g_k) = \\ &= \sum_{j=1}^n x_j p_j + \sum_{j=1}^n p_j \sum_{k=1}^m y_k g_k = m_1(x) + m_1(y) \end{aligned}$$

da cui si ricorda la:

- **REGOLA n. 1** Il valore medio della somma (o differenza) di due variabili casuali è la somma (differenza) dei valori medi delle variabili componenti

Il momento di secondo grado della variabile casuale somma è dato da:

$$m_2(s) = \sum_{i=1}^n s_i^2 f_i$$

da cui si ottiene:

$$\begin{aligned} m_2(s) &= \sum_{j=1}^n \sum_{k=1}^m (x_j + y_k)^2 p_{jk} = \sum_{j=1}^n \sum_{k=1}^m (x_j^2 + 2x_j y_k + y_k^2) p_{jk} = \\ &= \sum_{j=1}^n \sum_{k=1}^m x_j^2 p_{jk} + 2 \sum_{j=1}^n \sum_{k=1}^m x_j y_k p_{jk} + \sum_{j=1}^n \sum_{k=1}^m y_k^2 p_{jk} = m_2(x) + 2m_1(x)m_1(y) + m_2(y) \end{aligned}$$

la varianza della somma è data, ricordando la formula (3.6'), da:

$$\sigma^2(s) = m_2(s) - m_1^2(s) = m_2(x) + 2m_1(x)m_1(y) + m_2(y) - (m_1(x) + m_1(y))^2 = \sigma^2(x) + \sigma^2(y)$$

da cui ricaviamo la:

- **REGOLA n. 2** La varianza della somma (o differenza) di due variabili casuali è la somma (o la differenza) delle varianze delle variabili componenti.

Inoltre per la variabile prodotto si osserva che:

$$m_2(q) = \sum_{j=1}^n \sum_{k=1}^m x_j^2 y_k^2 p_{jk} = \sum_{j=1}^n x_j^2 p_j \sum_{k=1}^m y_k^2 g_k = m_2(x) \cdot m_2(y)$$

$$\sigma^2(q) = m_2(q) - m_1^2(q) = m_2(x)m_2(y) - m_1^2(x)m_1^2(y) =$$

$$= \sigma^2(x) \cdot \sigma^2(y) + m_1^2(x)\sigma^2(x) + m_1^2(y)\sigma^2(y)$$

da cui si ottengono le:

- REGOLA n.3: *Il valore medio del prodotto di due variabili casuali è il prodotto dei valori medi delle variabili casuali componenti.*
- REGOLA n. 4: *Il valore quadratico medio del prodotto di due variabili casuali è uguale al prodotto dei valori quadratici medi delle variabili componenti.*

Inoltre se si moltiplicano i valori argomentali di una variabile casuale per una costante c , è evidente che la media della nuova variabile è c volte la media della variabile originale, mentre il valore quadratico medio e la varianza risultano moltiplicati per c^2

- REGOLA n 5: *Moltiplicando una variabile casuale per una costante c , si moltiplica per c la media e per c^2 la varianza.*

Se una variabile casuale w è combinazione lineare delle variabili casuali x, y, z , ossia:

$$w = ax + by + cz$$

con a, b, c costanti diverse fra loro si ha che *il valore medio della variabile w è uguale alla somma dei valori medi delle singole variabili, moltiplicate rispettivamente per a, b, c mentre la varianza della w risulta uguale alla somma delle varianze moltiplicate per le costanti al quadrato (LEGGE DI PROPAGAZIONE DELLE VARIANZE):*

$$m_1(w) = am_1(x) + bm_1(y) + cm_1(z) \quad (4.5)$$

$$\sigma^2(w) = a^2\sigma^2(x) + b^2\sigma^2(y) + c^2\sigma^2(z)$$

Tutte le regole menzionate sono valide anche per le variabili casuali continue, basta infatti sostituire alle suddette formule le sommatorie con il segno di integrale.

5 - Distribuzioni di probabilità notevoli.

La distribuzione di probabilità di BERNOLLI è la variabile casuale che consente di prevedere come si distribuiscono i risultati di prove aleatorie ripetute eseguite su una popolazione reale o possibile caratterizzata da due soli valori argomentali; infatti un evento aleatorio caratterizzato da due sole

modalità è rappresentabile con una variabile casuale del tipo:

$$X \begin{cases} x_1 & x_2 \\ q_1 & q_2 \end{cases}$$

dove i valori argomentali associati ad ogni modalità possono essere qualsiasi. Si associa quindi alla modalità “E” il valore x_2 ed all’altra modalità “non E” il valore x_1 .

La variabile casuale può essere determinata censendo una popolazione e costruendo la variabile statistica che, con un semplice cambio di significato diventa, in base alla legge empirica del caso, una variabile casuale di previsione.

Consideriamo la variabile casuale e combiniamola n volte sommando i valori argomentali, come si è visto nel precedente paragrafo. In generale su una prova, consistente in n estrazioni a caso o esperimenti aleatori, si può calcolare, mediante le proprietà della probabilità composta e totale, la probabilità che si verifichi i volte l’evento “E” e $(n - i)$ volte l’evento “non E”. La probabilità associata alla modalità i volte E e $(n - i)$ volte “non E” per la proprietà della probabilità totale, sarà pari a:

$$P(i) = \binom{n}{i} p^i q^{n-i} = \frac{n!}{i!(n-i)!} p^i q^{n-i} \quad (5.1)$$

5.1 - Distribuzione Normale o Gaussiana.

E’ di fondamentale importanza per la teoria della combinazione delle misure e per numerosi altri fenomeni aleatori che rientrano agevolmente in questo modello matematico

Essa può considerarsi una rappresentazione approssimata di quella di BERNOLLI vista la difficoltà, praticamente insormontabile, di quest’ultima nel calcolare i fattoriali per i valori di n molto grandi.

Dalla (5.1) introducendo gli scarti $\nu = i - np$ (ipotizzando n molto grande e ν

sufficientemente piccolo in modo tale che $v < np$) otteniamo:

$$P(v) = \frac{1}{\sqrt{\pi} \cdot \sqrt{2npq}} e^{-\frac{v^2}{2npq}} = \frac{1}{\sqrt{\pi} \cdot \sqrt{2\sigma^2}} e^{-\frac{v^2}{2\sigma^2}}$$

Se n è molto grande i , e quindi lo scarto v , possono assumere un grande numero di valori e non si commette pertanto un errore sensibile considerando continua la variabile casuale così determinata; possiamo parlare, allora, di probabilità infinitesima dp di uno scarto compreso fra v e $v + dv$, ovvero:

$$dp = \frac{h}{\sqrt{\pi}} e^{-h^2 v^2} dv$$

dove si è posto $h = \frac{1}{\sqrt{2} \cdot \sigma}$

In definitiva la densità di probabilità di una variabile casuale normale o gaussiana x di media m e varianza σ^2 è

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

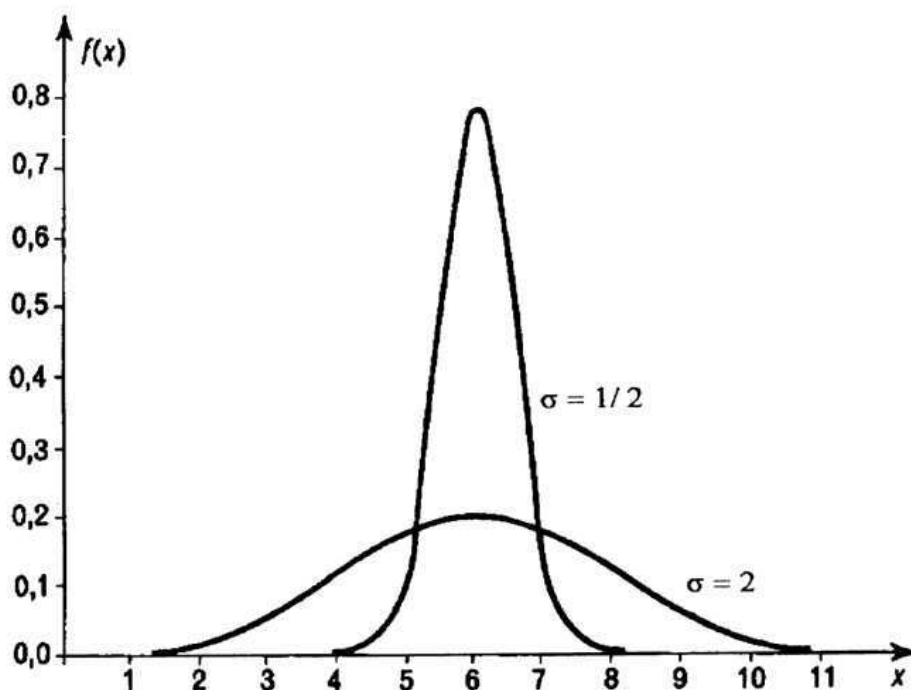


Fig. n. 5.1 - Distribuzioni gaussiane.

Due grafici di questa funzione sono riportati in figura con un valore di $m = 6$ e valori di σ rispettivamente uguali a $\frac{1}{2}$ e 2

La curva è simmetrica rispetto alla retta $x = m$, e si può dimostrare che ha due flessi per $x = m + \sigma$ e $x = m - \sigma$; al crescere di σ^2 ed a parità di media le curve si appiattiscono, ovvero si hanno distribuzioni più disperse, e viceversa diventano più alte e aguzze quando σ^2 diminuisce.

Il parametro $h = \text{misura di precisione}$ è inversamente proporzionale allo scarto quadratico medio.

La funzione di distribuzione

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

ha valore 0 per $x = -\infty$, valore 0,5 per $x = m$ e valore 1 per $x = +\infty$,

Nella seguente fig. 5.2 sono riportati i grafici delle funzioni di distribuzione relative alle due curve della fig. n.5. I

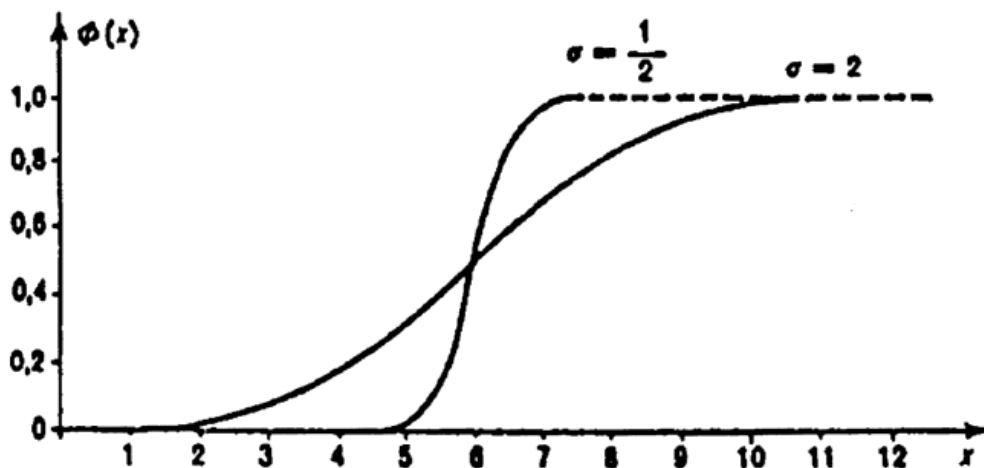


Fig. n. 5.2 - Funzioni di distribuzione delle distribuzioni gaussiane

Per poter tabellare la probabilità dei valori di x compresi fra due limiti a e b conviene riferirsi ad una variabile scarto standardizzata z così definita:

$$z = \frac{x - m}{\sigma}$$

la cui distribuzione è definita dalla densità di probabilità:

$$g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

La distribuzione della variabile standardizzata equivale ad una variabile normale avente media nulla e varianza uguale ad 1.

La funzione di distribuzione

$$G(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt$$

si trova tabulata nella maggior parte dei libri di Statistica Matematica.

Per calcolare la probabilità $P_{a,b}$ dei valori di x compresi fra due limiti a e b basta calcolare i valori

$$z_a = \frac{a - m}{\sigma} \quad z_b = \frac{b - m}{\sigma}$$

e determinare $G(z_a)$ e $G(z_b)$; il valore della probabilità cercata è:

$$P_{a,b} = G(z_b) - G(z_a)$$

E' importante determinare quali sono i valori della probabilità dei valori di x compresi fra i limiti:

$$m - \sigma \leq x \leq m + \sigma$$

$$m - 2\sigma \leq x \leq m + 2\sigma$$

$$m - 3\sigma \leq x \leq m + 3\sigma$$

e dalle tabelle di $G(z)$:

$$\begin{aligned}
P_{m-\sigma, m+\sigma} &= G(1) - G(-1) = 0,683 \\
P_{m-2\sigma, m+2\sigma} &= G(2) - G(-2) = 0,954 \\
P_{m-3\sigma, m+3\sigma} &= G(3) - G(-3) = 0,997
\end{aligned}
\tag{5.2}$$

Ciò significa che in ogni popolazione di eventi casuali, rappresentata da una distribuzione di tipo “normale” o gaussiana, il 68,3 % dei valori argomentali è compreso entro i limiti $m \pm \sigma$; il 95,4% entro i limiti $m \pm 2\sigma$; il 99,7% entro i limiti $m \pm 3\sigma$.

Si noti che il 99,7% dei valori di probabilità corrisponde “quasi” alla certezza ($P = 1$).

Questi valori sono superiori a quelli forniti dalla disuguaglianza di TCHEBYCHEFF ; infatti quella forniva i valori di distribuzione per una VARIABILE STATISTICA qualsiasi, mentre i valori ottenuti nella (5.2) si riferiscono ad una VARIABILE CASUALE di tipo NORMALE.

Da questo risultato scaturisce il concetto di tolleranza, *che costituisce il limite al di là del quale sono definiti “grossolani” gli scarti che superano il triplo dello S. Q.M. (come si vedrà in seguito):*

$$t = 3\sigma$$

6 - Misure dirette di una grandezza come variabili casuali ad una dimensione

Per ogni grandezza esiste un unico numero che ne esprime la misura vera; le fluttuazioni accidentali di misura sono imputabili, quindi, agli errori accidentali presenti in maniera imprevedibile in ogni operazione di misura (le fluttuazioni accidentali di misura sono ineliminabili per principio).

In generale sono a disposizione più osservazioni di una stessa grandezza, con risultati che possono essere leggermente diversi; è necessario vedere come combinare questi risultati in modo da ricavare una misura quanto più possibile vicina alla misura vera.

Considerando il risultato di una singola operazione di misura un evento aleatorio che si può configurare come un'estrazione a caso dalla popolazione di misure possibili e constatata la stabilità stocastica delle fluttuazioni accidentali si può affermare che una grandezza è definita quando è conosciuta la popolazione di misure possibili relative ad un determinato ambiente, ad una determinata strumentazione, ed ad una determinata materializzazione della grandezza. Misurare una grandezza equivale, quindi, a definire la variabile casuale che la rappresenta.

Una distribuzione tipica che rappresenta le fluttuazioni accidentali di misura è quella gaussiana, in quanto una numerosa serie di risultati sperimentali ha confermato che tale distribuzione si presta bene a rappresentare una popolazione di misure possibili.

Ad una legge di distribuzione di tipo gaussiano si arriva anche osservando che il prodursi delle fluttuazioni accidentali di misura sia dovuto al sommarsi di numerosi piccoli effetti aleatori, ognuno dei quali ha una distribuzione qualsiasi. Il teorema di convergenza stocastica afferma, infatti, che la distribuzione di probabilità della somma di n variabili casuali indipendenti tende, per n che va all'infinito, alla forma gaussiana.

La distribuzione di Gauss dipende da due parametri: la media m e la varianza σ^2 . Per definire una grandezza sarà necessario definire i parametri della distribuzione delle misure possibili, sulla base di un certo numero di prove ovvero di misure ripetute.

6.1 - Principio di massima verosimiglianza

Il *PRINCIPIO DI MASSIMA VEROSIMIGLIANZA* è un principio assiomatico in grado di fornire una stima dei parametri della distribuzione.

Sia $f(x)$ la densità di probabilità di una distribuzione dipendente da r parametri $\vartheta_1, \vartheta_2, \dots, \vartheta_r$ e sia x_1, x_2, \dots, x_n una n -pla di valori ottenuti da una estrazione a caso.

La densità di probabilità di questa distribuzione è:

$$V = P(x_1, x_2, \dots, x_n, \vartheta_1, \vartheta_2, \dots, \vartheta_r) = f(x_1)f(x_2)\dots f(x_n) \quad (6.1)$$

(prodotto delle probabilità - *funzione di verosimiglianza*)

I parametri ϑ possono essere variamente stimati ottenendo per ciascuna r -pla diversi valori della densità di probabilità; maggiore verosimiglianza o plausibilità hanno quelle stime dei parametri che rendono massima la P , ciò vuol dire soddisfare il sistema:

$$\frac{\partial P}{\partial \vartheta_1} = 0 \quad \frac{\partial P}{\partial \vartheta_2} = 0 \quad \dots \quad \frac{\partial P}{\partial \vartheta_r} = 0 \quad (6.2)$$

che definisce il massimo e consente di ricavare i valori più plausibili. AI

variare del campione x_1, x_2, \dots, x_n variano le stime più plausibili di $\vartheta_1, \vartheta_2, \dots, \theta_r$.
 Se le $f(x_i)$ sono distribuzioni di tipo gaussiano il principio di massima verosimiglianza conduce al *principio dei minimi quadrati*.
 In tal caso si ha:

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-m)^2}{2\sigma^2}} \quad (6.3)$$

$$P(x_1, x_2, \dots, x_n, m, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2} \quad (6.4)$$

che diviene massima quando è minimo l'esponente di e :

$$-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2 = \min \quad (6.5)$$

per σ^2 costante e ponendo $v_i = x_i - m$ si ha:

$$\sum_{i=1}^n v_i^2 = \min \quad (6.6) \quad \textbf{PRINCIPIO DEI MINIMI QUADRATI}$$

Attraverso questa condizione è possibile ricavare una stima della media. Qualunque funzione del campione x_1, x_2, \dots, x_n può fornire la stima dei parametri m e σ^2 ,

$$\begin{aligned} \overline{m} &= h(x_1, x_2, \dots, x_n) \\ \overline{\sigma^2} &= g(x_1, x_2, \dots, x_n) \end{aligned} \quad (6.7)$$

dove le funzioni h e g sono degli estimatori rispettivamente della media e della varianza (le stime sono indicate con il soprassegno). E' necessario stabilire dei criteri per confrontare tra loro stime diverse di uno stesso parametro, ovvero stabilirne delle proprietà.

La stima può, quindi, essere:

- *consistente*: se per n che tende all'infinito la stima tende al valore teorico;
- *non affetta da errore sistematico* se la media delle stime coincide con la media della popolazione da cui i campioni sono estratti;
- *efficiente* se caratterizzata da minima varianza (valutata relativamente ad altre stime).

6.2 - Stima della media e della varianza

Estratto a caso un campione x_1, x_2, \dots, x_n dell'insieme aleatorio caratterizzato da una distribuzione gaussiana, si può assumere come stima della media e della varianza rispettivamente le:

$$\bar{m} = O_m = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (6.8)$$

$$\bar{\sigma}^2 = s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{m})^2 \quad (6.9)$$

La stima (6.8) ottenuta per via aritmetica coincide con quella fornita dal principio dei minimi quadrati la quale è consistente e non è affetta da errore sistematico.

Derivando, infatti la (6.6) rispetto ad m si ha:

$$\frac{d}{dm} \sum_{i=1}^n (x_i - m)^2 = -2 \sum_{i=1}^n (x_i - m)$$

uguagliando a zero si ha:

$$\sum_{i=1}^n x_i = n\bar{m}$$

da cui si ricava la (6.8).

La stima (6.8) non è affetta da errore sistematico, infatti, con l'operatore media M si ottiene:

$$M[O_m] = \frac{1}{n} (M[x_1] + M[x_2] + \dots + M[x_n]) = \frac{nm}{n} = m$$

che dimostra la consistenza della stima dato che la relazione è valida per qualsiasi n . La popolazione delle stime O_m ha varianza:

$$\sigma_m^2 = \frac{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}{n^2}$$

ed essendo $\sigma_1 = \sigma_2 = \dots = \sigma_n$

$$\sigma_m^2 = n \frac{\sigma_n^2}{n^2} = \frac{\sigma^2}{n}$$

La stima della varianza s^2 è invece affetta da errore sistematico; introducendo, infatti, il valore m della media teorica si ha:

$$\begin{aligned} s^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - m + m - \bar{m})^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 + \frac{1}{n} \sum_{i=1}^n (m - \bar{m})^2 + \frac{2}{n} \sum_{i=1}^n (x_i - m)(m - \bar{m}) \end{aligned}$$

e poiché:

$$\begin{aligned} \frac{2}{n} \sum_{i=1}^n (x_i - m)(m - \bar{m}) &= \frac{2}{n} (m - \bar{m}) \sum_{i=1}^n (x_i - m) = \frac{2}{n} (m - \bar{m}) (\sum_{i=1}^n x_i - nm) = \\ &= \frac{2}{n} (m - \bar{m}) (n\bar{m} - nm) = -2(\bar{m} - m)^2 \end{aligned}$$

si ha quindi:

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 - (\bar{m} - m)^2$$

per cui utilizzando l'operatore M si ha:

$$M[s^2] = \frac{1}{n} \sum_{i=1}^n M[(x_i - m)^2] - M[(\bar{m} - m)^2] = \sigma^2 - \frac{\sigma^2}{n} = \sigma^2 \frac{n-1}{n}$$

La media della popolazione degli s^2 non coincide con σ^2 , pertanto la stima (6.9) è affetta da errore sistematico per difetto.

Per ottenere una stima non affetta da errore sistematico è sufficiente moltiplicare la (6.9) per $\frac{n}{n-1}$

pertanto:

$$\bar{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{m})^2 \quad (6.10)$$

Le stime ricavate con la (6.9) e la (6.10) sono entrambe consistenti; infatti per n sufficientemente grande forniscono quasi lo stesso risultato, la differenza è invece sensibile quando n vale poche unità.

La stima della varianza della media si può quindi scrivere:

$$\bar{\sigma}^2 = \frac{\sigma^2}{n} = \frac{\sum_{i=1}^n (x_i - \bar{m})^2}{n(n-1)}$$

Per misurare direttamente una grandezza si eseguono quindi n osservazioni (x_1, x_2, \dots, x_n) se ne effettua la media aritmetica ottenendo la stima \bar{m} , si calcola la stima della varianza della media $\bar{\sigma}_m^2$ e si sintetizza il risultato scrivendo che la misura è data da:

$$\bar{m} \pm \bar{\sigma}_m$$

La probabilità totale dei valori compresi fra

- $\bar{m} - \bar{\sigma}_m$ e $\bar{m} + \bar{\sigma}_m$ è 0.68
- $\bar{m} - 2\bar{\sigma}_m$ e $\bar{m} + 2\bar{\sigma}_m$ è 0.95
- $\bar{m} - 3\bar{\sigma}_m$ e $\bar{m} + 3\bar{\sigma}_m$ è circa uguale a 1

come già evidenziato a pag. 200.

L'intervallo $\bar{m} \pm 3\bar{\sigma}_m$, rappresenta la massima variazione possibile dei valori di \bar{m} .

Si assumerà affetta da errore grossolano quella osservazione che differisce dalla media \bar{m} per più di $3\bar{\sigma}_m$.

La variabilità di σ_m è inversamente proporzionale alla \sqrt{n} ; pertanto all'aumentare di n si ha la diminuzione dello s.q.m. della media e quindi l'aumento di precisione della misura.

Quest'aumento, tuttavia, non può essere molto spinto poiché i benefici crescono con la radice quadrata dei numeri della ripetizione delle osservazioni, pertanto all'aumentare del numero di osservazioni si hanno incrementi di precisione via via decrescenti (per 4 osservazioni lo s.q.m. si dimezza, per 9 osservazioni si riduce ad 1/3 e così via).

In conclusione non ha senso spingere il numero delle osservazioni oltre certi valori ($n = 25 \div 36 \dots$) essendo illusorio il vantaggio della riduzione del contributo degli errori accidentali, vista la possibile permanenza, a volte, di ulteriori contributi da parte dei sistematici.

6.3 - La media ponderata

Se le misure sono eseguite n volte con strumenti, metodi e condizioni operative differenti sarà necessario considerare non più la media aritmetica bensì la *media ponderata*.

Siano O_1, O_2, \dots, O_n i risultati di una media di un certo numero di estrazioni eseguite sulla stessa popolazione, aventi la stessa media teorica (perché riferite alla stessa misura) ma varianze diverse (in quanto di precisione diversa).

Le osservazioni O_1, O_2, \dots, O_n rappresentano quindi la *popolazione delle medie* ed a queste sono riferite le varianze σ_i^2 .

Relativamente agli elementi che caratterizzano la precisione delle n popolazioni si possono considerare due ipotesi:

1. sono note le varianze σ_i^2
2. sono note le varianze σ_i^2 a meno di un fattore di proporzionalità, ovvero sono note le quantità pi denominate *pesi* pari a:

$$p_i = \frac{\sigma_0^2}{\sigma_i^2}$$

dove σ_0^2 rappresenta la *varianza dell'unità di peso* ovvero la varianza della popolazione a cui è stato attribuito il peso unitario.

Nel primo caso la stima della media comune a tutta la popolazione e lo s.q.m.

possono derivarsi direttamente dalla conoscenza delle O_i e delle σ_i^2 ; nel secondo caso note le p_i occorre procedere ad una valutazione del valore della costante σ_0^2 .

In entrambi i casi, preliminarmente, si procede alla valutazione dei pesi delle singole misure:

$$p_1 = \frac{\sigma_0^2}{\sigma_1^2} \quad p_2 = \frac{\sigma_0^2}{\sigma_2^2}; \dots; \quad p_n = \frac{\sigma_0^2}{\sigma_n^2}$$

assumendo σ_0^2 arbitrariamente, ed in modo che i pesi risultino numeri espressi al massimo da due cifre (dell'ordine di qualche unità o decina).

La *densità di probabilità* relativa alla generica osservazione O_i è espressa in questo caso attraverso la:

$$f(O_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(O_i-m)^2}{2\sigma_i^2}} = \frac{1}{\sqrt{2\pi\sigma_0^2}} \sqrt{p_i} \cdot e^{-\frac{p_i}{2\sigma_0^2}(O_i-m)^2}$$

La funzione di *verosimiglianza* è data da:

$$V(O_1, O_2, \dots, O_n, m, \sigma_0^2) = \frac{\sqrt{p_1 p_2 \dots p_n}}{\sqrt{(2\pi\sigma_0^2)^n}} e^{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n p_i (O_i - m)^2}$$

Qualunque sia il valore di σ_0^2 il massimo della funzione di verosimiglianza si ottiene per:

$$Q = \sum_{i=1}^n p_i (O_i - m)^2 = \min$$

ovvero:

$$\frac{dQ}{dm} = -2 \sum_{i=1}^n p_i (O_i - m) = 0 \Rightarrow$$

$$m = \overline{m}_p = \frac{p_1 O_1 + p_2 O_2 + \dots + p_n O_n}{p_1 + p_2 + \dots + p_n} \quad (6.11) \text{ media ponderata}$$

Il valore \overline{m}_p , si chiama *media ponderata* delle n osservazioni O_1, O_2, \dots, O_n di diversa precisione.

Al variare del campione O_1, O_2, \dots, O_n la media ponderata \overline{m}_p definisce un insieme aleatorio di media m , infatti:

$$M[O_i] = m$$

$$M[\overline{m}]_p = \frac{p_1 M[O_1] + p_2 M[O_2] + \dots + p_n M[O_n]}{p_1 + p_2 + \dots + p_n} = \frac{m \sum_{i=1}^n p_i}{\sum_{i=1}^n p_i} = m$$

La media ponderata è una stima efficiente.

Per la varianza σ_p^2 si ha:

$$\sigma_p^2 = \left(\frac{p_1}{\sum p_i} \right)^2 \sigma_1^2 + \sigma_p^2 = \left(\frac{p_2}{\sum p_i} \right)^2 \sigma_2^2 + \dots + \left(\frac{p_n}{\sum p_i} \right)^2 \sigma_n^2 = \frac{1}{\left(\sum_{i=1}^n p_i \right)^2} \sum_{i=1}^n p_i^2 \sigma_i^2 \quad (6.12)$$

facilmente determinabile se sono note le σ_i^2 , altrimenti si dovrà procedere ad una stima di σ_0^2 , notando che la formula è indipendente da quest'ultimo:

$$\sigma_p^2 = \frac{1}{\left(\sum_{i=1}^n p_i \right)^2} \sum_{i=1}^n p_i \sigma_0^2 = \sigma_0^2 \frac{\sum_{i=1}^n p_i}{\left(\sum_{i=1}^n p_i \right)^2} = \sigma_0^2 \frac{1}{\sum_{i=1}^n p_i} \quad (6.13)$$

La stima di σ_0^2 deve basarsi sui campioni O_1, O_2, \dots, O_n per cui al variare dei campioni varia la stima della varianza dell'unità di peso che risulta pertanto una quantità aleatoria. Per tale stima è preferibile utilizzare il logaritmo della funzione di verosimiglianza:

$$\ln V = \frac{1}{2} \ln(p_1, p_2, \dots, p_n) - \frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma_0^2 - \frac{1}{2\sigma_0^2} \sum_{i=1}^n p_i (O_i - m)^2$$

che diventa massima per il valore di σ_0^2 che soddisfa l'equazione:

$$\begin{aligned} \frac{\partial \ln V}{\partial \sigma_0^2} &= 0 \Rightarrow \\ -\frac{n}{2\sigma_0^2} + \frac{1}{2\sigma_0^4} \sum_{i=1}^n p_i (O_i - m)^2 &= 0 \end{aligned}$$

ponendo in luogo della media teorica la stima \bar{m}_p si ha:

$$\sigma_0^2 = \frac{1}{n} \sum_{i=1}^n p_i (O_i - m)^2 = \frac{1}{n} \sum_{i=1}^n p_i v_i^2$$

Con procedimento analogo a quello del paragrafo precedente si trova che:

$$M[\sigma_0^2] = \frac{n-1}{n} \sigma_0^2$$

pertanto la precedente stima è affetta da errore sistematico, mentre non lo è la stima:

$$\bar{\sigma}_0^2 = \frac{n}{n-1} \sigma_0^2 = \frac{1}{n-1} \sum_{i=1}^n p_i v_i^2 \quad (6.14)$$

La stima $\bar{\sigma}_0^2$ dipende dai pesi che si sono scelti; moltiplicando questi ultimi per una costante anche la stima della varianza dell'unità di peso risulta moltiplicata per la stessa costante, mentre non cambia la media ponderata.

Si può concludere che nel caso in cui sono note le σ_i^2 la (6.12) fornisce la varianza della popolazione della media ponderata, indipendentemente dal campione usato, mentre la σ_p^2 calcolata con il valore stimato della (6.14) dà una indicazione sulla precisione di quel particolare valore della media

ponderata; è conveniente, quindi, procedere a quest'ultimo tipo di calcolo quando sono note le σ_i^2 , anche se i due valori non dovrebbero differire troppo (una forte differenza potrebbe essere imputata all'influenza di eventuali errori sistematici).

7 - Misura indiretta di una grandezza

Sia $X = f(X_1, X_2, \dots, X_n)$ una funzione che lega la grandezza X alle n grandezze direttamente o indirettamente misurabili.

Qualora le n grandezze siano direttamente misurabili risultano essere indipendenti e caratterizzate da una media X_{im} , ed una varianza σ_i^2 ; se invece sono a loro volta misure indirette in generale risultano essere correlate e caratterizzate dalla matrice di varianza e covarianza U della distribuzione ad n dimensioni X_1, X_2, \dots, X_n

$$U = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_1^2 & \dots & \sigma_{n2} \\ \dots & \dots & \dots & \dots \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_n^2 \end{pmatrix} \quad \text{matrice varianza-covarianza}$$

In entrambe i casi la misura X è definita come una variabile casuale le cui caratteristiche dipendono da quelle delle distribuzione.

Se il legame funzionale è espresso da una combinazione lineare delle X_i si ha:

$$X = a_1 X_1 + a_2 X_2 + \dots + a_n X_n \quad a_i = \text{costanti (note)}$$

pertanto la media e la varianza saranno date dalle seguenti espressioni:

$$\begin{aligned} X_m &= a_1 X_{1m} + a_2 X_{2m} + \dots + a_n X_{nm} \\ \sigma^2 X &= a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \dots + a_n^2 \sigma_n^2 \end{aligned}$$

Se invece le X_i , sono misure indirette, non potendo fare nessuna ipotesi sulla funzione, si procede all'estrazione, da ognuna delle popolazioni di misure possibili, di un campione O_1, O_2, \dots, O_n , ossia ogni O_i è il risultato di una determinata osservazione $O_i = X_{im} + v_i$.

Nella ipotesi che gli scarti v_i siano piccoli tanto da poter trascurare i loro quadrati, i prodotti misti e le potenze superiori si ha:

$$f(O_1, O_2, \dots, O_n) = f(X_1 + v_1, X_2 + v_2, \dots, X_n + v_n) =$$

$$f(X_{1m}, X_{2m}, \dots, X_{nm}) + \left(\frac{\partial f}{\partial X_1} \right)_m v_1 + \left(\frac{\partial f}{\partial X_2} \right)_m v_2 + \dots + \left(\frac{\partial f}{\partial X_n} \right)_m v_n \quad (7.1)$$

i valori delle derivate sono calcolati nei punti medi.

Utilizzando l'operatore media si può osservare che la media della popolazione $f = (O_1, O_2, \dots, O_n)$ è uguale a $f(X_{1m}, X_{2m}, \dots, X_{nm})$, visto che al secondo membro figurano gli scarti che sappiamo avere media nulla, in altre parole:

$$X_m = f(X_{1m}, X_{2m}, \dots, X_{nm})$$

ovvero la media teorica della grandezza X , determinata indirettamente, si ottiene introducendo nella funzione le medie teoriche delle popolazioni X_1, X_2, \dots, X_n .

Con procedimento analogo utilizzando l'operatore v.q.m. nella espressione (7.1), si ottiene che la varianza teorica della grandezza X è data dalla somma:

$$\sigma_x^2 = \left(\frac{\partial f}{\partial X_1} \right)_m^2 \sigma_1^2 + \left(\frac{\partial f}{\partial X_2} \right)_m^2 \sigma_2^2 + \dots + \left(\frac{\partial f}{\partial X_n} \right)_m^2 \sigma_n^2$$

(LEGGE DI PROPAGAZIONE DELLE VARIANZE)

CAPITOLO VI

COMPENSAZIONI

RIGOROSE

COMPENSAZIONI RIGOROSE

1 - Metodo delle osservazioni indirette

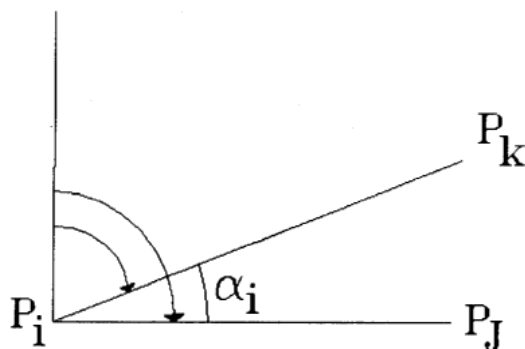
E' noto che alcune grandezze quali le coordinate, le quote, ecc. vengono determinate indirettamente in funzione di altre grandezze (distanze, angoli, dislivelli) che possono essere direttamente misurate sul terreno con strumentazione specifica, ovvero a loro volta ricavate indirettamente tramite ulteriori grandezze misurate (misure indirette). In tal caso è possibile scrivere delle relazioni che legano le grandezze incognite a quelle direttamente misurabili, del tipo:

$$f_i(X_1, X_2, \dots, X_m / L_i, M_i, N_i) = 0 \quad i = 1, 2, \dots, n \quad (1.1)$$

Citiamo ad esempio le equazioni planimetriche (non lineari) che legano le coordinate X,Y con i valori delle distanze e/o degli angoli orizzontali:

$$d_{ij} - \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} = 0$$

$$\alpha_i - \left[\arctg \frac{(X_i - X_j)}{(Y_i - Y_j)} - \arctg \frac{(X_k - X_j)}{(Y_k - Y_j)} \right] = 0$$



ovvero la più semplice relazione altimetrica lineare:

$$Q_j - Q_i + \Delta_{ij} = 0$$

Supposti noti i valori approssimati $X_{1m}^{(0)}, X_{2m}^{(0)}, \dots, X_{mm}^{(0)}$ ricavabili da un gruppo (m) qualunque delle (n) misure, si pone:

$$X_{1m} = X_{1m}^{(0)} + x_1, \quad X_{2m} = X_{2m}^{(0)} + x_2, \dots, \quad X_{mm} = X_{mm}^{(0)} + x_m \quad (1.2)$$

avendo indicato con x_i le *correzioni* (quantità abbastanza piccole da poter ritenere trascurabili i quadrati e le potenze superiori) Introducendo le (1.2) nelle (1.1) si ha:

$$f_i(X_{im}^{(0)} + x_1, X_{2m}^{(0)} + x_2, \dots, X_{mm}^{(0)} + x_m / L_i, M_i, N_i,) = 0 \quad (1.3)$$

Sviluppando la (1.3) in serie ed arrestando i termini a quelli lineari si ottiene:

$$\left(\frac{\partial f_i}{\partial X_1}\right)_0 x_1 + \left(\frac{\partial f_i}{\partial X_2}\right)_0 x_2 + \dots + \left(\frac{\partial f_i}{\partial X_m}\right)_0 x_m + f_i(X_{1m}^{(0)}, X_{2m}^{(0)}, \dots, X_{mm}^{(0)} / L_i, M_i, N_i,) = 0$$

ponendo:

$$a_i = \left(\frac{\partial f_i}{\partial X_1} \right)_0, \quad b_i = \left(\frac{\partial f_i}{\partial X_2} \right)_0, \dots, m_i = \left(\frac{\partial f_i}{\partial X_m} \right)_0$$

$$T_i = -f_i(X_{1m}^{(0)}, X_{2m}^{(0)}, \dots, X_{mm}^{(0)} / L_i, M_i, N_i)$$

dove T_i indica uno o più elementi noti (che in questo caso sono misure indirette di determinate grandezze dedotte da misure dirette di altre grandezze), si ottiene il sistema lineare

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 + \dots + m_1x_m &= T_1 \\ a_2x_1 + b_2x_2 + c_2x_3 + \dots + m_2x_m &= T_2 \\ \dots & \\ a_nx_1 + b_nx_2 + c_nx_3 + \dots + m_nx_m &= T_n \end{aligned} \quad (1.4)$$

che può essere scritto in forma matriciale attraverso:

$$CX = T$$

Si valutano le varianze dei termini noti attraverso la relazione (valida ove esiste indipendenza delle misure L_i, M_i, N_i):

$$\sigma_{T_i}^{-2} = \left(\frac{\partial f_i}{\partial L_i} \right)_0^2 \sigma_{L_i}^2 + \left(\frac{\partial f_i}{\partial M_i} \right)_0^2 \sigma_{M_i}^2 + \left(\frac{\partial f_i}{\partial N_i} \right)_0^2 \sigma_{N_i}^2$$

Se $\bar{L}_i, \bar{M}_i, \bar{N}_i$ sono le stime delle grandezze osservate direttamente e \bar{T}_i i valori che assumono in corrispondenza i termini T_i , si ha:

$$T_i = -f_i(X_{1m}^{(0)}, X_{2m}^{(0)}, \dots, X_{mm}^{(0)} / \bar{L}_i, \bar{M}_i, \bar{N}_i)$$

e

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 + \dots + m_1x_m - \bar{T}_1 &= \bar{v}_1 \\ a_2x_1 + b_2x_2 + c_2x_3 + \dots + m_2x_m - \bar{T}_2 &= \bar{v}_2 \\ &\vdots \\ a_nx_1 + b_nx_2 + c_nx_3 + \dots + m_nx_m - \bar{T}_n &= \bar{v}_n \end{aligned} \quad (1.5)^1$$

si ottiene così un sistema di equazioni alle correzioni che in forma matriciale è dato da:

$$CX - \bar{T} = V$$

Mentre il sistema delle equazioni generatrici (1.4) era iperdeterminato, il sistema generato (1.5) è indeterminato in quanto anche i v_i sono delle incognite (si hanno pertanto n equazioni nelle $m+n$ incognite).

¹ Si noti che le equazioni (1.5) dette “equazioni alle misure” non possono essere uguali a zero sia perché i valori X_{im}^0 sono approssimati, sia perché le misure $\overline{L}_i, \overline{M}_i, \overline{N}_i$ sono comunque affette da incertezza.

L'indeterminazione scompare se fra le infinite $\infty^{(n+m)^{-n}} = \infty^m$ soluzioni si considerano i valori delle x_i che godono delle più alta probabilità, ossia quelle soluzioni che, per il *principio dei minimi quadrati*, rendono minimi gli scarti:

$$\sum_{i=1}^n p_i v_i^2 = \min \quad (1.6)$$

Per la determinazione delle correzioni si procede in tale maniera:

- Si fissa un valore arbitrario σ_0^2 (a priori) e si valutano i pesi $p_i = \sigma_0^2 / \overline{\sigma_{T_i}^2}$, per cui i coefficienti ed i termini noti della (1.5) saranno moltiplicati per le rispettive radici quadrate dei pesi.
- Si ottiene successivamente il sistema normale unendo al sistema (1.5) le relazioni derivanti da (1.6):

$$\frac{\partial p_i v_i}{\partial x_i} = 0; \quad \frac{\partial p_i v_i}{\partial x_2} = 0; \quad ; \dots; \quad \frac{\partial p_i v_i}{\partial x_m} = 0;$$

pari ad “m” relazioni che unite alle “n” relazioni (1.5) conducono al:

$$\text{“SISTEMA NORMALE” : } \mathbf{NX}=\mathbf{Q} \quad (1.7)$$

dove:

$$N = \begin{pmatrix} [paa] & [pab] & \dots & [pam] \\ [pab] & [pbb] & \dots & [pbm] \\ \dots & \dots & \dots & \dots \\ [pau] & [pbu] & \dots & [pum] \end{pmatrix}$$

$$N = CPC^T$$

$$Q = \begin{pmatrix} [pa\bar{T}] \\ [pb\bar{T}] \\ \dots \\ [pu\bar{T}] \end{pmatrix} =$$

$$Q = TPC^T$$

- Si risolve il sistema (1.7) e si ottengono le correzioni x_i che sommate ai valori approssimati X_{im}^0 danno le stime delle X_i .

Qualora siano note le varianze $\sigma_{L_i}^2, \sigma_{M_i}^2, \sigma_{N_i}^2$, è possibile scrivere:

$$\sigma_{T_i}^2 = \left(\frac{\partial f_i}{\partial L_i} \right)_m^2 \sigma_{L_i}^2 + \left(\frac{\partial f_i}{\partial M_i} \right)_m^2 \sigma_{M_i}^2 + \left(\frac{\partial f_i}{\partial N_i} \right)_m^2 \sigma_{N_i}^2$$

secondo la “Legge di Propagazione delle Varianze”.

Essendo noti i pesi delle equazioni si procede al calcolo della stima $\bar{\sigma}_0^{-2}$ (a posteriori):

$$\bar{\sigma}_0^{-2} = \frac{\sum_{i=1}^n p_i v_i^2}{n-m}$$

dove per la determinazione degli scarti v_i si introducono le correzioni ottenute nel sistema (1.5)

Per cui si può ricavare l'incertezza delle correzioni trovate con:

$$\bar{\sigma}_{x_i}^{-2} = \bar{\sigma}_0^{-2} \alpha_{kk}$$

dove α_{kk} sono gli elementi della diagonale principale della matrice inversa dei coefficienti del sistema normale (N^{-1}).

La valutazione di $\bar{\sigma}_0^{-2}$ viene fatta anche se sono note le $\sigma_{T_i}^2$ per stabilire la precisione delle particolari stime ottenute.

Note le incertezze delle singole incognite (coordinate X_i, Y_i ovvero quote Q_i è possibile valutare la figura geometrica denominata “ellisse di errore” (avente semiassi a, b e direzione del semiasse maggiore pari a ϕ) del singolo punto P_i attraverso le relazioni²:

² Si noti che la presenza delle σ_{xy} (varianza-covarianza) del generico punto P_i che

implica la conoscenza delle relazioni mutue tra σ_x e σ_y di cui alle variabili casuali a due dimensioni.

$$a^2 = \frac{1}{2}(\sigma_x^2 + \sigma_y^2) + \frac{1}{2}\sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4\sigma_{xy}^2}$$

$$b^2 = \frac{1}{2}(\sigma_x^2 + \sigma_y^2) + \frac{1}{2}\sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4\sigma_{xy}^2}$$

$$\tan 2\varphi = -\frac{2\sigma_{xy}}{\sigma_y^2 - \sigma_x^2}$$

2 - Metodo delle osservazioni dirette condizionate

Si consideri un insieme di punti uniti a due a due da segmenti di rette in modo da formare una serie di figure poligonali chiuse.

La posizione relativa dei punti può essere definita misurando i lati e/o gli angoli di tali figure; si può vedere facilmente che, se p è il numero dei punti, per definirne la posizione relativa è sufficiente un numero $s = 2p - 3$ di misure di lati e/o di angoli.

Infatti la misura di un solo lato definisce la posizione relativa di due punti, mentre occorrono altre due misure indipendenti per definire la posizione degli altri $p - 2$ vertici.

Il totale di misure atte a definire completamente il sistema geometrico è quindi

$$s = 2(p - 2) + 1 = 2p - 3$$

Ogni misura di lato o di angolo che si esegue in più rispetto alle minime s indipendenti è esuberante e dà luogo ad una “condizione” a cui una parte o la totalità delle misure eseguite deve soddisfare.

Nel caso altimetrico dell'esercizio n.3 per ogni linea chiusa di livellazione si può scrivere l'equazione:

$$\sum \Delta_i = 0$$

se sono stati misurati n dislivelli corrispondenti agli n punti di cui si vogliono determinare le quote.

Il numero minimo di misure è in tal caso $n - 1$, e per ogni maglia chiusa è possibile scrivere una equazione di condizione.

In generale considerato un sistema geometrico in cui è definibile un numero t qualsivoglia di grandezze e supposto che si siano misurate n grandezze X_1, X_2, \dots, X_n , dove il numero n è maggiore del numero s di grandezze, le cui

misure sono atte a definire il sistema, sussisteranno $r = n - s$ condizioni a cui le grandezze misurate X_1, X_2, \dots, X_n devono soddisfare, e che in generale possono esprimersi con le

$$f_i(X_1, X_2, \dots, X_n) = 0 \quad i = 1, 2, \dots, r \quad (2.1)$$

Ad esempio la relazione angolare per un triangolo è del tipo:

$$\alpha + \beta + \gamma - \pi = 0$$

ovvero per un poligono:

$$\sum \alpha_i - (n - 2)\pi = 0$$

In generale saranno n (numero delle grandezze misurate) $> r$ (numero delle relazioni ovvero EQUAZIONI DI CONDIZIONE).

Se supponiamo che le misure dirette delle grandezze X_1, X_2, \dots, X_n sono delle variabili casuali di tipo normale e se X'_1, X'_2, \dots, X'_n rappresentano dei valori estratti a caso dalle popolazioni di misure possibili relative ad X_1, X_2, \dots, X_n , per l'aleatorietà dei valori estratti si avrà:

$$f_i(X'_1, X'_2, \dots, X'_n) = \Delta'_i \quad i = 1, 2, \dots, r \quad (2.2)$$

dove al residuo aleatorio Δ'_i si può dare il nome di “errore di chiusura”.

Consideriamo ora una grandezza A non misurata del sistema ed osserviamo che essendo $n > s$ si possono configurare, fra le n grandezze misurate, gruppi diversi di s grandezze indipendenti che danno “consistenza” al sistema.

A è suscettibile di essere ricavata indirettamente da gruppi diversi di s misure; se le medie $X_{1m}, X_{2m}, \dots, X_{nm}$ non soddisfano le r relazioni (2.1) avverrà che la media di A sarà definita in modo diverso a seconda del gruppo usato; se si può ammettere che a seconda del gruppo usato cambi la varianza di A , non si può ammettere che cambi la media di A .

E' necessario quindi che le medie $X_{1m}, X_{2m}, \dots, X_{nm}$ grandezze misurate direttamente soddisfino le condizioni (2.1); ciò si può ottenere correggendo ovvero “compensando” le n medie

$X_{1m}, X_{2m}, \dots, X_{nm}$ di quantità v_i tali che:

$$f_i(X_{1m} + v_1, X_{2m} + v_2, \dots, X_{nm} + v_n) = 0 \quad i = 1, 2, \dots, r \quad (2.3)$$

Si può notare subito che il problema di calcolare le n correzioni v_i , è indeterminato dato che $r < n$; inoltre si può notare che fra le infinite maniere di determinare le correzioni v_i ci può essere quella derivata dal principio di massima verisimiglianza con la quale si ottengono i valori “più plausibili” delle correzioni.

Infatti in presenza delle r relazioni (2.1), il campione di $n > r$ valori indipendenti $X_{1m}, X_{2m}, \dots, X_{nm}$ è esuberante per la determinazione delle medie “compensate”

$$X_1^* = X_{1m} + v_1, \quad X_2^* = X_{2m} + v_2, \dots, \quad X_n^* = X_{nm} + v_n$$

Supponiamo che gli scarti v_i siano sufficientemente piccoli in modo da poterne trascurare i quadrati e le potenze superiori; sviluppando in serie le (2.3) si ha

$$f_i(X_{1m}, X_{2m}, \dots, X_{nm}) + \left(\frac{\partial f_i}{\partial X_1} \right)_m v_1 + \left(\frac{\partial f_i}{\partial X_2} \right)_m v_2 + \dots + \left(\frac{\partial f_i}{\partial X_n} \right)_m v_n = 0$$

dove le derivate parziali sono calcolate con i valori delle stime X_{im} poniamo:

$$\left(\frac{\partial f_1}{\partial X_1} \right)_m = a_1 \quad \left(\frac{\partial f_1}{\partial X_2} \right)_m = a_2 \quad \dots \quad \left(\frac{\partial f_1}{\partial X_n} \right)_m = a_n$$

per cui tenendo presenti le (2.2) otteniamo il sistema

$$\begin{aligned} a_1 v_1 + a_2 v_2 + \dots + a_n v_n + \Delta_1 &= \bar{0} \\ b_1 v_1 + b_2 v_2 + \dots + b_n v_n + \Delta_2 &= \bar{0} \\ &\dots \dots \dots \\ r_1 v_1 + r_2 v_2 + \dots + r_n v_n + \Delta_r &= \bar{0} \end{aligned} \quad (2.4)$$

che prende nome di sistema di “equazioni di condizione delle correzioni”.
La condizione che conduce alle stime più plausibili degli scarti v_i risulta:

$$\sum p_i v_i^2 = \min \quad \text{dove} \quad p_i = \frac{\sigma_0^2}{\sigma_i^2}$$

è il peso di ogni misura diretta, essendo σ_i^2 la relativa varianza. Dobbiamo trovare i valori v_i che soddisfano le condizioni

$$\begin{aligned} a_1 v_1 + a_2 v_2 + \dots + a_n v_n + \bar{\Delta}_1 &= \bar{0} \\ b_1 v_1 + b_2 v_2 + \dots + b_n v_n + \bar{\Delta}_2 &= \bar{0} \\ \dots &\dots \\ r_1 v_1 + r_2 v_2 + \dots + r_n v_n + \bar{\Delta}_r &= \bar{0} \end{aligned}$$

$$\sum p_i v_i^2 = \min$$

Si tratta di risolvere un problema di minimo condizionato ovvero di ricercare il minimo della funzione $\sum p_i v_i^2 = \min$ fra i punti in cui sono verificate anche le r condizioni.

Un metodo semplice di risolvere questo problema è quello dei moltiplicatori di Lagrange detto anche dei “correlativi”.

Si può dimostrare che il minimo condizionato della funzione $\sum p_i v_i^2$ corrisponde al minimo della funzione

$$G(v_1, v_2, \dots, v_n) = \sum p_i v_i^2 - K_1(a_1 v_1 + a_2 v_2 + \dots + a_n v_n + \Delta_1) - K_2(b_1 v_1 + b_2 v_2 + \dots + b_n v_n + \Delta_2) + \dots - K_r(r_1 v_1 + r_2 v_2 + \dots + r_n v_n + \Delta_r)$$

dove K_1, K_2, \dots, K_r sono dei coefficienti indeterminati.

Il minimo della funzione $G(v_1, v_2, \dots, v_n)$ si ha ovviamente quando

$$\frac{\partial G}{\partial v_1} = 0, \quad \frac{\partial G}{\partial v_2} = 0, \quad \dots, \quad \frac{\partial G}{\partial v_n} = 0 \quad (2.5)$$

ovvero quando

$$\begin{aligned}
 p_1 v_1 - a_1 K_1 - b_1 K_2 - \dots - r_1 K_r &= 0 \\
 p_2 v_2 - a_2 K_1 - b_2 K_2 - \dots - r_2 K_r &= 0 \\
 &\dots\dots\dots \\
 p_n v_n - a_n K_1 - b_n K_2 - \dots - r_n K_r &= 0
 \end{aligned}
 \tag{2.6}$$

Il complesso delle equazioni (2.4) e (2.6), note con il nome di equazioni correlanti, risolve il nostro problema; infatti la soluzione dipende dalle $n + r$ equazioni nelle $n + r$ incognite $v_1, v_2, \dots, v_n, K_1, K_2, \dots, K_r$.

Consideriamo ora il sistema (2.6), scritto nella seguente maniera

$$\begin{aligned}
 p_1 v_1 &= a_1 K_1 + b_1 K_2 + \dots + r_1 K_r \\
 p_2 v_2 &= a_2 K_1 + b_2 K_2 + \dots + r_2 K_r \\
 &\dots\dots\dots \\
 p_n v_n &= a_n K_1 + b_n K_2 + \dots + r_n K_r
 \end{aligned}$$

Moltiplicando ambo i membri della prima equazione per a_1/p_1 quello della seconda per a_2/p_2 quelli dell'ennesima per a_n/p_n e sommando, analogamente, moltiplicando la prima per b_1/p_1 , la seconda per b_2/p_2 la ennesima per b_n/p_n , e sommando e così via³ si ottiene un sistema lineare di r equazioni che contiene come incognite gli r correlativi chiamato *sistema normale ai correlativi*:

$\dots + \dots$

³ E' possibile scrivere il tutto in forma matriciale infatti la (2.4) può scriversi:

$$AV + \Delta = 0$$

inoltre indicata con K la matrice dei correlativi, con P quella dei pesi e con N la matrice normalizzata ($N = AP^{-1}A^T$) si ottiene:

$$K = -N^{-1}\Delta \quad \bar{V} = P^{-1}A^T K = -P^{-1}A^T N^{-1}\Delta$$

mentre la matrice varianza-covarianza è data da:

$$U = \sigma_0^2 (P^{-1} - P^{-1}A^T N^{-1}AP^{-1})$$

$$\begin{aligned} & \left[\frac{aa}{p} \right] K_1 + \left[\frac{ab}{p} \right] K_2 + \dots + \left[\frac{ar}{p} \right] K_r + \Delta_1 = 0 \\ & \left[\frac{aa}{p} \right] K_1 + \left[\frac{ab}{p} \right] K_2 + \dots + \left[\frac{ar}{p} \right] K_r + \Delta_1 = 0 \\ & \dots \dots \dots \\ & \left[\frac{aa}{p} \right] K_1 + \left[\frac{ab}{p} \right] K_2 + \dots + \left[\frac{ar}{p} \right] K_r + \Delta_1 = 0 \end{aligned}$$

Calcolati i correlativi K_1, K_2, \dots, K_r , si ricavano agevolmente dalle (2.6) le correzioni v_1, v_2, \dots, v_n .

Si può inoltre calcolare una stima della varianza dell'unità di peso

$$\overline{\sigma}_0^2 = \frac{1}{n-r} \sum_{i=1}^n p_i v_i^2$$

E' quindi possibile verificare che la compensazione non solo elimina le incongruenze delle misure eseguite ma conduce ad una omogeneizzazione della precisione.

Le misure compensate sono definite da distribuzioni che hanno una varianza minore delle corrispondenti misure dirette.

Vi sono dei problemi di compensazione di misure che si possono risolvere utilizzando sia il metodo delle osservazioni indirette, che il metodo delle osservazioni dirette condizionate.

Nei casi in cui il problema di compensazione trova soluzione in ambedue i metodi si può facilmente verificare che gli scarti delle equazioni alle misure coincidono con le correzioni fornite dalle equazioni correlanti.

CAPITOLO VII

STRUMENTI TOPOGRAFICI

Capitolo VII

STRUMENTI TOPOGRAFICI

1 - Il teodolite¹ (Misuratore di angoli)

Occorre anzitutto definire l'angolo azimutale (orizzontale) e l'angolo zenitale (verticale). Dati tre punti O, A e B della superficie terrestre, si definisce angolo azimutale fra A e B misurato in O, (fig. 1.1) la *sezione retta dell'angolo diedro* formato dal piano contenente la verticale per O e per il punto A ed il piano contenente la verticale per O ed il punto B.

Questo angolo coincide, a meno di correzioni in genere trascurabili (superiori ai minimi errori di misura), con l'angolo fra le due sezioni normali O_oA_o e O_oB_o dove con A_o , O_o e B_o si indicano le proiezioni dei punti A, O e B sulla superficie di riferimento (ellissoide).

Si definisce angolo zenitale l'angolo che la direzione OA forma con la verticale in O; il suo complemento è l'angolo di altezza.

Il **teodolite** è lo strumento che essenzialmente misura angoli azimutali e zenitali e che con determinati accessori, può anche misurare distanze ed azimut (magnetici o geografici).

¹ L'origine del nome è incerta; la denominazione si trova per la prima volta in un libro inglese del 1570: "...an instrument called theodolitus".

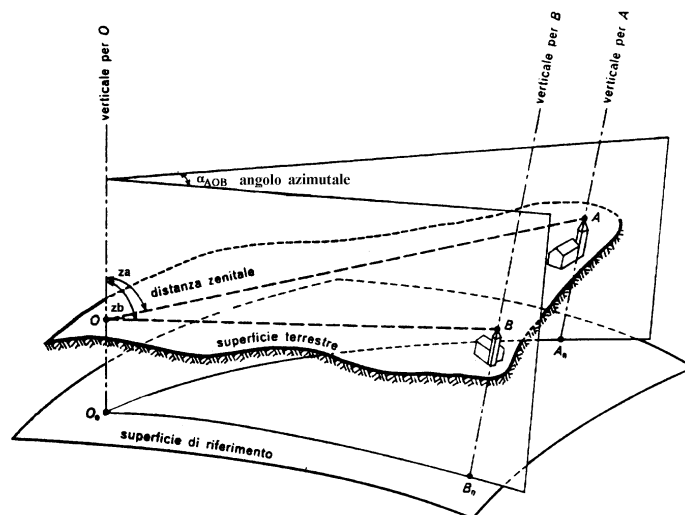


Fig. n. 1.1 - Definizione operativa dell'angolo azimutale e della distanza zenitale.

Il teodolite è costituito da una base “**b**”, solidale al terreno attraverso un treppiede, munita di tre “viti calanti” che permettono l’orientamento dell’asse primario “**a**”, quello intorno a cui ruota l’alidada. Sull’alidada è montato l’asse secondario “**s**”, intorno a cui ruota il cannocchiale, all’interno del quale è definito un terzo asse “**c**”, detto di collimazione (fig. 1.2).

Se lo strumento è rettificato l’asse secondario “**s**” è normale all’asse primario “**a**” e l’asse di collimazione “**c**” è normale all’asse “**s**”; quando lo strumento è messo in stazione “**a**” viene posto lungo la verticale e, di conseguenza, “**s**” diviene orizzontale. In queste condizioni l’asse di collimazione “**c**” ruotando attorno ad “**s**” descrive un piano verticale. I tre assi si incontrano in generale in un punto C, detto centro dello strumento.

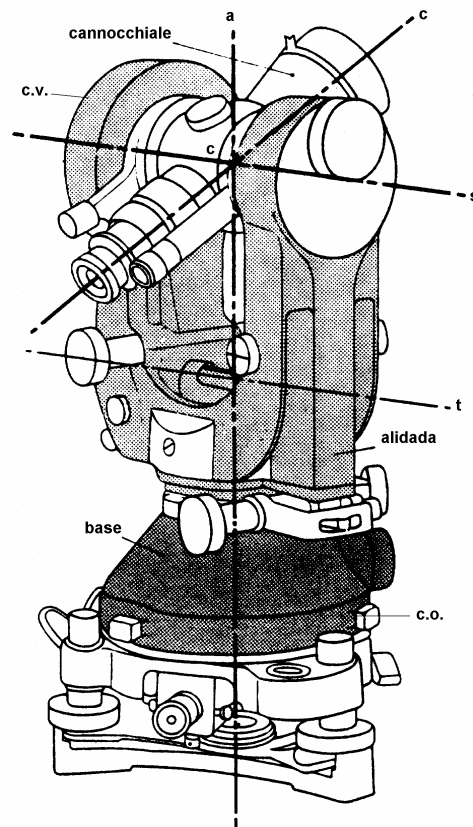


Fig. n. 1.2 - Assi principali del teodolite

Le letture degli angoli vengono effettuate mediante letture su cerchi graduati; quello per le misure azimutali è per costruzione normale all'asse "a" e risulta, quindi, orizzontale quando lo strumento è in stazione (C.O.); il cerchio per la misura delle distanze zenitali è, invece, normale ad "s" e risulta, perciò, in un piano verticale, se anche l'asse "a" è tale (C.V.).

Durante una semplice misura il C.O. resta solidale alla base, mentre il C.V. è solitamente fissato all'asse "s" e ruota insieme al cannocchiale (quando l'asse di collimazione descrive un piano verticale); gli indici di lettura sono in entrambi i casi fissati all'alidada.

Le graduazioni dei due cerchi sono sempre numerate in senso orario e possono essere sessagesimali o centesimali². A tutti i punti situati su uno stesso piano verticale in O, corrisponde la stessa posizione dell'alidada e quindi la stessa lettura sul C.O.

Ne deriva che, dirigendo l'asse di collimazione prima verso A e poi verso B ed eseguendo sul cerchio le misure L_A ed L_B , la differenza $L_B - L_A$ è uguale all'angolo azimutale AOB, misurata in senso orario da A verso B ($\alpha_{AOB} = L_B - L_A$).

² SISTEMI DI MISURA DEGLI ANGOLI

- 1) ANALITICO (S.I. sistema internazionale): adotta come unità di misura il **radiante**, che si denota con il simbolo 1^R e sottende un arco di circonferenza uguale al raggio. Con questo sistema gli angoli hanno l'inconveniente di essere espressi da numeri irrazionali (π , $\pi/2$, $\pi/4$, ...).
- 2) SESSAGESIMALE: la circonferenza è divisa in 360 parti, per ciascuna delle quali si assegna il valore di $1^\circ = \pi/180$. Il **grado** è diviso in 60 parti (**primi**) o minuti di angolo ($1' = 1^\circ/60$) ed il primo in 60 parti (**secondi**) o secondi di angolo ($1'' = 1'/60$).
- 3) CENTESIMALE: la circonferenza è divisa in 400 parti, (**gradi centesimali**) che si indicano con "g", l'ulteriore frazionamento prosegue con sottomultipli decimali.

E' possibile stabilire le seguenti relazioni tra i sistemi di misura descritti:

$$\frac{\alpha^R}{2\pi} = \frac{\alpha^\circ}{360} = \frac{\alpha^g}{400}$$

da cui si ricava che:

$$\alpha^\circ = \alpha^R / \text{arc}1^\circ, \alpha' = \alpha^R / \text{arc}1', \alpha'' = \alpha^R / \text{arc}1'' \quad \text{con } \text{arc}1' = 2.9 \times 10^{-4} \text{ e } \text{arc}1'' = 4.8 \times 10^{-6}$$

$$\alpha^g = \alpha^R / \text{arc}1^g, \alpha^C = \alpha^R / \text{arc}1^C, \alpha^{CC} = \alpha^R / \text{arc}1^{CC}$$

$$1^G = (9/10)1^\circ = 0^\circ,9 = 0^\circ54' = 3.240''$$

$$1^C = (9/1000)1^\circ = 0^\circ,009 = 32'',4$$

$$1^{CC} = (9/100000)1^\circ = 0'',324$$

Sul C.V. si leggono, invece, le rotazioni dell'asse "c" intorno ad "s". Supponiamo, allora, di conoscere la lettura "Z" in corrispondenza della direzione verticale dell'asse "c"; se ruotiamo l'asse "c" nel piano verticale contenente un punto A, fino a quando l'asse stesso non passi per A, la rotazione è uguale alla "distanza zenitale" ed è data dalla differenza tra la lettura S al C.V. e la lettura Z, ossia:

$$z = S - Z$$

Un angolo si esprime, in ogni caso, come differenza di due letture.

La misura degli angoli azimutali e zenitali raggiunge un valore di incertezza che dipende dallo strumento utilizzato, dall'ambiente in cui si opera, dai segnali che si collimano e da altri fattori che verranno esaminati in seguito.

La precisione strumentale di un teodolite varia entro limiti molto ampi, poiché lo scarto quadratico medio strumentale della misura di un angolo varia da circa 3×10^{-5} gon (cioè circa 0.1") a 2×10^{-2} gon (cioè circa 1').

Lo scarto quadratico medio ottenibile nelle misure effettive è praticamente uguale a quello strumentale quando si usano strumenti di precisione modesta e si collimano punti a breve distanza, non scende invece al di sotto di $1 \div 1.5 \times 10^{-4}$ gon (0.2"- 0.5") per gli angoli azimutali quando, pur usando strumenti di alta precisione, si collimano punti a notevole distanza, malgrado particolari cautele e malgrado si ripeta la misura un notevole numero di volte.

L'accuratezza ottenibile nella misura degli angoli zenitali risulta inferiore sia perché, per ragioni d'ingombro, il diametro del cerchio verticale è in genere più piccolo di quello azimutale, sia a causa dell'influenza negativa della rifrazione atmosferica. E' ovvio che tutte le parti che costituiscono il teodolite devono essere progettate e realizzate adeguatamente rispetto alla precisione globale dello strumento.

E' rimasta ancora in uso, ma è in via di eliminazione, la distinzione fra **teodoliti** e **tacheometri**; infatti mentre nei vecchi strumenti le caratteristiche erano diverse e le funzioni separate, oggi si dovrebbero chiamare con due nomi diversi strumenti che, come aspetto, peso ed operazioni non sono affatto diversi, essendo unica essenziale differenza quella della precisione.

E' più logico chiamare lo strumento teodolite ed indicare sempre lo scarto quadratico medio strumentale; ad es. teodolite da 1", teodolite da 10^{cc}, ecc.

2 - Il cannocchiale

Il cannocchiale astronomico fornito di reticolo fornisce il mezzo migliore per realizzare l'asse di collimazione "c".

Il cannocchiale astronomico è costituito da due lenti convergenti, la prima, **obbiettivo**, ha una distanza focale molto più grande della seconda, chiamata **oculare**; di un oggetto posto al di là del doppio della distanza focale l'obbiettivo dà un'immagine reale, capovolta e rimpicciolita; se la distanza tra l'oggetto, valutata sull'asse ottico dell'obbiettivo (congiungente i centri di curvatura delle superfici ottiche di separazione) ed il 1° punto principale dell'obbiettivo N₁ è **d**, l'immagine si forma alla distanza **q** dal secondo punto principale N₂, fornita dalla prima equazione fondamentale delle lenti:

$$(2.1) \quad \frac{1}{d} + \frac{1}{q} = \frac{1}{f}$$

Se l'oggetto, supposto puntiforme, ha una distanza H dall'asse ottico dell'obbiettivo, la distanza h dell'immagine dallo stesso asse è data da:

$$(2.2) \quad \frac{h}{H} = \frac{q}{d} = I$$

dove con “T” si indica l'**ingrandimento trasversale** della lente, ovvero il rapporto tra la dimensione dell'immagine e quella dell'oggetto, per una data posizione dell'oggetto stesso.

L'immagine data dall'obiettivo viene fatta cadere tra il fuoco ed il 1° punto principale dell'oculare: ne risulta un'immagine virtuale, diritta ed ingrandita (rispetto alla prima immagine), che può essere osservata ponendo l'occhio dietro l'oculare.

Il reticolo è costituito da un vetrino con sottili linee incise che, con diverse modalità, permettono di individuare un punto centrale chiamato “centro” del reticolo. (fig. 2.1)

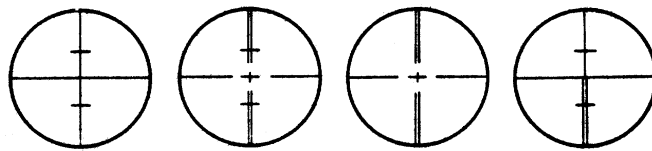


Fig. n. 2.1 - Esempi di reticolo

Il reticolo va posto nello stesso piano in cui si forma l'immagine reale attraverso l'obiettivo, il suo centro è sempre vicino all'asse ottico dell'obiettivo; un punto è collimato quando la sua immagine cade sul centro del reticolo (V. regole di collimazione).

L'asse “c” di collimazione è la congiungente il centro del reticolo con il secondo punto nodale dell'obiettivo (quello interno); infatti, la congiungente il centro del reticolo con il 2° punto nodale dell'obiettivo è parallela alla congiungente il 1° punto nodale con il punto collimato.

Quando un particolare oggetto è “collimato” l'immagine sul centro del reticolo, il punto nodale interno dell'obiettivo e l'oggetto collimato sono su una stessa retta (trascurando la piccolissima traslazione dovuta alla distanza fra i due punti nodali dell'obiettivo); l'asse di collimazione è definito, quindi, all'interno dello strumento.

Per collimare oggetti posti a distanze diverse è necessario poter variare la distanza q fra il reticolo e l'obbiettivo: si può quindi montare il reticolo sul tubo principale in posizione fissa e muovere l'obbiettivo o, come accade più spesso, rendere mobile il complesso reticolo-oculare. Comunque l'oculare deve essere mobile rispetto al reticolo, per permettere la visione distinta del reticolo e dell'immagine ad esso sovrapposta.

Le caratteristiche principali del cannocchiale sono: l'ingrandimento, il diametro dell'obbiettivo, il campo e la chiarezza.

L'**ingrandimento I** di un cannocchiale è definito dal rapporto tra la tangente dell'angolo " α " sotto cui l'occhio dell'osservatore vede l'immagine di un oggetto attraverso il cannocchiale e la tangente dell'angolo ε sotto cui l'occhio nudo dell'osservatore, nella stessa posizione, vede l'oggetto, e cioè:

$$(2.3) \quad I = \frac{\operatorname{tg} \alpha}{\operatorname{tg} \varepsilon}$$

L'ingrandimento dipende principalmente dalle distanze focali dell'obbiettivo e dell'oculare, ma anche dalla posizione di quest'ultimo rispetto al reticolo e dalla distanza dell'oggetto osservato; si definisce ingrandimento "**normale**", quello che si ha osservando un oggetto a distanza praticamente infinita, per cui $q = f_1$ con l'oculare accomodato per l'infinito, cioè posto in modo che il primo fuoco cada sul reticolo. Si trova così (fig. 2.2)

$$(2.4) \quad \operatorname{tg} \varepsilon = \frac{h}{f_1}, \quad \operatorname{tg} \alpha = \frac{h}{f_2}$$

da cui

$$(2.5) \quad I = \frac{\operatorname{tg} \alpha}{\operatorname{tg} \varepsilon} = \frac{f_1}{f_2}$$

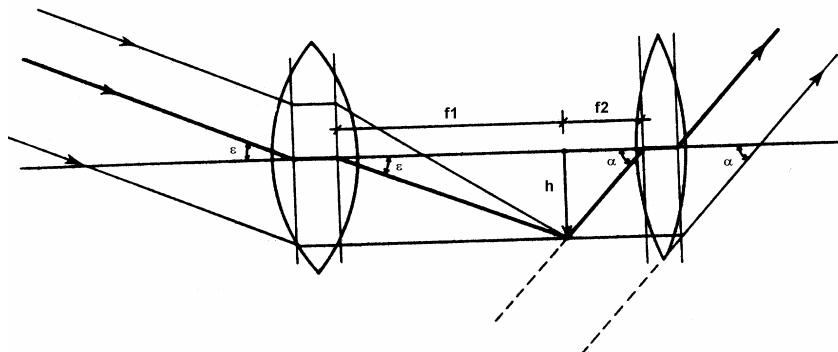


Fig. n. 2.2 - Ingrandimento *normale* del cannocchiale astronomico.

Negli strumenti topografici **I** varia da 10 a 60 volte circa.

Il diametro dell'obbiettivo interviene nel **potere di puntamento**, a sua volta legato al **potere risolutivo**.

E' noto che le due leggi (2.5) e (2.6) rappresentano uno schema di prima approssimazione dei fenomeni ottici relativi alle lenti, sufficientemente vicino alla realtà quando gli angoli che i raggi luminosi formano con l'asse ottico o con le normali alle superfici delle lenti sono abbastanza piccoli; questo è il caso dei cannocchiali.

Supponendo, quindi, che ad un punto oggetto corrisponda geometricamente un punto immagine, a causa della natura ondulatoria della luce, l'immagine sarà, in realtà, costituita da un dischetto di diffrazione il cui diametro r è pari a:

$$(2.6) \quad r = \frac{1.22\lambda}{D} f$$

con D diametro dell'obbiettivo, f distanza focale e λ lunghezza d'onda della luce, supposta monocromatica.

Per $D = 5$ cm, $f = 30$ cm e $\lambda = 5 \times 10^{-5} \mu$ (luce verde) si ha:

$$r = 3.7 \mu\text{m}$$

E' ovvio che, se sul piano del reticolo le immagini di due punti distano meno di r , i due punti non sono più distinguibili.

Il **potere risolutivo** γ è l'angolo con vertice nel centro dell'obbiettivo che sottende due punti distanti r nel piano focale; quindi il potere risolutivo:

$$(2.7) \quad \gamma \approx \frac{r}{f} = \frac{1.22\lambda}{D}$$

dipende solo dal diametro dell'obbiettivo (inversamente proporzionale).

Il **potere di puntamento** di un cannocchiale è numericamente uguale all'errore angolare che si commette quando, secondo il giudizio dell'osservatore, il punto collimato è sul filo del reticolo e può essere valutato circa 4 volte inferiore a γ . Si verifica facilmente che la precisione delle collimazioni effettuate con un cannocchiale provvisto di obbiettivo avente un diametro di 5 cm è di 1", con 6cm è di 0.5" e così via.

E' ovvio, quindi, che il diametro dell'obbiettivo debba essere commisurato alla precisione dello strumento e che siano armonizzati l'ingrandimento e il potere di puntamento.

La **chiarezza** è il rapporto tra la quantità di luce Q che, provenendo dall'oggetto, entra nell'occhio e l'area dell'immagine sulla retina; la chiarezza dell'oggetto visto attraverso il cannocchiale può essere definita in maniera uguale: la chiarezza di un cannocchiale è il rapporto tra la chiarezza dell'oggetto visto attraverso il cannocchiale e la chiarezza dell'oggetto visto ad occhio nudo; si può dimostrare che la chiarezza di un cannocchiale è inversamente proporzionale al quadrato dell'ingrandimento:

$$C \approx K/I^2$$

Praticamente i moderni strumenti topografici non utilizzano il cannocchiale astronomico, in quanto il cannocchiale a lunghezza costante presenta numerosi vantaggi:

a) è fissa la distanza fra obbiettivo e reticolo (una lente divergente, mobile entro il cannocchiale, provvede a far sì che le immagini reali degli oggetti collimati a varie distanze si formino sempre sul piano del reticolo);

b) ha maggiori ingrandimenti, a parità di lunghezza del cannocchiale;

c) il suo asse di collimazione, durante la messa a fuoco, è univocamente determinato (al contrario di quello del cannocchiale astronomico);

d) è praticamente stagno alla polvere e all'umidità, ecc..

Si indicano di seguito le regole per la collimazione, da utilizzare per un corretto uso del teodolite:

1. Adattamento alla visione distinta del reticolo. Si ottiene rivolgendo l'obbiettivo verso una superficie chiara (cielo) e spostando l'oculare, fino a vedere nitidamente il reticolo.
2. Collimazione approssimata con il mirino esterno (cannocchiale cercatore).
3. Bloccaggio delle viti di fermo dell'alidada e del cannocchiale.
4. Adattamento alla distanza (spostando la vite di messa a fuoco) e controllo della coincidenza dei piani su cui giacciono le immagini del reticolo e dell'oggetto collimato (assenza di parallasse).
5. Collimazione esatta con le viti micrometriche.

Se le operazioni sono state eseguite correttamente, spostando l'occhio davanti all'oculare non devono avvenire spostamenti relativi apprezzabili tra i tratti del reticolo e l'immagine dell'oggetto collimato.

3 - Mezzi di lettura ai cerchi

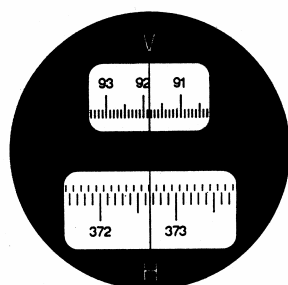
I cerchi graduati dei teodoliti hanno oggi dimensioni assai ridotte per poter realizzare strumenti di piccolo ingombro.

Lo spessore dei tratti del cerchio, riportanti una finissima graduazione, deve essere molto piccolo e, quindi, l'osservazione dei cerchi non può che essere effettuata attraverso un microscopio, il quale serve anche come organo di lettura (indice).

Una “**lettura**” viene eseguita leggendo sempre i gradi e le parti di grado incise sul cerchio e valutando le frazioni di intervallo con mezzi diversi; distinguiamo anzitutto i mezzi in due gruppi: il gruppo in cui la valutazione delle frazioni è effettuata con una stima e quello in cui si ricorre ad un sistema micrometrico.

Esempi del primo gruppo sono:

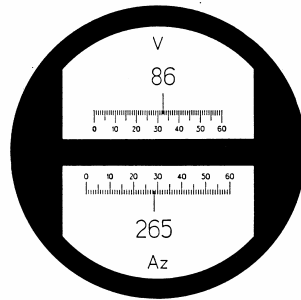
- a) la semplice stima diretta sulla graduazione dei cerchi;
microscopio a stima (fig. 3.1);



Lettura centesimale $V = 91.85^g$ $H = 372.66^g$

Fig. n. 3.1 - Lettura a stima su graduazione centesimale

- b) Il microscopio a scala (fig. 3.2). Sul reticolo del microscopio di lettura è incisa una scala, divisa in un certo numero di parti, la cui lunghezza è uguale alla lunghezza di un intervallo della graduazione del cerchio; si leggono così direttamente le parti della scala comprese tra l'origine ed il tratto di graduazione precedente e si può stimare una porzione della parte della scala.



Lettura sessagesimale $V = 86^{\circ}32',5$ $H = 265^{\circ}28',5$

Fig. n. 3.2 - Microscopio a scala.

I mezzi di lettura di questo primo gruppo non consentono una grande precisione, perchè al massimo si giunge ad apprezzare 0.5^C (15").

3.1 - Lettura con micrometro

Molto usata nei microscopi micrometrici è la lastra piano-parallela (fig. 3.3); come si vede applicando le leggi sulla rifrazione, ad un raggio che incide sulla lastra, in modo da formare un angolo di incidenza i “piccolo” con la normale, corrisponde un raggio emergente parallelo e spostato proporzionalmente all’angolo i della quantità “d”:

$$(3.1) \quad d = s \frac{n-1}{n} i$$

con n = indice di rifrazione della “lastra”; s = spessore della “lastra”.

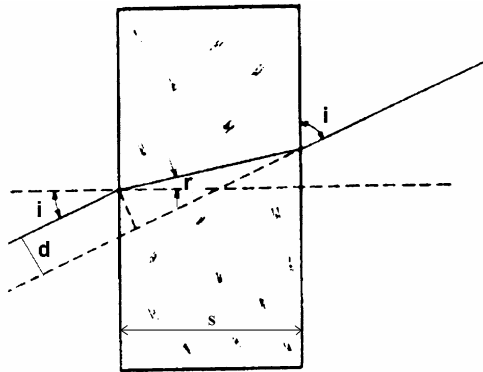


Fig. n. 3.3 - Lastra piano parallela

Si pone, allora, nel microscopio micrometrico (fig. 3.4), una lastra piano-parallela in posizione opportuna lungo il percorso dei raggi di lettura del cerchio, girevole intorno ad un asse ortogonale alla direzione dei raggi, e si fa in modo che nel campo dell'oculare del microscopio possa essere effettuata la lettura dell'angolo di cui la lastra piano-parallela è ruotata.

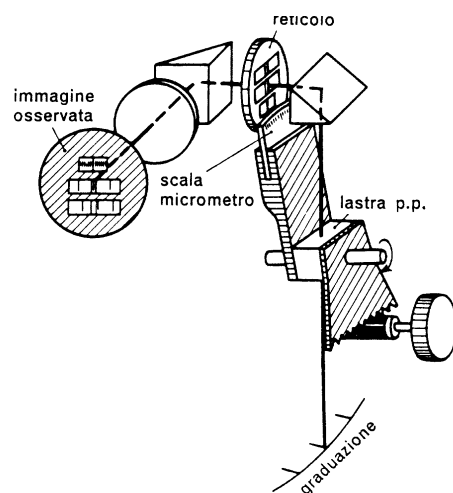
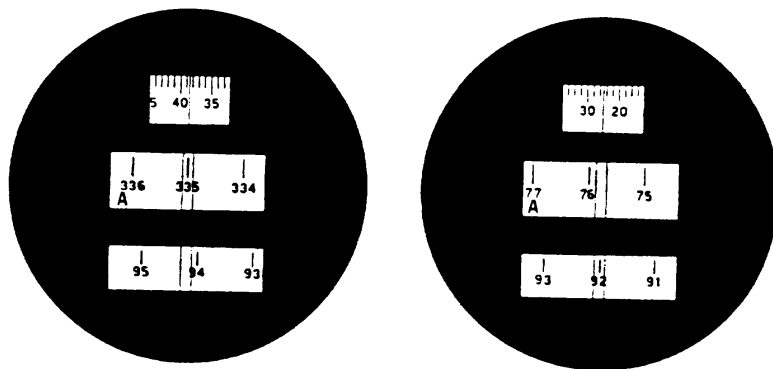


Fig. n. 3.4 - Micrometro a lastra piano parallela

Sul reticolo è posto un semplice doppio tratto; ruotando la lastra, l'immagine del cerchio si sposta fin tanto che un tratto della graduazione viene ad essere centrato dal "doppio tratto" del reticolo; l'angolo di cui la lastra ha ruotato rispetto alla posizione per cui lo spostamento è nullo, è proporzionale alla frazione di intervallo che si vuole valutare. Le letture possono essere fatte o sul bottone di comando delle rotazioni della lastra, oppure su un piccolo cerchio graduato che viene fatto ruotare insieme alla lastra, la cui immagine appare nell'oculare (solitamente entrambe le letture permettono di leggere direttamente le frazioni di intervallo). Con questo sistema possono essere effettuate letture con la precisione fino a 0.1" (fig. 3.5).



lettura $335^{\circ} 38' 30''$

lettura $92^{\circ} 24' 50''$

Fig. n. 3.5 - Esempi di lettura con micrometro a lastra p.p.

Nei moderni teodoliti, come in quelli elettronici, la funzione dell'operatore non è più quella di una stima, ma di eseguire un puntamento e realizzare la coincidenza.

3.2 - Lettura digitale

Una diversa concezione si ha con la lettura a contatore digitale: alla periferia dei cerchi sono riportati dei tratti opachi separati da intervalli trasparenti dello stesso spessore; il sistema di lettura è formato da una sorgente luminosa e da una sottostante cellula fotoelettrica; al ruotare del sistema di lettura si produce nel circuito della cellula una variazione di corrente in corrispondenza di ogni passaggio luce-buio; tali variazioni di corrente definiscono pertanto un certo numero di impulsi che vengono sommati e registrati in un contatore digitale di impulsi.

La lettura automatica digitale è legata al valore della precisione (il numero dei tratti alla periferia del cerchio non può essere aumentato a piacere), che raramente può scendere sotto il secondo centesimale (almeno per ora). (fig. 3.6)

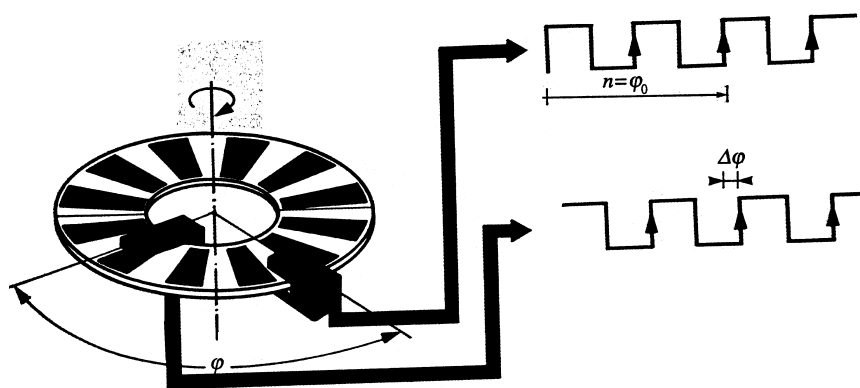


Fig. n. 3.6 - Lettore ottico digitale

4 - Messa in stazione di un teodolite

La messa in stazione di uno strumento in corrispondenza di un punto P materializzato con un “picchetto” (o con un segno che lo identifichi) consiste nell’importante operazione preliminare di: (1) rendere verticale, e (2) coincidente con la verticale passante per P (individuata dal filo a piombo), l’asse principale “a” dell’alidada (blocco ruotante dello strumento).

Sistemi di centramento: il punto di stazione, come pure i segnali da collimare, devono essere definiti con una precisione adeguata al tipo di misura da eseguire:

- con i sistemi di centramento “forzato”, che realizzano la intercambiabilità tra strumento e segnale, si ottengono precisioni di centramento inferiori al mm; (fig. 4.1)

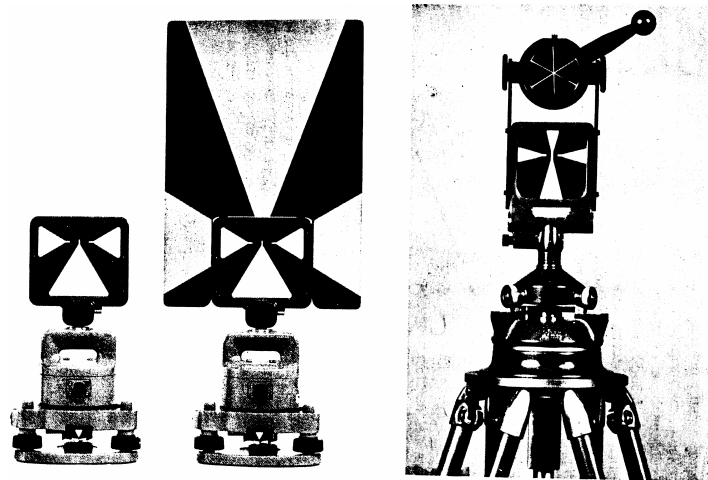


Fig. n. 4.1 - Alcuni tipi di segnali o mire

- con il “piombino ottico” si raggiungono precisioni di centramento di qualche mm (esistono anche particolari treppiedi con un bastone munito di livella sferica e testa a snodo);
- con il “filo a piombo” (fig. 4.2) sospeso alla vite di attacco dello strumento al treppiede, in assenza di vento, si raggiunge una precisione di circa 1cm.

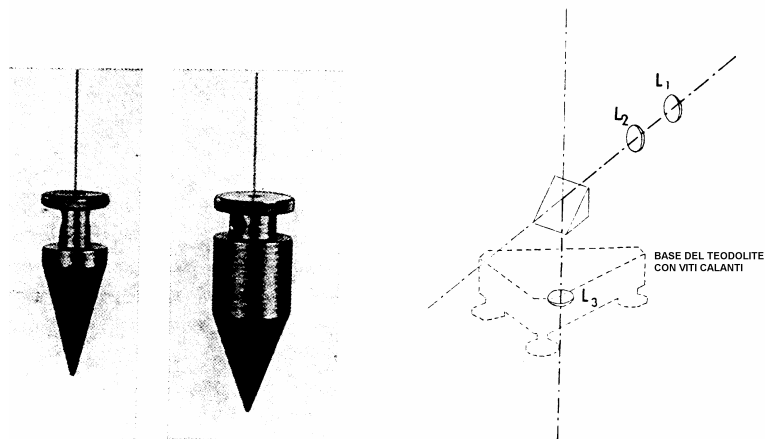


Fig. n. 4.2 - Tipi di piombini a gravità-schema di piombino ottico

Si badi che la precisione di centramento è importante sia per le misure di controllo delle deformazioni (ripetute più volte nel tempo) sia per misure di elevata precisione.

5 - Livelle

La livella sferica consiste in una fiala di vetro avente la superficie interna a forma di calotta sferica; il piano tangente alla sfera nel centro della bolla è orizzontale ed è, a “bolla” centrata, parallelo al piano di appoggio a condizione che la livella sia rettificata.

La sua sensibilità è piuttosto bassa e può raggiungere un'approssimazione di 2'-3'.

E' usata negli strumenti topografici e negli accessori (basette, stadie, ecc.) per permettere orientamenti rapidi, ma poco precisi di assi.

Molto più precisa è, invece, la livella torica (fig. 5.1), fiala di vetro riempita parzialmente di un liquido poco viscoso (alcool,

etere o benzina) avente la superficie interna di “forma toroidale”³; poichè il liquido la riempie solo parzialmente, nella parte più alta del “toroide” si forma una “bolla” costituita, prevalentemente, dai vapori del liquido; la superficie del liquido che delimita la bolla, a parte i menischi, è orizzontale; ne deriva che la tangente centrale alla curva direttrice del “toroide” nel punto di mezzo della bolla è sempre orizzontale.

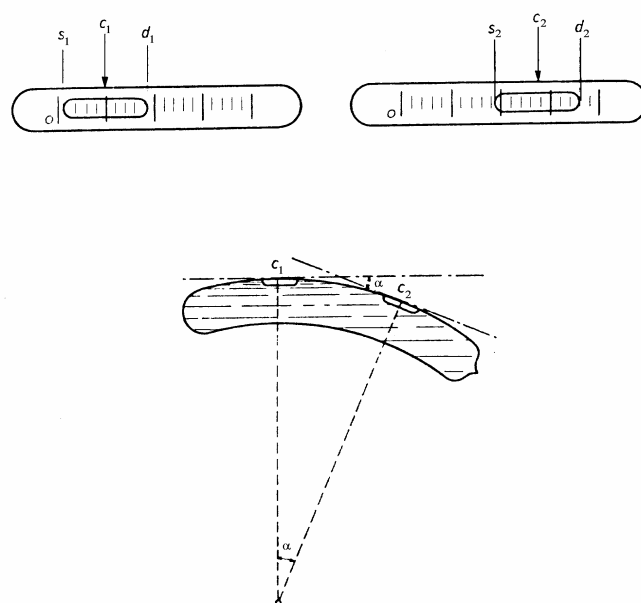


Fig. n. 5.1 - Livella torica

Sulla fiala e' incisa una graduazione i cui tratti distano di uno o due millimetri. Centrare la bolla significa ruotare la livella finchè la mezzeria della bolla si porta sulla mezzeria della graduazione (in queste condizioni la tangente al “toroide” nel punto di mezzo della graduazione, o tangente centrale, è orizzontale); il punto di mezzo

³ Superficie Toroidale è quella generata dalla rotazione di una circonferenza attorno ad un asse non passante per essa.

della bolla non è individuabile con precisione, mentre lo sono, più facilmente, i menischi laterali, che dovranno, ovviamente, essere simmetrici rispetto al centro della graduazione, cioè equidistanti dai tratti simmetrici della stessa. Quindi, la bolla si può ritenere “centrata” se i suoi estremi sono equidistanti dai tratti simmetrici della graduazione.

Esiste una proporzionalità tra spostamenti della bolla e rotazioni della livella; la sensibilità della livella (fig. 5.2) e' definita come l'angolo di cui deve ruotare la tangente centrale perchè la bolla si sposti di 1mm; è evidente che quanto più alto e' il valore del raggio R della sezione mediana del toroide, tanto più elevata è la sensibilità (cioè più basso il valore che la rappresenta), cosicchè:

$$\nu'' = \frac{P}{R} = \frac{(mm)}{(mm)} \times \frac{1}{arc1''} \quad (\text{sensibilità in sec. sessagesimali})$$

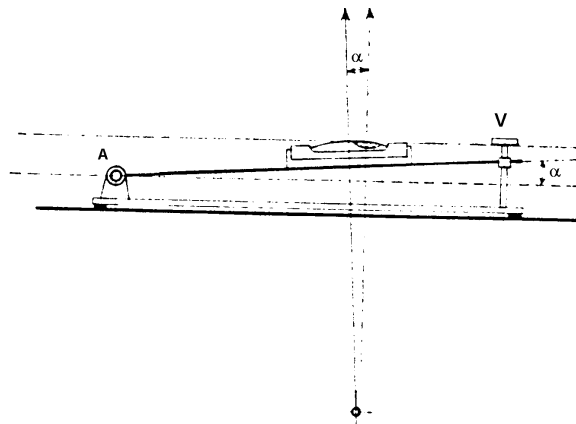


Fig. n. 5.2 - Comparatore di livelle

dove P è l'intervallo della graduazione.

La sensibilità delle comuni livelle toriche è di circa 10-20".

Conosciuta la sensibilità e/o il “valore angolare di una parte” si può sempre valutare la rotazione che ha subito la tangente centrale della livella, quando la bolla si è spostata da un punto ad un altro.

Le operazioni che vengono comunemente eseguite con la livella torica sono:

a) orizzontalità di una retta (fig. 5.3)

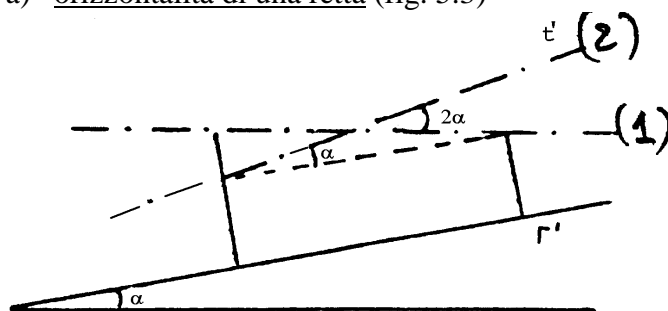


Fig. n. 5.3 - Livella per assi orizzontali

In questo caso la livella è vincolata, in qualche maniera, all'asse in modo da rendere orizzontale, mediante due (o una) viti di rettifica che consentono di far variare l'assetto della fiala rispetto alla retta di appoggio.

La livella si dice “rettificata” quando la tangente centrale è parallela alla retta di appoggio. E' facile quindi con una livella rettificata rendere orizzontale un asse: basta appoggiarla a questo e far ruotare quest'ultimo finchè la bolla si centra.

Se la livella non è rettificata, la tangente centrale forma un angolo α con l'asse su cui è appoggiata. Quindi, supposto di avere centrato la bolla, l'asse di appoggio (r') è inclinato dell'angolo α sull'orizzontale (pos.1). Invertendo la livella sugli appoggi, la tangente centrale risulterà inclinata dell'angolo 2α (pos.2) ed evidentemente la bolla subirà uno spostamento corrispondente; ovvero lo spostamento della bolla di una livella, rettificata o meno, appoggiata ad un asse e successivamente invertita di π , è proporzionale al doppio dell'angolo che l'asse forma con l'orizzontale.

Bisognerà quindi ricentrare la bolla per metà, ruotando l'asse r' (che, corrispondentemente all'angolo α , diventa orizzontale) e per

l'altra metà manovrando le viti di rettifica, per rendere la tangente centrale parallela alla retta di appoggio (ottenendo così anche la rettifica della livella);

b) orizzontalità di un piano

E' sufficiente rendere orizzontali due rette del piano fra loro ortogonali con il metodo descritto al punto precedente (condizione teorica che si verifica solo per piccole porzioni di superfici appartenenti a materiali poco deformabili); questa tecnica non viene utilizzata in topografia.

c) verticalità di un asse (fig. 5.4 e 5.5)

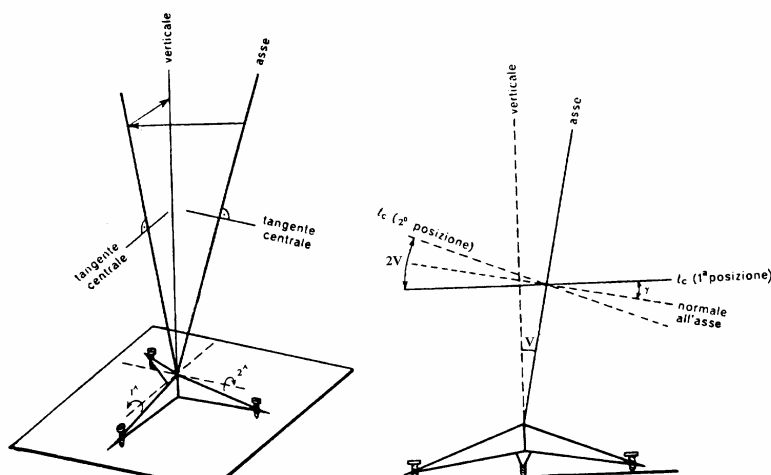


Fig. n. 5.4-5.5 - Operazioni per rendere verticale un'asse con una livella

E' il caso che più interessa per la messa in stazione di uno strumento.

In questo caso la livella deve essere montata su un supporto (base dell'alidada), che si può ritenere ortogonale all'asse che si vuole rendere verticale, ed è rettificata quando la sua tangente centrale è normale all'asse considerato (intorno a cui la livella è libera di ruotare).

Il teodolite è munito di una base triangolare con tre viti calanti agli estremi.

- 1) Si dispone la livella nella direzione di una vite calante e si centra la bolla manovrando la vite stessa.
- 2) Ruotando la livella di 180° , se la livella è rettificata, la bolla rimane centrata e si passa all'operazione successiva.

Altrimenti lo spostamento della bolla è nuovamente proporzionale al doppio dell'angolo di rettifica v , e quindi si riconduce la bolla in centro per metà, ruotando l'asse con la stessa vite di base (l'asse si dispone nel piano verticale normale al piano di lavoro della livella), per l'altra metà con le viti di rettifica della livella (la tangente centrale si dispone normale all'asse). (fig. 5.4-5.5).

- 3) La terza operazione, a livella già rettificata, si esegue ruotando la livella di $\pi/2$ e centrando sempre la bolla con movimenti "simultanei e contrapposti" delle altre due viti. Infine, come verifica, si può ruotare la livella in una direzione generica: se le operazioni sono state condotte con cautela, essa rimarrà centrata.

In definitiva, supponendo la livella rettificata, si eseguono le operazioni 1) e 3) con la verifica ed eventuali ripetizioni più accurate.

Solo quando, nonostante ripetuti tentativi, non si consegue il centramento della bolla in una generica posizione della livella, si dovrà controllare lo stato di rettifica della stessa, con operazioni da eseguire in ambiente controllato (base dello strumento molto stabile, non sul treppiede ma su un pilastro o muratura, temperatura costante, assenza di cause perturbanti come sole, vento, ecc...).

6 - Errori sistematici nella misura degli angoli

Esaminiamo ora le condizioni di rettifica di un teodolite.

- 1) Asse “a” verticale e passante per il punto di stazione (messa in stazione);
- 2) Asse “s” perpendicolare all'asse “a”;
- 3) Asse “c” perpendicolare all'asse “s”;
- 4) C.O. perpendicolare all'asse “a”;
- 5) Asse “a” passante per il centro del C.O.;
- 6) C.V. perpendicolare all'asse “s”;
- 7) Asse “s” passante per il centro C.V.;
- 8) Asse “a” incidente asse “s”;
- 9) Asse “s” incidente asse “c”; assi “a”, “s”, e “c” incidenti in un punto C;
- 10) Graduazione esatta nei cerchi graduati.

Da un punto di vista generale, le misure di angoli saranno affette da errori sistematici, dovuti a cause preventivamente determinabili (e quindi in gran parte eliminabili), e da errori accidentali, regolati dalla casualità e dalle leggi della probabilità.

L'unico modo per prendere in considerazione questi ultimi errori è aumentare il numero di misure tenendo presente che, se σ_s è la precisione dello strumento, sarà:

$$(6.1) \quad \sigma_m = \sigma_s / \sqrt{n}$$

e potremo considerare attendibile la misura espressa dal valore più probabile:

$$(6.2) \quad \alpha_m \pm \sigma_m$$

Gli errori dovuti alla non perfetta rettifica dello strumento misuratore rientrano tra gli errori sistematici. Poichè di questi 10 errori conosciamo le cause, possiamo studiarli e, se possibile, tentare di eliminarli.

E' necessario anzitutto osservare che la condizione (1) di verticalità dell'asse “a” e le condizioni (2) di perpendicolarità tra “a” ed “s” e (3) di perpendicolarità “s” e “c”, non possono essere mai realizzate con esattezza.

Si potrebbe pensare che le tre condizioni si possano considerare realizzate quando tali errori siano dell'ordine di grandezza della precisione strumentale del teodolite, ma occorre tener presente che: 1) l'influenza di tali errori varia con le condizioni operative; 2) mentre negli strumenti meno precisi le rettifiche possono essere spinte sino al punto da far ritenere lo strumento privo di errori, ciò non si verifica mai negli strumenti di alta precisione.

Immaginiamo quindi che nello strumento siano presenti i tre errori: ϵ_v (di verticalità), ϵ_i (di inclinazione) e ϵ_c (di collimazione) dando per scontato che sia stata già effettuata dalla casa costruttrice la rettifica migliore (parleremo cioè di errori “residui” di rettifica) e che conviene al più effettuare, di tanto in tanto, delle operazioni che ci consentano di verificare se la rettifica dello strumento sia rimasta inalterata.

Vediamo come tali errori influenzano le letture al C.O. (angoli azimutali) quando venga collimato un punto C in una direzione qualunque, che sarà definita da un angolo azimutale A (riferito al piano di inclinazione dell'asse **a** rispetto alla verticale) e dalla distanza zenitale o dal complementare (angolo di altezza ϕ).

Teniamo anche presente che la collimazione al punto C può essere eseguita in due situazioni diverse dello strumento, una con il C.V. **a sinistra** del cannocchiale (C.S.) ed una con il C.V. **a destra** (C.D.). Per studiare l'influenza degli errori residui di rettifica è comodo riferirsi ad una sfera delle direzioni di raggio unitario, operando con le formule della trigonometria sferica. Tale studio è molto complesso se si vuole calcolare l'influenza globale dei tre errori, qualunque siano i loro valori.

Se, però, supponiamo in prima approssimazione che ϵ_v , ϵ_i , ed ϵ_c siano abbastanza piccoli da poterne trascurarne le potenze superiori, l'influenza di ciascun errore potrà essere determinata sperimentalmente.

Se infatti L è la lettura che si farebbe nell'ipotesi di strumento perfettamente rettificato (non determinabile quindi) ed L' la lettura reale, possiamo scrivere che l'errore nella lettura è:

$\delta = L' - L = f(v, i, c, A, \varphi)$; sviluppando in serie di MacLaurin, trascurando le potenze superiori e considerando che $f(0) = 0$ si ottiene:

$$(6.3) \quad L' - L = \left(\frac{\partial f}{\partial v} \right) \varepsilon_v + \left(\frac{\partial f}{\partial i} \right) \varepsilon_i + \left(\frac{\partial f}{\partial c} \right) \varepsilon_c + \dots$$

ove le derivate sono funzioni di A e di φ e sono calcolate per $\varepsilon_v = \varepsilon_i = \varepsilon_c = 0$.

Come si vede, gli effetti dei vari errori sulle letture possono essere studiati separatamente (legge di indipendenza dei piccoli errori).

Consideriamo anzitutto l'influenza dell'errore di verticalità ε_v . E' evidente che collimando punti nel piano verticale BOV non si ha nessun errore nella lettura; viceversa, (fig. 6.1) collimando un generico punto P, l'alidada dovrebbe ruotare dell'angolo A , mentre ruota effettivamente dell'angolo A' .

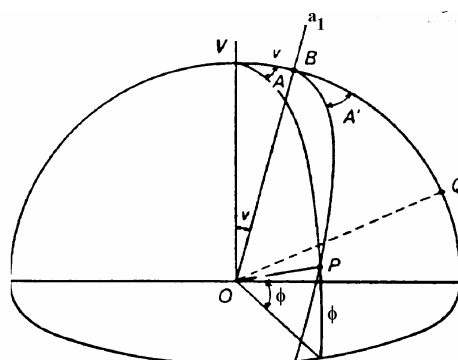


Fig. n. 6.1 - Influenza dell'errore di verticalità

Si dimostra che:

$$(6.4) \quad \varepsilon_v = A' - A = v (\operatorname{tg} \varphi - \operatorname{tg} A)$$

Con considerazioni analoghe si valutano gli errori $\varepsilon_i = \pm i \cdot \operatorname{tg} \varphi$ ed $\varepsilon_c = \pm c \cdot \sec \varphi$.

Se supponiamo di ruotare l'alidada di π ed effettuare le letture in questa seconda posizione (quindi con C.D.) noteremo che (fig. 6.1) mentre gli errori ε_i ed ε_c si invertono di segno, permane solo l'errore ε_v . Possiamo quindi esplicitare la:

$$(6.5) \quad L - L' = v \cdot \operatorname{sen} A \cdot \operatorname{tg} \varphi \pm i \cdot \operatorname{tg} \varphi \pm c \cdot \sec \varphi$$

dove i doppi segni si riferiscono agli errori nelle due posizioni coniugate C.S. e C.D. (il passaggio non comporta invece alcuna variazione dell'asse a e quindi di ε_v).

6.1 - Misura degli angoli azimutali

Già sappiamo che l'angolo orizzontale AOB si definisce come la sezione retta dell'angolo diedro avente per costola l'asse verticale dello strumento passante per O e per facce i piani (verticali) per i punti A e B. Sappiamo anche che un angolo si misura come differenza tra due letture (collimando A e poi B, $\alpha = L_B - L_A$).

In pratica si preferisce sempre effettuare almeno due misure, nelle posizioni C.S. e C.D., ottenendo così due vantaggi: 1) il possesso di due valori dell'angolo e, quindi, una diminuita possibilità di errori grossolani; 2) l'eliminazione degli errori ε_i ed ε_c che influiscono di quantità contrarie sulle letture C.D. e C.S.

Se ne deduce la “regola di Bessel”:

“La misura di un angolo azimutale, con la media dei valori ottenuti nelle due posizioni coniugate dello strumento, non è influenzata dalla presenza di errori di collimazione e di inclinazione dell'asse di rotazione del cannocchiale”.

La misura dell'angolo più attendibile sarà quindi:

$$\alpha = \hat{AOB} = \frac{(L_B - L_A)}{2} + \frac{(L'_B - L'_A)}{2} - v(\operatorname{sen} A_B \cdot \operatorname{tg} \varphi_B - \operatorname{sen} A_A \cdot \operatorname{tg} \varphi_A)$$

avendo indicato con L e L' le letture sul C.O. nelle due posizioni coniugate dello strumento (C.D. e C.S.).

Si ricordi che “ φ ” è l'angolo di elevazione del cannocchiale rispetto al piano orizzontale, mentre A è l'angolo tra la posizione iniziale di riferimento, coincidente con il piano di inclinazione della verticale (evidentemente non determinabile) ed il piano verticale contenente l'asse di collimazione.

La presenza dell'errore residuo di verticalità può essere resa del tutto trascurabile sia per piccole inclinazioni del cannocchiale ($\varphi=0 \rightarrow \text{tg}\varphi=0$), sia con l'utilizzo di una livella molto sensibile a “coincidenza” per cui “ v ” diventa molto piccolo (da utilizzare per le misure di controllo delle deformazioni delle strutture).

Continuiamo ora a valutare l'influenza degli altri errori dovuti alla non perfetta rettifica dello strumento:

- L'errore calcolato attraverso la (6.4) raggiunge al massimo ordini di grandezza di $\cong 0''.09$, ed è ininfluente (minore del limite di precisione conseguibile nella misura degli angoli, che è di $\cong 0''.1$).

- L'errore dovuto alla eccentricità dell'asse “ a ” rispetto al cerchio orizzontale, può essere eliminato facendo due “letture agli indici opposti” e, con una certa approssimazione, anche con due letture coniugate (fig. 6.2):

$$(6.6) \quad \alpha' = \alpha - \varepsilon; \quad \text{sen}\varepsilon \approx \varepsilon = e/r \times \text{sen}\alpha$$

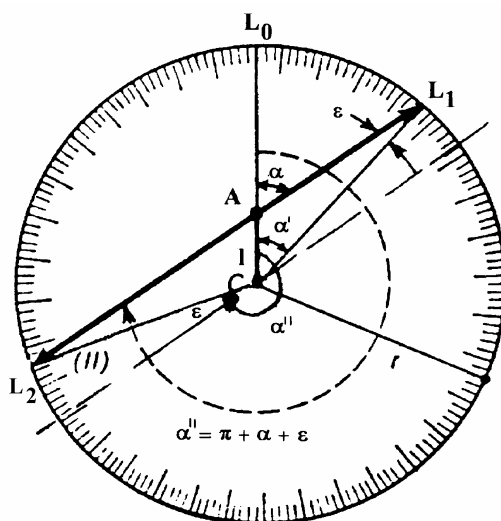


Fig. n. 6.2 - Errore di lettura dovuto all'eccentricità tra centro del cerchio graduato e asse dell'alidada.

Per esempio posto $\sin \alpha = 1$; $e = 0,01 \text{ mm}$; $r = 50 \text{ mm}$ risulta:

$$\varepsilon' = \frac{0.01 \text{ mm}}{50 \text{ mm}} \times 206.265 \approx 40''$$

(errore notevole anche per teodoliti di bassa precisione).

Se, però, consideriamo anche un'altra lettura con un indice diametralmente opposto al primo (microscopio a coincidenza):

$$(6.7) \quad \alpha'' = \alpha' + \pi + 2\varepsilon = \alpha + \pi + \varepsilon;$$

sommando le (6.6) e (6.7) si ha:

$$(6.8) \quad 2\alpha = \alpha' + \alpha'' - \pi \rightarrow \alpha = \frac{\alpha' + \alpha'' - \pi}{2}$$

- Gli errori (6) e (7), riguardanti letture al C.V., non vengono presi in considerazione a causa di quelli (di ben maggiore entità) dovuti al fenomeno della rifrazione nelle misure degli angoli zenitali.

- Gli errori (8) e (9), al pari degli errori (2) e (3), vengono eliminati con le due letture coniugate.

- Per cercare di eliminare la maggior parte dell'errore (10) di graduazione del cerchio, consideriamo anzitutto che, qualunque sia la disposizione dei tratti (più o meno precisa), la loro somma deve comunque valere 2π (360 o 400 a seconda del sistema di misura).

Nei teodoliti in cui è applicabile il meccanismo di “reiterazione” il C.O. non è fissato alla base, ma viene ad esso collegato con un sistema a frizione, e può quindi essere ruotato.

Stabilita la precisione che si vuole ottenere, e quindi l'errore della media σ_m , si valuta anzitutto il numero $n = \frac{\sigma_s^2}{\sigma_m^2}$ di reiterazioni

(ricavandolo dal valore dell'incertezza dello strumento e dal valore dell'errore della media prefissato); si collima, quindi, il primo

punto e si esegue una prima misura dell'angolo, applicando naturalmente la regola di Bessel (4 letture). Si ricollima quindi il primo punto e si sposta il C.O. di $\gamma = \frac{\pi}{n}$ mediante la “**vite di reiterazione**”⁴; si esegue la seconda misura e si sposta nuovamente il C.O. di un angolo γ , e così via.

La media aritmetica delle misure fornirà il valore cercato e riterremo:

$$(6.9) \quad \alpha = \alpha_m \pm \sigma_m$$

Alcuni teodoliti usano, invece, il meccanismo di ripetizione: in tal caso il C.O. può essere fissato o alla base o all'alidada; per la misura dell'angolo azimutale tra A_1 ed A_2 si procede così: fissato il cerchio alla base si esegue la lettura L_1 ; si collima A_2 e si ottiene la lettura $L_2 = L_1 + \alpha$; la differenza $L_2 - L_1 = \alpha$ dà un valore approssimato dell'angolo.

Si fissa il cerchio all'alidada e si collima di nuovo A_1 ruotando il complesso C.O. alidada rispetto alla base; la lettura è ancora L_2 ; unito il cerchio alla base, si collima A_2 , si legge l'indice ($L_1 + 2\alpha$); fissato nuovamente il cerchio dell'alidada ci si riporta nuovamente su A_1 e così via per il numero di volte previsto.

Quando si collima per l'ultima volta A_2 si legge di nuovo il cerchio, ottenendo:

$$(6.10) \quad L_{n+1} = L_1 + n\alpha$$

Il valore medio dell'angolo misurato è dato in tal caso da:

$$(6.11) \quad L_{n+1} + k \times 400^g - L_1 = \alpha_m$$

K si determina facilmente, conoscendo già il valore approssimato dell'angolo α (le stesse misure devono farsi al C.S. ed al C.D.).

Tale metodo è stato per lungo tempo abbandonato a causa di piccole variazioni di lettura durante le operazioni di bloccaggio;

⁴ La vite deve essere protetta, con uno sportellino, per evitare una sua casuale rotazione durante le misure.

recentemente sono stati però realizzati sistemi di blocco del cerchio più efficaci, che sembrerebbero riproporre tale metodo.

Notiamo comunque che, con la ripetizione, le letture non sono distribuite regolarmente lungo la graduazione e che, quindi, l'eliminazione degli errori di graduazione è meno efficace che con il metodo della reiterazione (anche se si ottiene una maggiore rapidità dovuta al minor numero di letture).

Per la misura di più angoli con centro in un unico punto di stazione si usano due metodi (fig. 6.3):

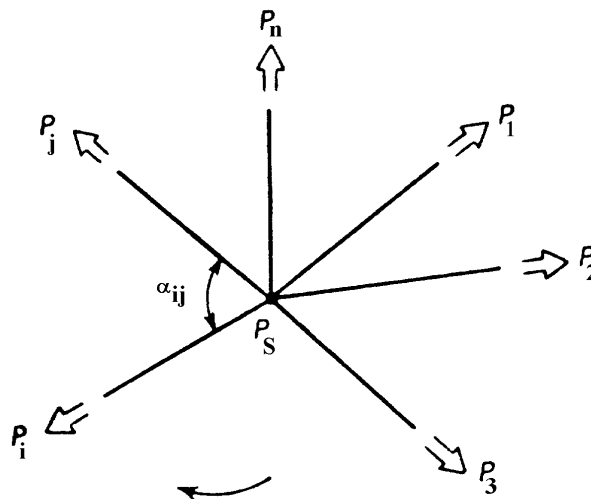


Fig. n. 6.3 - Misure di angoli azimutali da una singola stazione

- quello delle **direzioni isolate** in cui si misurano gli angoli relativi alle vere direzioni con riferimento ad un'altra direzione, individuata con un segnale a distanza «modesta» e scelta solitamente verso il nord geografico (per compensare leggeri errori di collimazione dovuti a misure effettuate in diverse ore della giornata);

- il metodo **a strati** che sceglie, invece, come origine una delle direzioni stesse (verso nord), misurando di seguito tutti gli angoli.

Il metodo delle direzioni isolate è il più preciso. In entrambi i casi, ogni misura sarà il risultato di quattro letture (regola di Bessel), la misura viene effettuata più volte con metodo reiterativo (spostando di γ il C.O.) .

Per ogni altro errore sistematico e quindi conosciuto, si usa una metodologia analoga: dapprima si studia «quantitativamente» la causa che lo determina; apprezzato l'errore, se questo supera i limiti di precisione dello strumento, bisognerà tenerne conto nei calcoli e/o fare in modo di eliminarlo.

Dal punto di vista pratico gli errori accidentali dipenderanno da alcuni elementi operativi, quali il centramento del punto di stazione S, la collimazione dei vari punti (solitamente effettuata servendosi dei cosiddetti “scopi” di collimazione), le condizioni atmosferiche, gli orari di rilevamento, piccoli movimenti dello strumento e/o degli scopi, ecc.

Si tenga presente, in quanto segue, che la misura degli angoli fatta con estrema precisione ha conservato molta importanza non tanto per le misure di rilievo (per le quali vengono utilizzati soprattutto i più moderni e precisi distanziometri), quanto per lo studio delle deformazioni di una struttura, al fine di tenere sotto controllo i fenomeni deformativi della stessa e prevenirne i dissesti (dighe, muri di sostegno, cedimenti di manufatti, ecc.).

In tal caso diventa molto importante valutare il valore di σ_α prima di effettuare le misure.

6.2 - Misura degli angoli zenitali

Consideriamo oraria la graduazione del C.V. e ricordiamo che questo è mobile insieme al cannocchiale, mentre gli indici di lettura sono fissi all'alidada (il verso crescente delle letture C.S. è quello antiorario, opposto alla graduazione); supponiamo anche, per ora, il teodolite perfettamente rettificato.

Si voglia determinare la distanza zenitale di un punto A rispetto al centro O dello strumento. Si collima il punto nella posizione C.S. e si esegue la lettura S sul cerchio, in corrispondenza dell'indice. In queste condizioni l'asse di collimazione forma con l'asse "a" (e quindi con la verticale) l'angolo "z". (fig. 6.4)

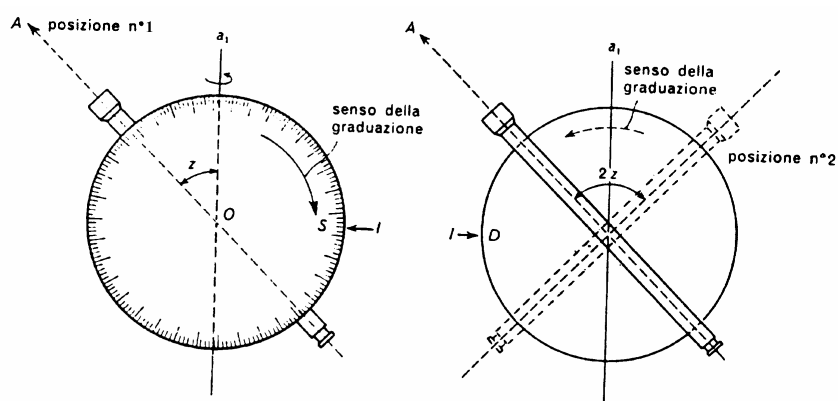


Fig. n. 6.4 - Determinazione di una distanza zenitale

Ruotiamo l'alidada di 200° e ruotiamo il cannocchiale fino a collimare di nuovo il punto A. L'angolo di cui è ruotato il cannocchiale è $2z$; detta D la lettura nella posizione C.D., tale angolo $2z$ è dato da $S-D$.

Si ha dunque: $2z = S - D$ (a meno di un multiplo dell'angolo giro). Si noti che a seconda della posizione delle due letture si deve aggiungere 400° ad S.

Osserviamo che la media delle letture $X = \frac{S+D}{2}$ rappresenta la

lettura che si farebbe con il cannocchiale diretto secondo la bisettrice dell'angolo, cioè secondo l'asse "v". Infatti:

$$z = S - X; \quad z = (S - D)/2; \quad D = S - 2z = S - 2(S - X); \quad X = (S + D)/2.$$

X è lo «zenit strumentale» ed ha importanza per le misure rapide delle distanze zenitali, perchè se esso fosse nullo, si avrebbe $D = -S$ e quindi $z = S$.

Molti teodoliti hanno in effetti il C.V. montato in modo che X sia o possa essere reso nullo con un piccolo spostamento degli indici; in tal caso è possibile determinare X , renderlo uguale a zero e misurare gli angoli zenitali con una sola lettura nella posizione C.S.

Per rettificare X , collimato un punto nella posizione C.S., si spostano gli indici fino a fare una lettura uguale a z ; si tenga tuttavia presente che tale rettifica non sempre si conserva nel tempo per l'instabilità dello strumento ed un errore residuo nel valore 0 di X , sarebbe causa di uno scostamento sistematico in tutti gli angoli misurati.

Supponiamo, invece, che l'asse principale a sia inclinato di un angolo v nel piano verticale che contiene il punto collimato, cosa che avviene di solito. (fig. 6.5)

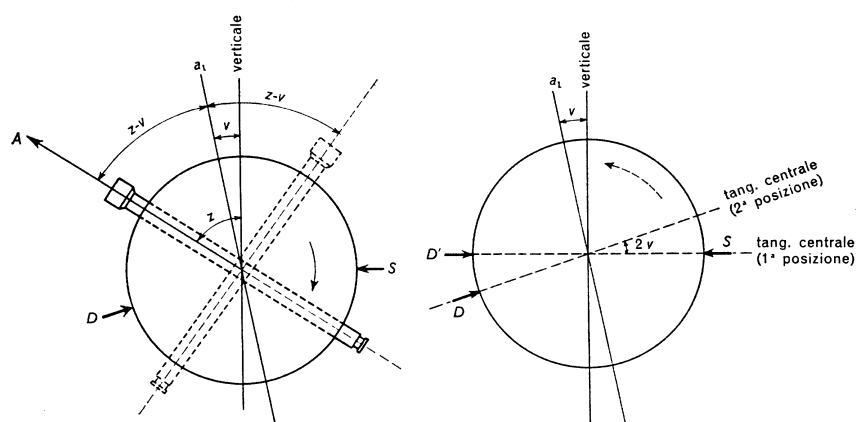


Fig. n. 6.5 - Influenza dell'errore di verticalità nella misura di una distanza zenitale.

Collimato il punto A, **c** forma con **a** un angolo $z - v$ (se l'asse **a** è inclinato dalla parte del punto). Sia S la lettura in posizione C.S., quando si ruota l'alidada di 200^g intorno all'asse **a**, l'asse **c** assume una posizione simmetrica della prima rispetto ad **a**, cioè forma con esso ancora l'angolo $z - v$; per ricollimare il punto occorre, quindi, ruotare il cannocchiale di $2(z - v)$, per cui si ha $S - D = 2(z - v)$ cioè: $(S - D)/2 = z - v$.

Calcolando la semidifferenza delle due letture non si ottiene più la distanza zenitale ma $z - v$ (se l'asse **a** è inclinato in una direzione qualunque, v è la componente dell'inclinazione nel piano di collimazione).

Per eliminare tale errore si fa ricorso ad una livella posta in prossimità del C.V. (negli strumenti più precisi una livella a coincidenza, con sensibilità dimezzata), con la tangente centrale parallela al piano descritto da C per ogni posizione dell'alidada e collegata agli indici di lettura.

Prima di eseguire ciascuna lettura, la livella viene centrata (tramite una vite a piccolo passo che fa ruotare anche gli indici solidali alla livella): in queste condizioni la semidifferenza delle letture dà il valore corretto della distanza zenitale.

Infatti sia S la lettura nella posizione C.S. dopo il centramento della livella (la quale ha quindi la t.c. orizzontale). Quando si ruota di 200^g l'alidada, la t.c. si inclina di $2v$; indichiamo con D la lettura che si farebbe in queste condizioni dopo aver collimato il punto.

Se si centra la livella, l'indice si sposta in senso tale da diminuire la lettura di $2v$; se D' è la nuova lettura: $D' = D - 2v$ e quindi:

$$(6.12) \quad \frac{S - D'}{2} = \frac{S - (D - 2v)}{2} = \frac{S - D + 2v}{2} = z - v + \frac{2v}{2} = z$$

Vediamo cosa accade quando si effettuano le misure nella sola posizione C.S. Anzitutto osserviamo che, in tal caso $(S + D)/2$

fornisce una lettura $Z'=Z+v$ che si avrebbe quando “c” è diretto secondo l’asse “a”.

Se abbiamo eseguito le letture con bolla centrata si ha:

$$(S+D')/2 = (S+D)/2 - v = Z' - v = Z, \text{ ovvero «zenit strumentale»}.$$

Per correggerlo bisogna, quindi, spostare gli indici e dopo ricentrare la livella con la sua vite di rettifica.

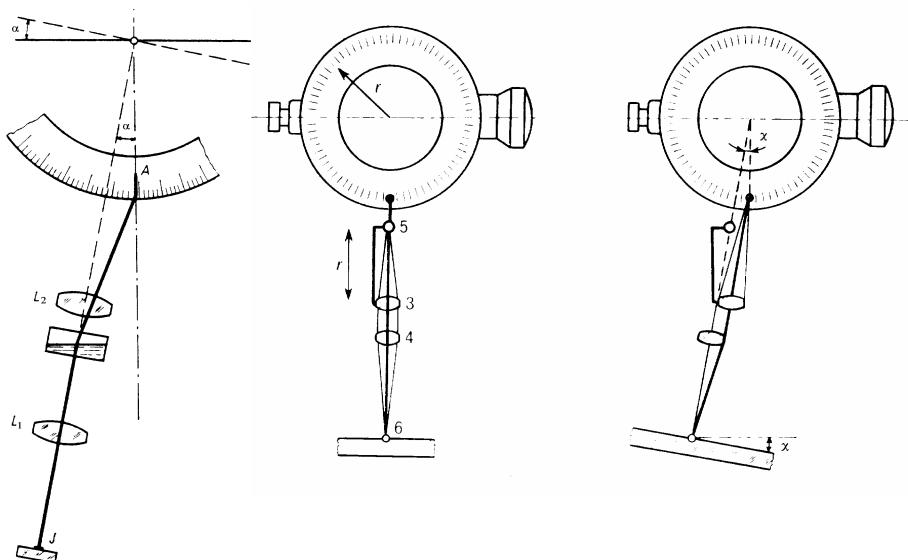
Se lo zenit strumentale è così rettificato, le letture S , *sempre a bolla centrata*, sono uguali alle distanze zenitali, qualunque sia l’inclinazione dell’asse “a” (sia infatti $S-Z$ la lettura con asse verticale e la livella centrata: pensiamo di inclinare “a” verso A di v ; il cannocchiale ruota dello stesso angolo e per collimare il punto occorre ruotare il cannocchiale della stessa quantità in senso inverso, per cui si farà una lettura $S' = S - V$).

Anche la t.c. della livella si è però inclinata dello stesso angolo; ricentrando la bolla, gli indici si spostano e la lettura aumenta di V , cioè torna uguale ad S .

6.3 - Indice Zenitale Automatico

La livella zenitale viene spesso sostituita da un sistema a pendolo, che corregge automaticamente l’influenza dell’errore di verticalità.

Il principio è semplice: il pendolo ha l’asse di oscillazione sull’asse di rotazione del cannocchiale ed è collegato con gli indici di lettura. Poichè il pendolo è sempre verticale, indipendentemente dall’inclinazione dell’asse “a”, il sistema equivale ad una livella zenitale sempre centrata, cioè sempre con la tangente centrale orizzontale.(fig. 6.6-6.7)



a)

b)

Fig. n. 6.6-6.7 - Indice zenitale automatico

- a) In questo caso è il liquido che è sempre orizzontale e funge da “compensatore”. (Wild)
- b) In questo caso il pendolo causa la riflessione del percorso ottico di lettura sul C.V. (Kern)

CAPITOLO VIII

MISURA DELLE DISTANZE

Capitolo VIII

Misura delle distanze

1 - Allineamenti

Ricordiamo anzitutto che per distanza tra due punti della superficie terrestre si intende la **lunghezza dell'arco di geodetica che congiunge le proiezioni dei due punti sulla superficie di riferimento**; si è visto, inoltre, che anche per distanze di parecchie centinaia di chilometri tale lunghezza può ritenersi uguale alla lunghezza di una delle due sezioni normali che congiungono tali punti sull'ellissoide di riferimento.

E' necessario pertanto stabilire:

- a) le modalità per individuare sulla superficie terrestre la traccia di una sezione normale;
- b) le modalità per riportare una misura effettivamente eseguita sulla superficie del terreno alla superficie di riferimento.

In merito al primo punto, se la distanza è inferiore o uguale alla portata dello strumento di misura (ed è quindi misurata con una sola operazione), l'individuazione della sezione normale è intrinsecamente realizzata nel metodo di misura; se la distanza viene misurata suddividendola in tronchi per ognuno dei quali lo strumento è utilizzato per un valore minore o uguale alla sua portata, si individua la traccia richiesta sul terreno tramite un'operazione chiamata "allineamento", facilmente eseguibile con un teodolite.

Dovendo misurare la distanza A-B, si pone in A un teodolite ed in B un segnale; collimato B e fissata l'alidada, il piano descritto da "c" supposto lo strumento rettificato, interseca la superficie terrestre tra A e B secondo una traccia corrispondente alla sezione normale; sui punti del terreno individuati con il centro del reticolo, a distanze opportune, si pongono dei segnali (paline), che consentono di eseguire la misura lungo l'allineamento.

Solo se i singoli segnali non si discostano dalla sezione normale più di 3-4cm, l'errore sistematico ottenuto per imprecisione di allineamento è trascurabile usando un campione di 20m.

In merito al punto b) ricordiamo anzitutto che, se la distanza non eccede i 150Km, la superficie ellissoidica si discosta poco da una sfera di raggio $R = \sqrt{\rho N}$ ed è sempre consentita l'ipotesi che la distanza sia assimilabile ad un arco di cerchio massimo sulla sfera locale.

Esaminiamo pertanto il caso più generale della riduzione sulla superficie di riferimento di una distanza notevole misurata in un'unica operazione (per esempio con un distanziometro ad onde).

Sia d_e la distanza effettiva misurata tra A e B, e z la distanza zenitale di B rispetto ad A. Sia d_a la distanza AB valutata sul piano orizzontale di A e d_0 la distanza sulla superficie di riferimento, ovvero topograficamente la distanza AB (fig. 1.1).

Si ha $d_a = d_e \times \text{senz.}$ Indicando con " Δ " la correzione di convergenza, ovvero la differenza fra d_a ed il tratto compreso tra le due verticali per A e B, con alcune approssimazioni si ottiene:

$$(1.1) \quad d_0 = d_a - \Delta - (d_a - \Delta) \times (Q/R)$$

(con Q si è indicata la quota di A).

Per Q piccola (<6m) l'ultimo termine è trascurabile.

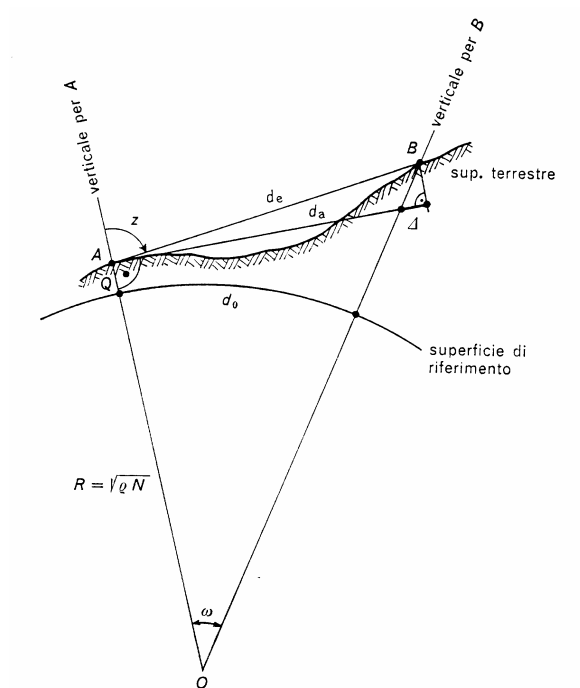


Fig. n. 1.1 - Riduzione alla superficie di riferimento di una distanza.

E' bene osservare che la riduzione delle distanze alla superficie di riferimento deve essere eseguita solo per rilievi a scopo cartografico; quando, invece, le misure hanno come scopo la esecuzione di lavori, quali tracciamento di gallerie, funivie, rilievi di bacini idroelettrici ecc. la riduzione non deve essere apportata, poichè interessano le distanze reali.

2 - Metodi diretti di misura delle distanze

Consistono nel riporto di un campione sull'allineamento stabilito e nella valutazione di quante volte il campione e i suoi sottomultipli sono contenuti nella distanza da misurare. Il campione è, generalmente, costituito da una "rotella" metrica da 10, 20, 50, 100m.

Tali misure dirette, con fili o nastri, devono essere sempre fatte su terreno piano o possibilmente poco inclinato, specialmente se si desidera raggiungere una elevata precisione. Esaminiamo le diverse cause di errori in tali misure:

Errori sistematici.

Sono dovuti a:

1. Taratura del campione (non costante nel tempo, per l'elasticità dei materiali costituenti i campioni);
2. Temperatura. Le variazioni di temperatura possono essere causa di errori abbastanza rilevanti; per diminuirne molto l'influenza si costruiscono, nastri o fili di un materiale come l'INVAR (36% Fe + 63% Ni + C e altre sostanze) che ha un coefficiente di dilatazione di $1\mu\text{m}$ al metro e al grado ($10^{-6}\text{m/m}^\circ\text{C}$). Le rotelle metriche sono dotate di una scala graduata con la indicazione delle correzioni necessarie al variare della temperatura.
Inoltre, nelle misure di precisione, si fa in modo che il nastro si trovi in condizioni ambientali più uniformi possibili (evitando, ad esempio, che una parte di esso si trovi al sole ed un'altra all'ombra);
3. Inesatta valutazione della tensione. Per ottenere precisioni dell'ordine di 10^{-6} bisogna servirsi dell'apparato di Jaderin: con fili o nastri di "INVAR", lunghi 24 m e graduati al mm. Le "rotelle" metalliche devono essere ben tese durante la misura e sono preferibili a quelle di plastica o di stoffa che, con il tempo, si allungano.

4. Inesatto allineamento; risulta l'errore più pericoloso e deve essere eliminato con uno strumento ottico: sull'allineamento si devono disporre dei segnali (Paline) muniti di tratti di riferimento, tra i quali si determina la distanza.
5. Correzione di convergenza e/o riduzione all'orizzonte per le distanze maggiori di qualche km.

Errori accidentali.

Se d è la lunghezza del campione e $D = nd$ la lunghezza misurata, supposte le n misure di egual precisione σ_0 , si avrà per σ_D (varianza della variabile casuale D) l'espressione:

$$\sigma_D^2 = n \times \sigma_0^2 = \frac{D}{d} \sigma_0^2 \quad \text{e quindi} \quad \sigma_D = \pm k \times \sqrt{D}, \quad \text{ovvero l'errore}$$

(s.q.m) è proporzionale alla radice quadrata della distanza D .

3 - La misura delle distanze mediante onde elettromagnetiche

Ricordiamo che le onde elettromagnetiche sono generate da una corrente elettrica variabile nel tempo in un circuito che termina con un'antenna.

Un'onda sinusoidale viene emessa da uno strumento posto nell'estremo A della distanza AB da misurare, si riflette su un "prisma ottico" posto in B e torna al punto A impiegando un tempo Δt .

Poichè tale intervallo di tempo non può essere facilmente misurato con precisione (tenendo conto che $c = 3 \times 10^8$ m/sec, per avere la precisione di 1m bisognerebbe valutare Δt con precisione di 10^{-8}), si paragonano nel punto A l'onda emessa e quella riflessa, determinando lo sfasamento ϕ tra le due onde; utilizzando lunghezze d'onda diverse, dalla misura di ϕ si può ottenere la misura della distanza AB (fig. 3.1).

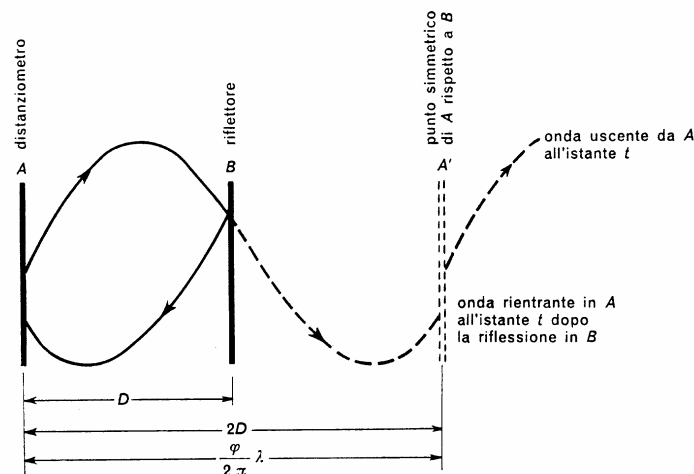


Fig. n. 3.1 - Schema della misura di una distanza D mediante determinazione dello sfasamento di un'onda

Supponiamo che in A si abbia un'oscillazione trasversale sinusoidale: $S=A \times \sin(\omega t + \phi_0)$ dove ω è la pulsazione dell'oscillazione e T il suo periodo (intervallo di tempo necessario affinché si ristabilisca, con le stesse modalità, uno stesso valore dell'oscillazione).

Supponiamo dapprima che $d = AB < \lambda/2$. In un istante t lo stato dell'oscillazione in A è $S_1 = A \times \sin(\omega t + \phi_0)$; nello stesso istante e nello stesso punto l'onda riflessa ha uno stato di oscillazione pari a $S_2 = A \times \sin[\omega(t + \Delta t) + \phi_0] = A \times \sin(\omega t + \omega \Delta t + \phi_0)$ dove Δt è il tempo impiegato dall'onda a coprire la distanza 2d.

Si produce, quindi tra le due onde uno sfasamento:

$$(3.1) \quad \phi = \omega \Delta t = \frac{2\pi}{T} \times \frac{2d}{v} = \frac{2\pi}{\lambda} \times 2d$$

ed è chiaro che dalla misura dello sfasamento può ottenersi:

$$(3.2) \quad d = \frac{\varphi}{2\pi} \times \frac{\lambda}{2}.$$

Se l'estremo B della distanza, quello sul quale l'onda si riflette, viene spostato di mezza lunghezza d'onda lungo il percorso, φ non cambia in quanto sul percorso ABA viene ad inserirsi un'onda completa, ed altrettanto dicasi se B viene spostato di n mezze lunghezze d'onda. Possiamo, quindi, stabilire l'equazione fondamentale del distanziometro ad onde:

$$(3.3) \quad d = n \times \frac{\lambda}{2} + \frac{\varphi}{2\pi} \times \frac{\lambda}{2}$$

Un distanziometro deve allora essere costituito da due parti:

- 1) una posta in A, che deve avere un oscillatore, un rilevatore ed un discriminatore di fase (strumento);
- 2) una parte posta in B, sensibile all'onda in arrivo ed in grado di rinviarla in direzione di A (prisma ottico in vetro).

Tenendo presente che il discriminatore di fase consente in genere di misurare la differenza di fase φ con $0 < \varphi < 2\pi$ con $\sigma_{\varphi} = \frac{2\pi}{2000} \frac{\varphi}{2\pi} \times \frac{\lambda}{2}$, φ può essere misurata con uno s.q.m. che va da $0.5 \times 10^{-3} \lambda/2$ a $10^{-3} \lambda/2$ e si evince che volendo una distanza con la precisione del centimetro occorre generare un'onda di misura che abbia una lunghezza dell'ordine di 20m.

Per la determinazione di n basterebbe utilizzare due lunghezze d'onda λ_1 e λ_2 purchè n non vari nei due casi; ciò comporterebbe l'utilizzo dello strumento entro limiti molto modesti (circa 110m). Si fornisce pertanto l'apparecchio di parecchie frequenze solitamente corrispondenti a lunghezze d'onda multiple di 10.

Supponiamo che l'apparecchio abbia la frequenza di 15 kHz: ($\lambda_1/2=10$ km), 150 kHz ($\lambda_2/2=1$ km), 1.5 MHz ($\lambda_3/2=100$ m), 15 MHz ($\lambda_4/2=10$ m). La misura si esegue utilizzando dapprima $\lambda_1 = 20$ Km con il che si ottiene la distanza supposta (< 20 km e quindi con $n=0$) con uno s.q.m. di ± 10 m; di questa determinazione si

conservano le cifre dei chilometri e degli ettometri; si passa quindi alla lunghezza d'onda $\lambda_2/2=1$ km e si ottiene una distanza diversa da quella effettiva per un numero intero di chilometri, ma in cui le cifre corrispondenti agli ettometri ed ai decametri sono quelle effettive; si prosegue con $\lambda_3/2 = 100$ m ottenendo i metri ed infine con $\lambda_4/2 = 10$ m che consente di valutare i decimetri ed i centimetri (con uno s.q.m. di ± 1 cm).

L'uso delle onde elettromagnetiche per la misura delle distanze comporta la soluzione di alcuni problemi. Per evitare che l'energia si propaghi e si disperda in tutte le direzioni occorre convogliare le onde in ristretti angoli solidi mediante l'uso di opportune antenne; ciò è favorito da alte frequenze e basse lunghezze d'onda.

Onde molto corte non possono però essere usate per la misura delle distanze per la qual cosa sono adatte, come si è visto, onde con $\lambda > 10$ m. Per risolvere la questione si ricorre alla modulazione delle onde e .m.

Si impiega cioè un'onda “portante” di frequenza molto elevata (di facile convogliabilità), ma si fa variare nel tempo una caratteristica di quest'onda (ad es. l'ampiezza) con una legge periodica che riproduce un'onda di lunghezza decisamente più elevata (onda modulata di frequenza molto minore) (fig. 3.2).

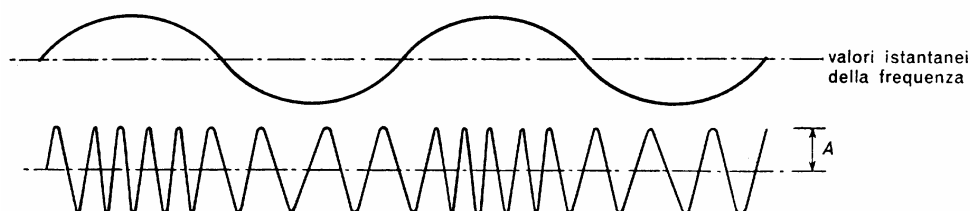


Fig. n. 3.2 - Onda modulata in frequenza

E' ovvio che esistono dei circuiti elettronici chiamati modulatori che consentono, una volta prodotta l'oscillazione corrispondente

all'onda portante e quella corrispondente alla modulante, di sovrapporre le due, come pure vi sono dei circuiti demodulatori che consentono di separarle; è chiaro anche che la funzione dell'onda di misura. deve essere assolta dall'onda modulante.

Sia l'onda portante che quella modulante sono prodotte da oscillatori, ovvero da circuiti aventi un generatore di frequenza (quarzo piezoelettrico) in grado di produrre e mantenere l'oscillazione sinusoidale di una grandezza elettrica. Se le onde utilizzate sono quelle luminose solitamente esse vengono modulate per polarizzazione.

La luce può essere modulata anche direttamente, quando la stessa sorgente luminosa emette una radiazione la cui intensità varia con legge sinusoidale. Ciò si può ottenere con un diodo elettro - luminescente di Arseniuro di Gallio (GaAs); questo semiconduttore emette una radiazione la cui intensità è proporzionale alla corrente (detta di "eccitazione") cui è attraversato (fig. 3.3).

Potendo variare sinusoidalmente tale corrente di eccitazione si ottiene una luce modulata in intensità con la stessa caratteristica della corrente.

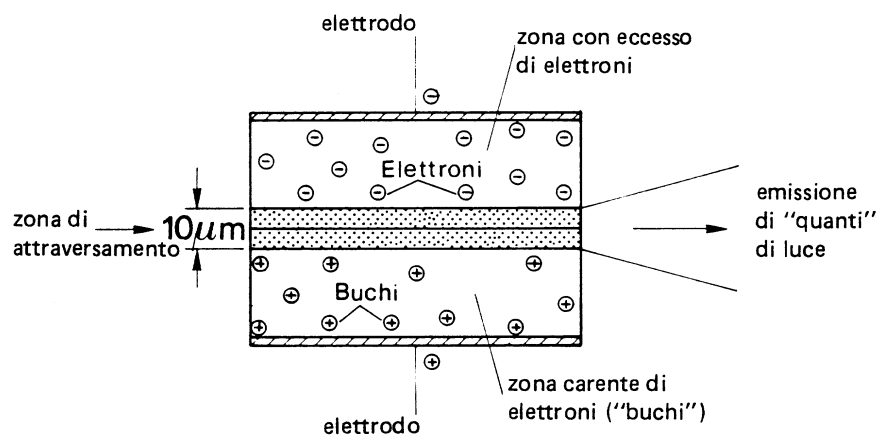


Fig. n. 3.3 - Schema semplificato di un diodo fotoemittente

Il riflettore (per le onde non luminose esso doveva svolgere anche la funzione attiva di amplificarle) è in genere costituito da uno o più prismi retrodirettivi, godenti della proprietà di poter rinviare un fascio luminoso nella stessa direzione da cui esso proviene.¹ (fig. 3.4-3.5)

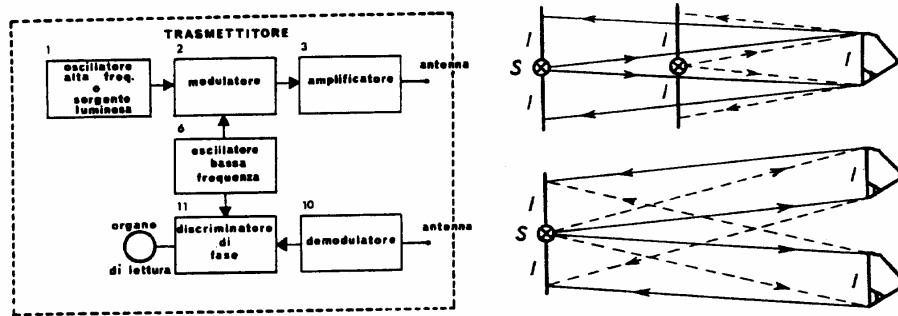


Fig. n. 3.4-3.5 - Schema Trasmettitore – Riflessione della luce su prismi retrodirettivi.

Se si osservano le indicazioni fornite da un discriminatore di fase durante un certo intervallo di tempo, si possono notare delle variazioni in genere lente ma superiori a quelle che ci si attenderebbe in relazione agli $s.q.m. = \frac{2\pi}{1000} 0 \frac{2\pi}{2000}$.

¹ Si diffondono sempre di più, negli ultimi tempi, distanziometri che consentono, in virtù della potenza di emissione “laser”, di effettuare misure senza riflettore, fino ad una certa distanza (200÷300m) in funzione delle caratteristiche di “riflettanza” della superficie interessata.

Per “portate” maggiori è necessario l’uso di uno o più prismi, che consentono di misurare distanze anche fino a 10km.

Ciò è dovuto da una parte all'instabilità dell'atmosfera lungo il percorso, ma in maniera più rilevante all'instabilità dei circuiti elettronici. Lo strumento stesso si impone, quindi, un controllo “verificando” ogni misura con una distanza predeterminata al suo interno, detta “linea di calibrazione”, e mediando un numero molto elevato di misure (sempre in un tempo molto breve).

Indicazioni date dalle case costruttrici (abachi) permettono di correggere le misure.

3.1 - Errori Sistematici

Gli errori sistematici nella misura elettronica delle distanze sono dovuti a due cause:

- errori di centramento dello strumento e/o del prisma, tra i quali possiamo inserire l'errore dovuto alla “costante” del prisma, che dovrebbe essere uguale a zero quando il percorso delle onde all'interno del prisma è uguale allo scostamento tra punto di riflessione e punto “B” estremo della distanza (materializzato sul terreno).
- differenza di velocità di percorso delle diverse onde elettromagnetiche aventi diverse lunghezze d'onda, a causa del diverso indice di rifrazione dell'atmosfera. Questo indice è influenzato dai fattori ambientali (Temperatura, Pressione, Umidità, ecc.) e può essere calcolato con formule apposite (Barrel e Sears) solo nei due estremi A e B della distanza da misurare, dopo aver misurato i rispettivi parametri.(fig. 3.6)

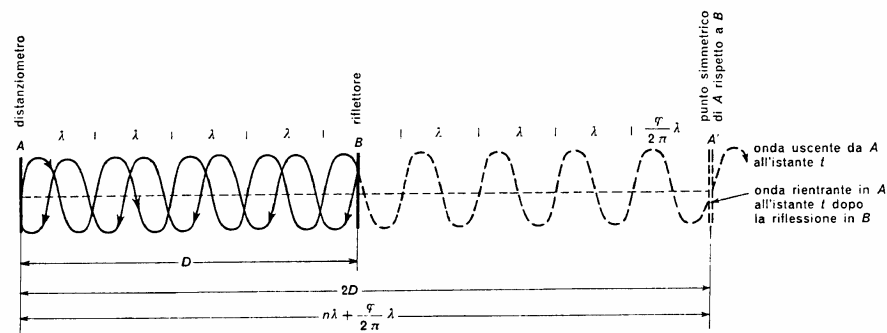


Fig. n. 3.6 - Schema della misura di una distanza mediante determinazione dello sfasamento di un'onda e determinazione del numero intero di lunghezze d'onda ($D > \lambda/2$).

Lo strumento richiede, quindi, delle indicazioni sommarie di T, P, u%, che vengono memorizzate dal computer e permettono di calcolare automaticamente le correzioni da apportare alla misura.

In conclusione, volendo ricavare l'incertezza della singola misura si indicherà come:

$$(3.4) \quad \sigma_d = \pm(m + kD^{-6})$$

dove la parte fissa "m" può variare da 0.1mm fino a qualche mm, a seconda del "costo" e delle caratteristiche del distanziometro; invece, k esprime la quantità proporzionale alla distanza (espressa in km) variabile da frazioni di mm fino a qualche mm, in funzione delle caratteristiche ambientali.

CAPITOLO IX

MISURA DEI DISLIVELLI

Capitolo IX

MISURA DEI DISLIVELLI

1 - Principio della livellazione geometrica e livelli

E' noto che ciascun punto della superficie fisica terrestre può essere individuato altimetricamente tramite la sua quota, che è la distanza del punto stesso dalla superficie di riferimento (geoide) misurata lungo la verticale (per ora si chiamerà quota o quota "ortometrica").

Le livellazioni sono operazioni che consentono di misurare la differenza di quota tra punti o, come si suol dire, il loro dislivello; la misura assoluta delle singole quote è, infatti, impossibile. Per misurare queste ultime è necessario, allora, conoscere le quote assolute di alcuni punti di riferimento (capisaldi di livellazione) determinate in altro modo.

La superficie libera del mare, qualora non esistessero il moto ondoso, le azioni di marea ed altre azioni perturbanti (correnti, variazioni di densità, ecc.), coinciderebbe con la superficie del geoide, essendo superficie di equilibrio di un fluido. D'altra parte, i movimenti del mare hanno carattere periodico e, quindi, fanno oscillare la superficie libera attorno ad un valore medio, il quale corrisponde al livello del mare supposto in quiete.

Esso viene determinato con appositi strumenti, quali: i mareografi, i medimarimetri, etc.

I **mareografi** sono costituiti da un galleggiante disposto in un pozzo in comunicazione col mare, ma in luogo molto riparato (moto ondoso insensibile). Il galleggiante è collegato da un sistema di leve e di carrucole ad una punta scrivente sopra un foglio di carta

parametrato, avvolto attorno ad un cilindro che ruota con velocità uniforme. La penna traccia così, in una scala opportuna, il diagramma del livello marino; determinando l'ordinata media del diagramma si ottiene il livello medio, esteso al periodo di integrazione. Tale strumento è stato sostituito, nel tempo, da opportuni contatori digitali che registrano ed eventualmente, trasmettono la posizione del l. m. m.

Le osservazioni e le integrazioni si ripetono giornalmente e poi si mediano le ordinate trovate per periodi di tempo uguali al periodo della più lunga marea, ottenendo così il livello medio. Per misure più precise, i periodi di osservazione diventano maggiori di 18 anni (tenendo presente che, comunque, si hanno variazioni alle varie epoche).

In Italia il mareografo fondamentale, che stabilisce, mediante il collegamento ad un apposito caposaldo, l'origine di tutte le linee di livellazione (e cioè di tutti i capisaldi), è quello di Genova (fig. 1.1).

I capisaldi possono essere orizzontali (basamento di calcestruzzo fondato nel terreno sul quale è appoggiata una piramide tronco-conica di ceramica al cui vertice è una sfera anch'essa di ceramica, il tutto protetto da un chiusino metallico), o verticali (piccole mensole incassate nelle murature e terminanti con una sfera): in entrambi i casi la quota, tabellata con precisione del decimo di mm, corredata di numerosi dati accessori (v. monografia) per l'individuazione del punto, viene riferita alla sommità della sfera (fig. 1.2 e 1.3).

Le linee fondamentali di livellazione (fig. 1.1) sono situate sulle strade statali (ad es. in Puglia la "vecchia" S.S. n.16).

I capisaldi sono di quattro tipi diversi:

1. **di linea**, posti alla distanza media di 1 km;
2. **principali**, posti alla distanza media di 2-3 km;
3. **fondamentali**, posti ogni 25 km;
4. **nodali**, posti all'incrocio di più linee; tutti devono essere in sede propria (orizzontali). I capisaldi verticali sono solitamente aggiunti ai precedenti per verificarne gli eventuali spostamenti; la loro distanza media è solitamente di $\cong 3$ km.

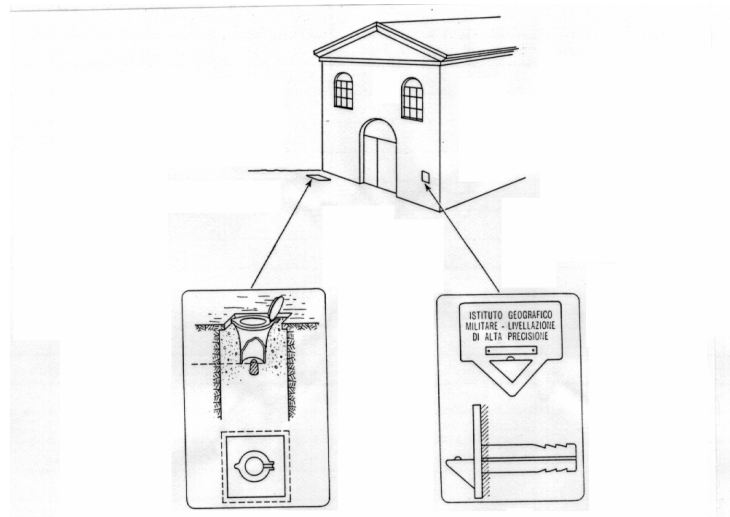
ISTITUTO GEOGRAFICO MILITARE
DIVISIONE GEODETICA



Fig. n. 1.1 - Linee di livellazione geometrica di alta precisione.

La monografia di un caposaldo (fig.3) può essere richiesta all'I.G.M. specificando la più vicina linea di livellazione.

Fig. n. 1.2 - Caposaldi di livellazione: orizzontale e verticale



Istituto Geografico Militare		mod. 282/G-AR	
Comando Regione Militare Nord-Ovest-Torino			
Staz. CC. Gallarate			
Comune Gallarate		Prov. Varese	
Proprietario Autorità Ecclesiastica			
Quota		Coordinate Gauss-Boaga	
H = 238.6505		N	
Riferita a: csoc		E	
		Coordinate geografiche	
		φ	
		λ	
		45° 39' 44" - 8° 39' 41"	
MONOGRAFIA (1953)		Pianta o Prospettivo	
021380	GALLARATE. Chiesa di S. Francesco.	Lv - III - 044 - 59/20	
● A 0238 8710	alla base della facciata. spigolo di sinistra		
Nome	Cl. Cat.	F ^a	N ^o
GALLARATE. Chiesa di S. Francesco	Lv III	044	● 59/20

Fig. n. 1.3 - Monografia di un caposaldo di livellazione di alta precisione

Il principio geometrico della livellazione è molto semplice: si considerino due punti A e B a breve distanza (non superiore ai 100 - 150 m); poniamo su essi due stadie verticali e nel punto medio M un **livello**, strumento che realizza un asse di collimazione orizzontale, ossia parallelo al piano tangente in M_0 al geoide. (fig. 1.4)

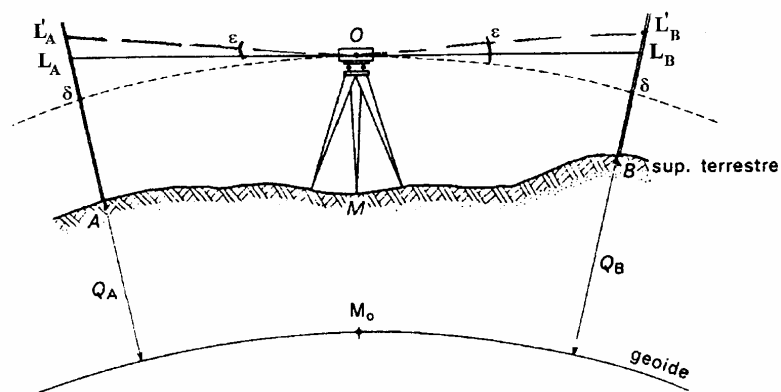


Fig. n. 1.4 - Misura diretta di un dislivello mediante livello a cannocchiale.

Nell'ipotesi che il geoide si confonda, nel breve tratto considerato, con la sfera locale in M_o (o addirittura con il piano tangente):

$$(1.1) \quad Q_A + L_A = Q_B + L_B \Rightarrow \Delta_{AB} = Q_B - Q_A = L_A - L_B$$

È questa l'operazione elementare di una livellazione geometrica chiamata, normalmente, "battuta" di livellazione.

Se la lunghezza della battuta è modesta, tale da rendere trascurabile l'influenza della curvatura terrestre rispetto alla precisione della misura (per una distanza di 100 m: $\delta = d/R \Rightarrow (100\text{m}/6,4 \times 10^6) \times 206265 \cong 3''$, solitamente trascurabile rispetto all'errore di verticalità della stadia, dell'ordine dei primi), è necessario che l'asse di collimazione venga disposto secondo l'orizzontale del punto M, e lo strumento deve essere messo in una posizione intermedia alle due stadi, possibilmente nel mezzo.

Un livello (fig. 1.5) è costituito schematicamente da una traversa, girevole attorno ad un asse "a" e impernata quindi su una base B con viti calanti per rendere grossolanamente verticale l'asse "a"; il cannocchiale che materializza l'asse di collimazione è collegato alla traversa con un sistema costituito da una vite di elevazione E e (schematicamente) da una molla M che gli permette piccole rotazioni attorno ad un asse orizzontale per O.

Rigidamente collegata al cannocchiale, a parte le viti di rettifica, è invece la livella torica L.

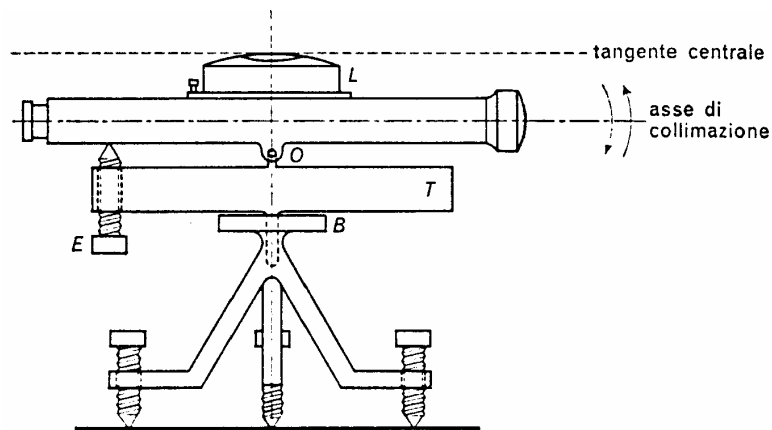


Fig. n. 1.5 - Schema del livello con vite di elevazione.

Solitamente la base B dispone anche di una livella sferica S che consente di disporre grossolanamente verticale l'asse "a". Collimata la stadia, si agisce sulla vite di elevazione per il centramento della livella torica. Questa può essere anche a coincidenza (fig. 1.6-1.7). Lo strumento é "rettificato" quando l'asse di collimazione e la tangente centrale della livella torica sono paralleli. È evidente, allora, che centrando la bolla la tangente centrale si dispone orizzontale e, conseguentemente, anche l'asse di collimazione risulta orizzontale.

Quando la traversa viene ruotata per collimare "B", a causa dell'errore di verticalità di "a", la bolla non rimane centrata ed una nuova manovra sulla vite E consente di ripristinare l'orizzontalità dell'asse di collimazione.

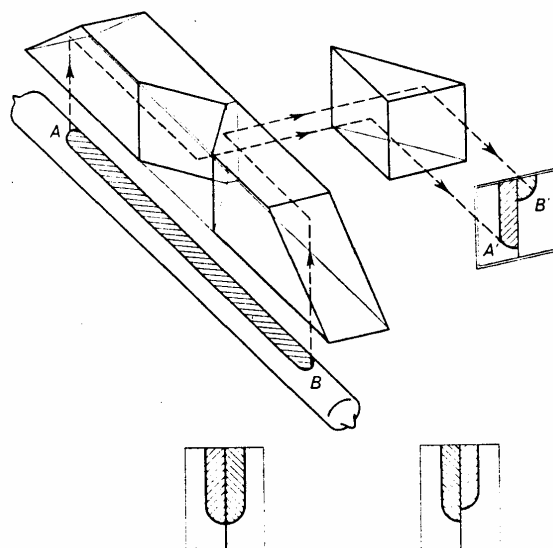


Fig. n. 1.6 - Schema della livella a coincidenza.

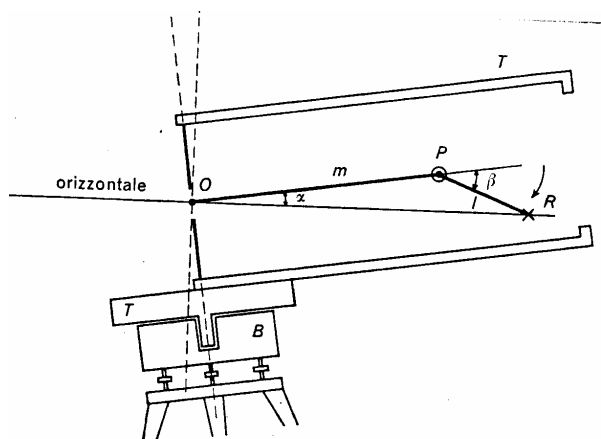


Fig. n. 1.7 - Principio del livello autolivellante.

Quando nei livelli “da cantiere” manca la vite di elevazione, possiamo, eventualmente, centrare la bolla della livella torica con una delle viti di base, attraverso una operazione che rende orizzontale l’asse di collimazione.

1.1 - Livelli Automatici

Supponiamo che il cannocchiale sia sostituito da un tubo “T” nel quale sono disposti un foro “O” (obbiettivo) e il reticolo “R”, vincolato ad un braccio mobile “I” ruotante attorno a P distante m da O.

Se il tubo (ed anche OP) ruotano rispetto all’orizzontale di un angolo α basta che “I” ruoti di un angolo β che si ricava da:

$$(1.2) \quad \frac{m}{\sin(\beta - \alpha)} = \frac{l}{\sin \alpha}; \beta - \alpha = \frac{m}{l} \alpha; \beta \approx \alpha \left(\frac{m}{l} + 1 \right)$$

Per realizzare un livello automatico è sufficiente quindi un meccanismo che, al ruotare di α del tubo, faccia ruotare di β il braccio “I”.

In particolare, se $m = 0$ (“I” ruota attorno ad O), dovrà essere $\beta = \alpha$. Uno schema funzionale che realizza tale soluzione è mostrato in fig. 1.8; è sufficiente rendere solidale il braccio “I” ad un **pendolo** sospeso nel punto O.

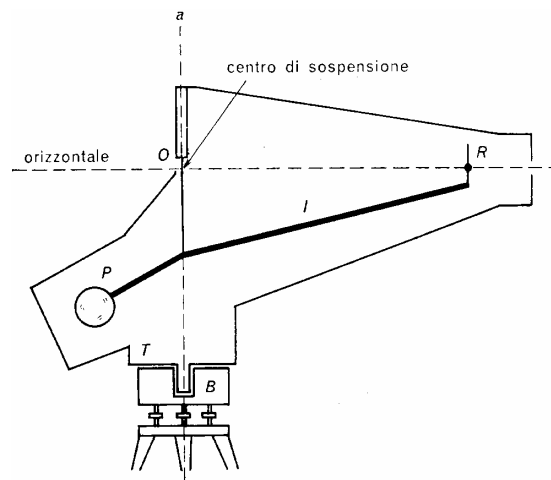


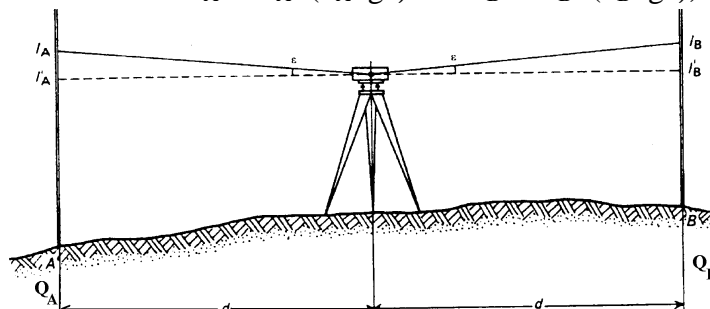
Fig. n. 1.8 - Schema costruttivo di principio del livello autolivellante.

Tutti i livelli “automatici” (forniti di pendolo) hanno bisogno, per non avere disturbo nelle letture e non dover attendere troppo tempo, di uno smorzatore delle oscillazioni; con tale accorgimento essi garantiscono una notevole rapidità di azione, e richiedono solo il centramento della livella sferica .

Esaminiamo ora, come al solito, gli errori sistematici e accidentali di una battuta di livellazione che si verificano per qualsiasi tipo di livello (automatico, inglese, da cantiere, digitali¹, laser², ecc).

Se lo strumento non è perfettamente rettificato (errore residuo), l’asse di collimazione e la tangente centrale della livella (che supponiamo orizzontale) formano un piccolo angolo ϵ di cui l’asse di collimazione resta costantemente inclinato, indipendentemente dalla direzione in cui si osserva (purché in ogni lettura si centri la livella). (fig. 1.9)

Le letture saranno: $L'_A = L_A \pm (d_A \cdot \text{tg} \epsilon)$ ed $L'_B = L_B \pm (d_B \cdot \text{tg} \epsilon)$, con



¹ Il livello “digitale” è un autolivello dotato di una telecamera che “legge” su di una stadia graduata con “codici a barre” e, quindi, permette di registrare la lettura in un apposito registratore di dati.

² Il livello “laser” è realizzato con un emettitore laser sospeso ad un pendolo che, mediante un prisma ruotante, realizza un asse orizzontale che viene materializzato con un opportuno rivelatore dotato di “beep” sonoro. Ha una portata notevole (fino a 300m) e presenta il notevole vantaggio di richiedere un solo operatore.

il doppio segno \pm a seconda dell'inclinazione di ϵ .

$$\Delta = Q_B - Q_A = L_A - L_B = (L'_A - d_A \cdot \text{tg}\epsilon) - (L'_B - d_B \cdot \text{tg}\epsilon)$$

$$\Delta = L'_A - L'_B \pm (d_A - d_B) \text{tg}\epsilon.$$

Fig. n. 1.9 - Eliminazione dell'influenza dell'errore di rettifica mediante battuta del mezzo.

Il metodo più semplice e più sicuro per l'eliminazione di tale errore residuo di rettifica, pari a $[\pm(d_A - d_B) \text{tg}\epsilon]$, consiste nel porre lo strumento ad uguale distanza dalle due stadie (**livellazione geometrica dal mezzo**), in modo che $d_A = d_B$ da cui: $\Delta = L'_A - L'_B$.

Si tenga presente che se $(d_A - d_B) \cdot \text{tg}\epsilon$ risulta dell'ordine di $2 \text{ m} \cdot 10''$ cioè $\frac{2 \cdot 10}{206.265} \approx \pm 0.1 \text{ mm}$; essa risulta una quantità trascurabile nelle operazioni tecniche, ma non per misure di alta precisione (le quote dei caposaldi sono fornite al decimo di mm).

La livellazione dal mezzo ha anche i seguenti vantaggi:

- consente di determinare la differenza di quota eliminando anche l'errore dovuto alla curvatura della superficie di riferimento (che in realtà non è mai un piano);
- rende più agevole la collimazione (non c'è bisogno dell'adattamento alla distanza, che rimane invariata);
- elimina in gran parte il piccolo effetto della rifrazione atmosferica, la quale tende ad incurvare leggermente il percorso dei raggi ottici verso il basso.

Inoltre, non è necessario misurare l'altezza strumentale che risulta di difficile ed imprecisa determinazione.

È opportuno notare che la livellazione del mezzo esige solo l'uguaglianza della distanza fra lo strumento e le stadie, ma non che il livello sia allineato con queste ultime.

Qualora non si possa usare la livellazione dal mezzo per la presenza di qualche ostacolo, si ricorre alla livellazione reciproca (fig. 1.10), che consiste nell'eseguire le letture da due stazioni intermedie,

ciascuna prossima ad una stadia e tali che risultino simmetriche rispetto al punto di mezzo:

posiz. M: $Q_B - Q_A = L_1 - L_2 - \varepsilon(D-d)$; (*)

posiz. N: $Q_B - Q_A = L_3 - L_4 + \varepsilon(D-d)$; (*)

da cui:
$$Q_B - Q_A = \frac{L_1 - L_2 + L_3 - L_4}{2}$$

(il dislivello corretto è dato dalla differenza delle medie fra le battute indietro e quelle avanti).

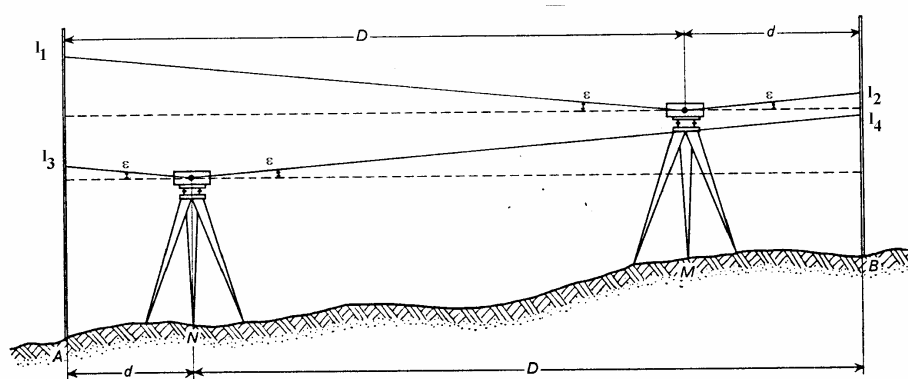


Fig. n. 1.10 - Battuta di livellazione reciproca

La precisione della livellazione reciproca è tuttavia inferiore a quella del mezzo, sia perché le distanze di collimazione sono più grandi, sia per eventuali piccoli spostamenti del livello e/o delle stadie.

1.2 - Rettifica del livello

Se si esegue dapprima una livellazione dal mezzo e, sempre dagli stessi punti, anche una livellazione a distanza qualunque, per differenza dalle (*) si ottiene: $(L_1 - L_2) - (L_3 - L_4) = 2\varepsilon(D-d)$, da cui si ricava l'errore di rettifica (fig. 1.11)

$$(1.3) \quad \varepsilon = \frac{(L_A - L_B) - (L'_A - L'_B)}{(D - d)}$$

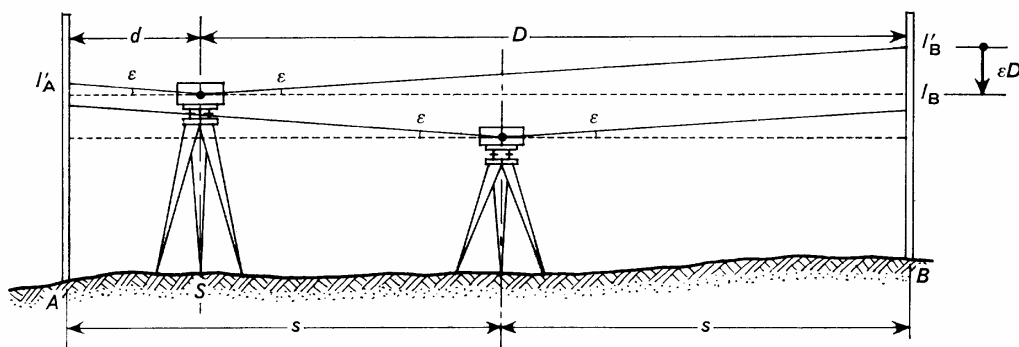


Fig. n. 1.11 - Rettifica di un livello

Poiché la condizione di rettifica è una condizione di parallelismo, si può agire o sul reticolo o sulla livella, ma è più semplice ruotare la tangente centrale. Dalla formula si ricava il valore di ε e, quindi, si può ricavare il valore ($d_A \cdot \text{tg} \varepsilon$) della stadia più vicina.

In tal modo, si corregge la lettura L'_A muovendo la vite di elevazione fino ad ottenere $L_A = L'_A - (d \cdot \text{tg} \varepsilon)$.

A questo punto si centra la livella con la sua “vite di rettifica”.

Le letture sulle stadiie vengono effettuate a stima, per livelli meno precisi, ovvero con un sistema micrometrico per le misure di alta precisione, nel primo caso l'approssimazione consentita è di 1/10 dell'intervallo di graduazione (di solito ± 1 mm); nel secondo caso la precisione può raggiungere $\pm 0,1$ mm fino a distanze di 50 m.

Il micrometro più diffuso (fig. 1.12) consiste in una lastra pianparallela posta davanti all'obiettivo e girevole intorno ad un asse orizzontale, normale all'asse di collimazione. La rotazione della lastra

sposta in un piano verticale la linea di mira di quantità proporzionali alla rotazione stessa, mantenendola sempre in posizione orizzontale.

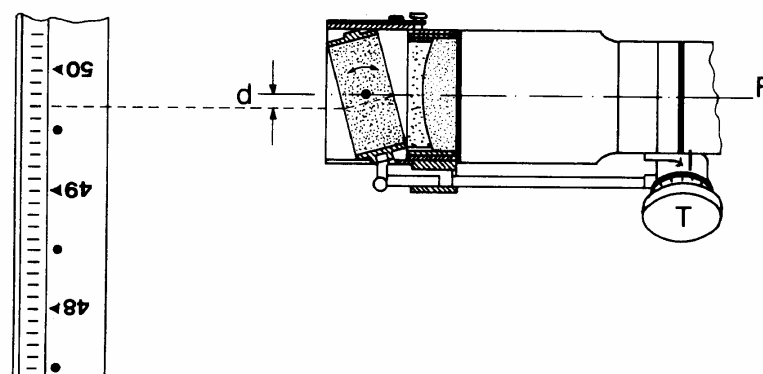


Fig. n. 1.12 - Micrometro a lastra piano - parallela

Sul bottone stesso che comanda la rotazione, o in una scaletta graduata visibile nell'interno del cannocchiale, si leggono le frazioni d'intervallo, dopo aver ruotato la lastra sino a collimare una "unità" della stadia, corrispondenti allo "spostamento" dell'asse di collimazione.

Per le livellazioni comuni si usano generalmente stadie di legno, lunghe 3 m, provviste di livella sferica; per le livellazioni di precisione si devono invece impiegare speciali stadie consistenti in una cassetta di legno o metallica, a cui è opportunamente fissato un nastro di "invar" su cui sono tracciate una o due graduazioni, con tratti di spessore prefissato, in genere pari a 5 mm.

Quando si desidera la differenza di livello fra punti molto distanti, si ricorre alla cosiddetta **livellazione composta** (fig. 1.13).

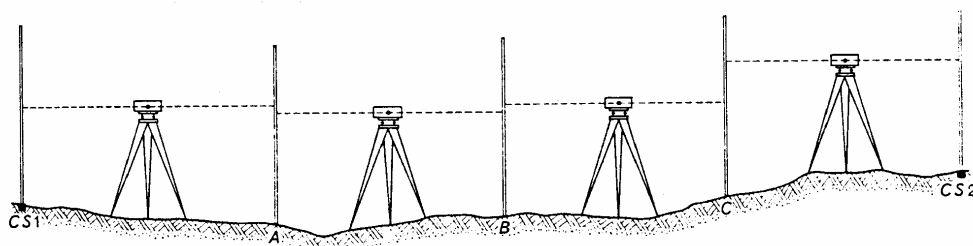


Fig. n. 1.13 - Schema operativo di livellazione “composta”.

Il percorso che porta da A a B del tratto da livellare, non necessariamente rettilineo, si divide in tanti tratti di 50÷100 m e sui punti di divisione si pongono le staffe.

Il dislivello totale è dato da $Q_B - Q_A = \sum (L_i - L_a)$ (i = indietro; a = avanti).

Il possibile “affondamento” della stadia durante la rotazione (necessaria per effettuare le livellazioni parziali precedente e seguente), che determinerebbe una possibile e pericolosa propagazione degli errori, può essere evitato utilizzando come punto di appoggio una apposita “piastra” di ghisa (per controllo si possono eseguire due livellazioni composte, una in andata ed una in ritorno e ricavarne la media).

Gli errori di lettura sulla stadia hanno una espressione globale del tipo $\sigma = C \sqrt{\frac{u \cdot d}{i \cdot s}}$ dove C è una costante che dipende dal tipo di strumento, u l’unità di stadia, d la distanza dalla stadia, i l’ingrandimento del cannocchiale ed s la sensibilità della livella in secondi.

Gli errori (accidentali) di misura di un dislivello hanno quattro cause principali e cioè:

- errori di lettura con stima o puntamento (micrometro) della stadia;
- errori di centramento della livella (o di funzionamento del sistema compensatore);
- errori di verticalità delle stadie;
- errori di graduazione delle stadie.

Il terzo errore diventa del tutto trascurabile se la stadia è tenuta verticale con una livelletta sferica. Il quarto errore è assai piccolo, purché si usino stadie appropriate ed eventualmente controllate (previa taratura).

L'errore di lettura a stima di una frazione dell'unità di stadia dipende dall'ampiezza dell'unità stessa, dalla distanza della stadia, nonché dal potere separatore e dall'ingrandimento del cannocchiale. L'errore minimo che si può commettere con una ordinaria stadia centimetrata è di circa 0.3 mm: esso cresce proporzionalmente alla radice quadrata della distanza della stadia fino a 50-60 m, dopo di che aumenta più rapidamente, per arrivare a circa 1 mm per $d = 100$ m.

Molto inferiore (< 0.1 mm) è invece l'errore di puntamento, specialmente se eseguito su graduazioni a tratti, mediante un apposito reticolo a "cuneo" e su distanze di stadia < 50 m (dopo di che aumenta sensibilmente). (fig. 1.14)

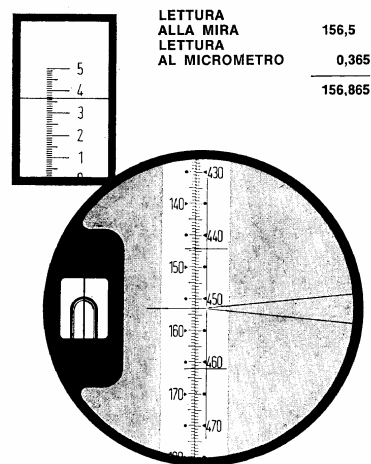


Fig. n. 1.14 - Campo dell'oculare di un livello di precisione.

L'errore dovuto al centramento della livella è proporzionale alla distanza d ; per $d = 50$ m va da 0.06 mm per i livelli di alta precisione a 0.5 mm per i livelli con livelle graduate.

In conclusione, si possono dare le seguenti cifre indicative per l'errore medio complessivo σ_0 di una lettura su una stadia posta a 50 m dallo strumento:

- livelli di alta precisione con micrometro: ± 0.1 mm
- livelli di precisione ± 0.2 mm
- livelli medi: ± 0.5 mm
- piccoli livelli da cantiere: ± 1 mm.

L'errore medio di una battuta si ottiene moltiplicando i numeri precedenti per $\sqrt{2}$, poiché il dislivello si ottiene dalla differenza di due letture.

Se ora consideriamo un'intera livellazione (composta da n battute), supponendo che l'errore medio di ogni singola battuta sia costante e uguale a σ_b , poiché la differenza totale di livello è uguale

alla somma dei livelli parziali $\Delta = \sum \Delta_i$, l'errore medio di tutta la livellazione è $\sigma_{\Delta}^2 = n\sigma^2 = 2n\sigma_b^2 = \frac{2D}{2d}\sigma_b^2 = \frac{D}{d}\sigma_b^2 =$ (d = distanza della stadia = distanza di battuta).

Se anche l'errore medio di una lettura, come si è visto, può ritenersi proporzionale a \sqrt{d} , e cioè della forma $\sigma_b = \mu\sqrt{d}$, (il che equivale a trascurare l'errore dovuto al centramento della bolla):

$$\sigma_{\Delta}^2 = \frac{D}{d} \mu^2 d = \mu^2 D \Rightarrow \sigma_{\Delta} = \mu \sqrt{D}$$

indipendentemente dalla lunghezza delle singole battute.

Poiché abbiamo considerato trascurabile l'errore dovuto alla livella, tale risultato richiede battute non eccessivamente lunghe.

1.3 - Quote Dinamiche

La trattazione vista finora consentirebbe la valutazione della precisione della livellazione se gli errori fossero puramente accidentali; ciò non accade effettivamente, perché sono stati messi in luce altri errori pseudo - sistematici (non dovuti alle cause già note).

Se si esegue una livellazione di una lunga sezione, di 20 o 30 km di lunghezza, prima in andata poi in ritorno, si constata che le differenze progressive (accumulate) tra i risultati delle due misure crescono con legge pressappoco lineare. Per lunghezze maggiori anche questi errori cambiano segno e si comportano come accidentali.

Si può calcolare, quindi, oltre all'errore chilometrico puramente accidentale, σ , un secondo errore chilometrico, che per lunghi tratti (30÷50 km) è anch'esso proporzionale a \sqrt{D} ; cosicché l'errore medio chilometrico totale di una lunga linea di livellazione è dato da: (fig. 1.15)

$$(1.5) \quad \tau^2 = \sigma^2 + \xi^2$$

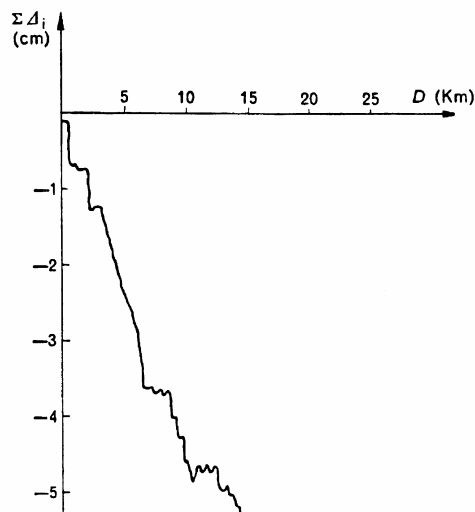


Fig. n. 1.15 - Diagramma delle differenze cumulate lungo una linea di livellazione.

Per un tratto di lunghezza $D \geq 30$ km l'errore medio sarà ancora della forma $\tau\sqrt{D}$, mentre per tronchi inferiori ξ è praticamente proporzionale alla lunghezza (trascurabile tuttavia fino a lunghezze di qualche chilometro). Il risultato di una livellazione geometrica, quindi, non fornisce il vero dislivello ortometrico a causa del non parallelismo delle superfici equipotenziali della gravità (fig. 1.16).

Così nella situazione in figura è:

$$dW = g_1 \cdot dh_1 = g_2 \cdot dh_2; \quad dh_1/dh_2 = g_2/g_1,$$

per cui se $g_1 \neq g_2$ anche $dh_1 \neq dh_2$.

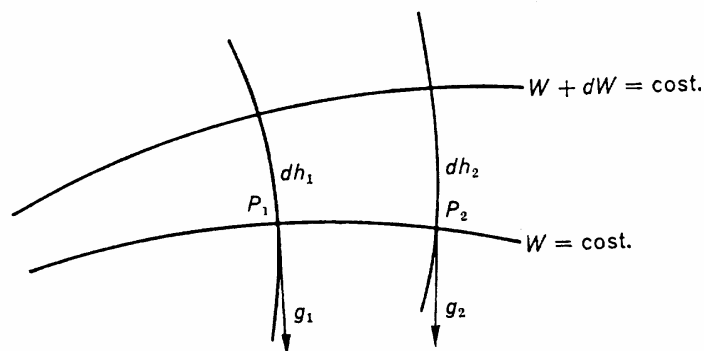


Fig. n. 1.16 - Non parallelismo delle superfici equipotenziiali della gravità.

D'altra parte, se supponiamo di eseguire una livellazione lungo una superficie equipotenziiale, siccome la verticale è in ogni punto normale a tale superficie, si ottiene un dislivello nullo.

La ragione analitica di questo fatto è la seguente: per due punti a distanza infinitesima vale la $dh = -\frac{dW}{g}$ e, quindi, il dislivello tra due

punti qualunque è espresso dalla relazione $h_2 - h_1 = \int_1^2 \frac{dW}{g}$ essendo

l'integrale esteso al cammino percorso (che potrebbe anche essere diverso). Ciò significa che eseguendo due livellazioni indipendenti tra due punti, si ottengono due risultati diversi se vengono eseguiti due percorsi diversi o, il che è lo stesso, che in un percorso chiuso si ottiene un dislivello totale diverso da zero, anche prescindendo dagli errori di misura (fig. 1.17).

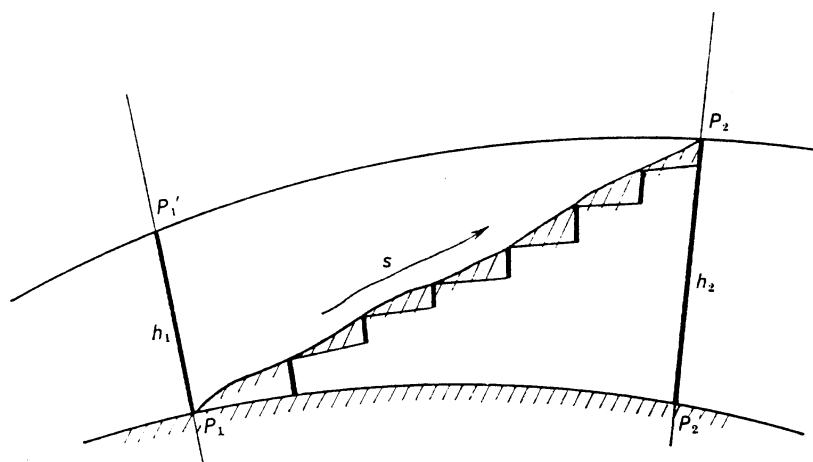


Fig. n. 1.17 - Dipendenza dal cammino percorso del risultato della livellazione geometrica fra due punti.

Di qui la necessità di introdurre le cosiddette “quote dinamiche”³, ovvero le “quote geopotenziali”, che tengono conto appunto della variazione di “ g ”, definite attraverso la differenza del potenziale di gravità tra il geoide ed il punto di cui si vuole definire la quota (calcolata lungo la verticale dal punto P al geoide).

2 - Livellazione trigonometrica

³ La quota “dinamica” di un punto è il lavoro che l’unità di massa compie nell’andare dal geoide al punto, pertanto ha le dimensioni di un lavoro diviso una massa. Per ottenere valori delle quote dinamiche abbastanza prossimi ai valori delle quote ortometriche, si è deciso di dividere le quote dinamiche per il valore medio della gravità che è pari a 980 gal (1 gal = 1 cm/sec²), o in maniera più semplice per 1000; si deducono così le “quote geopotenziali”, mediamente inferiori alle quote ortometriche del 2%.

È basata sull'impiego del teodolite, mediante misure di distanze zenitali. (fig. 2.1)

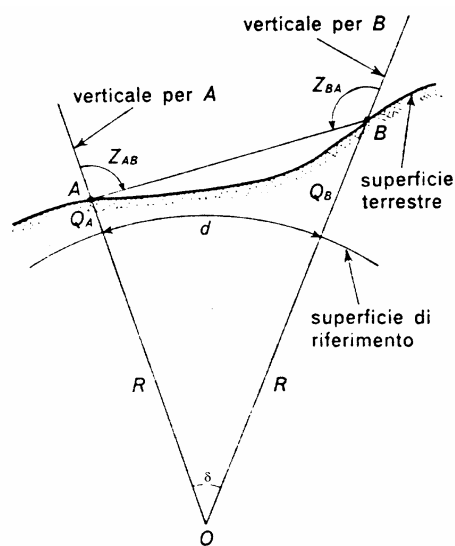


Fig. n. 2.1 - Schema della livellazione trigonometrica.

Viene applicata per distanze qualsiasi, fino a 10 km, richiede che sia nota la distanza tra i punti dei quali si misura il dislivello; poiché tale distanza non supera i 15 km, i calcoli possono essere sempre eseguiti sulla sfera locale.

Nella situazione in figura, dalla trigonometria si ricava l'espressione:

$$(2.1) \quad Q_B - Q_A = d \cdot \left(1 + \frac{Q_m}{R}\right) \cdot \operatorname{tg} \frac{1}{2}(z_B - z_A)$$

dove per Q_m , che rappresenta la quota media dei due punti, si può assumere un valore approssimato, essendo il rapporto Q_m/R assai piccolo (tale termine può essere spesso trascurato).

Se si misura una sola distanza zenitale, ponendo $z_B = \pi + \delta - z_A$ si ottiene:

$$\operatorname{tg} \frac{1}{2}(z_B - z_A) = \operatorname{tg} \frac{1}{2}(\pi + \delta - 2z_A) = \operatorname{tg} \left[\frac{\pi}{2} - (z_A - \frac{\delta}{2}) \right] = \cot g(z_A - \frac{\delta}{2}) =$$

$$\cot g(z_A - \frac{d}{2R}),$$

$$\text{da cui:} \quad Q_B - Q_A = d \cdot \left(1 + \frac{Q_m}{R}\right) \cdot \cot g\left(z_A - \frac{d}{2R}\right).$$

In effetti, non è possibile misurare le vere distanze zenitali, perché le traiettorie ottiche che congiungono i due punti non sono rette, a causa della “rifrazione atmosferica”.

La densità dell’aria diminuisce con la quota e quindi diminuisce anche l’indice di rifrazione; i raggi luminosi, attraversando zone di differente densità, subiscono delle rifrazioni (verso il basso se dirette verso l’alto e viceversa) e quindi si incurvano (fig. 2.2).

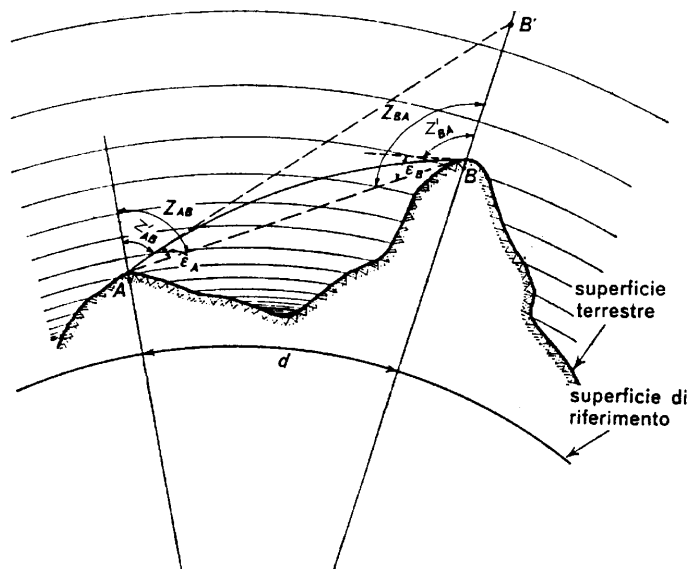


Fig. n. 2.2 - Influenza della rifrazione atmosferica nella misura delle distanze zenitali.

Se teniamo conto di questo errore sistematico, l'asse di collimazione del teodolite posto in A si dispone secondo la tangente in A alla curva di rifrazione, e quindi si ottiene:

$z_A = z'_A + \epsilon_A$, come pure $z_B = z'_B + \epsilon_B$, (collimando B da A), dove possiamo ritenere $\epsilon_A = \frac{1}{2} K_A \delta$ e $\epsilon_B = \frac{1}{2} K_B \delta$.

K_A e K_B dipendono sia dalle posizioni dei punti A e B che dalle condizioni fisiche (istantanee) dell'atmosfera, e vengono denominati "coefficienti di rifrazione" (fig. 2.3).

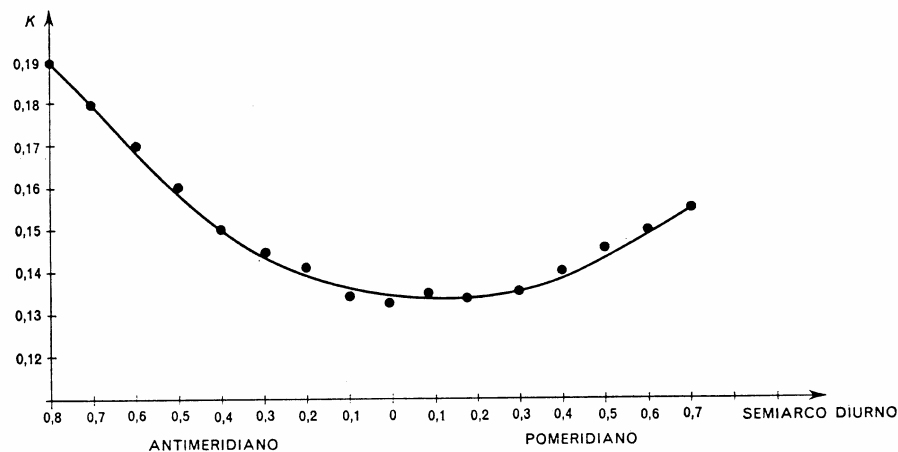


Fig. n. 2.3 - Variazione diurna del coefficiente di rifrazione (determinazione sperimentale)

Essi in genere hanno valori massimi la mattina, decrescono per raggiungere un minimo nelle ore pomeridiane, durante le quali si mantengono quasi stazionari, per poi risalire verso il tramonto senza però raggiungere i valori del mattino.

Se si misurano contemporaneamente z'_A e z'_B tra due punti non troppo distanti, si può ritenere $K_A = K_B \Rightarrow \epsilon_A = \epsilon_B$, dopo di che la formula (1) continua a sussistere con le letture z'_A e z'_B .

Per contro, tale metodo richiede due osservatori, muniti ciascuno di un teodolite, che operino contemporaneamente⁴.

Se invece vogliamo tener conto dell'errore sistematico, inserendolo nei calcoli (senza eliminarlo operativamente), bisogna modificare la (2) ponendo:

⁴ E' possibile, in tal caso, calcolare il coefficiente di rifrazione considerando il triangolo $A\hat{O}B$: $[\pi - (z'_A + \epsilon_A)] + \delta + \pi - (z'_B + \epsilon_B) = \pi$, da cui supponendo

$$k_A = k_B = k \text{ si ha: } k = 1 - \frac{z'_A + z'_B - \pi}{d} \cdot R.$$

$$z_A = z'_A + \varepsilon_A = z'_A + K \frac{d}{2R}; z_A - \frac{d}{2R} = z'_A + K \frac{d}{2R} - \frac{d}{2R} = z'_A - \frac{1-K}{2R} d.$$

Osservando poi che il termine $\frac{1-K}{2R} d$ è molto piccolo, si possono sostituire alla cotangente i primi due termini dello sviluppo in serie di Taylor:

$$\cot g\left(z'_A - \frac{1-K}{2R} d\right) \approx \cot g z'_A + \frac{1}{\sin^2 z'_A} \frac{1-K}{2R} d \approx \cot g z'_A + \frac{1-K}{2R} d.$$

La (2.1) diviene in definitiva:

$$(2.2) \quad Q_A - Q_B = d \cdot \left(1 + \frac{Q_m}{R}\right) \cdot \cot g z'_A + \frac{1-K}{2R} d^2$$

potendosi trascurare $\frac{Q_m}{R}$ nel secondo termine, che ha già valore più piccolo rispetto al primo.

2.1 - Precisione conseguibile nella livellazione trigonometrica

Applicando la legge di propagazione delle varianze, si ha:

$$\sigma_{\Delta}^2 = \left(\frac{\partial f}{\partial d} \right)^2 \sigma_d^2 + \left(\frac{\partial f}{\partial z} \right)^2 \sigma_z^2 + \left(\frac{\partial f}{\partial k} \right)^2 \sigma_k^2 \approx$$

$$\approx \left(\cot g z'_A + \frac{1-K}{R} d \right)^2 \sigma_d^2 + \left(\frac{d}{\sin^2 z'_A} \right)^2 \sigma_z^2 + \left(-\frac{d^2}{2R} \right) \sigma_k^2$$

trascurando $\frac{Q_m}{R}$. L'influenza di σ_d non è molto preoccupante (di ordine inferiore all'errore totale della livellazione), purché sia Δ notevolmente minore di d ; σ_k influisce su σ_{Δ} proporzionalmente al quadrato della distanza ed è trascurabile solo fino a qualche chilometro (per $d = 10$ km è dell'ordine di 8 cm; per $d = 20$ km è di circa 32 cm), dopo di che cresce notevolmente e può divenire predominante; l'influenza della rifrazione, tuttavia, si può eliminare con il metodo delle distanze zenitali reciproche e contemporanee.

Gli errori dovuti alla misura delle distanze zenitali sono in genere predominanti sugli altri, per distanze ≤ 10 km e, quindi, per raggiungere una precisione discreta su distanze di alcuni chilometri occorrono teodoliti di notevole precisione.

Per distanze inferiori ai 10 km, considerando trascurabili le influenze di σ_d e di σ_k , si vede che gli errori nella livellazione trigonometrica crescono proporzionalmente alla distanza ($\sigma_{\Delta} \cong 1 \div 2 \text{ cm/km}$), al contrario della livellazione geometrica dove crescono assai più lentamente ($\mu \sqrt{D}$).

3 - Livellazione distanziometrica

Si applica per punti distanti fino a qualche km e differisce dalla precedente solo perché la distanza viene determinata insieme al dislivello mediante un distanziometro elettronico.

E' necessario, inoltre, misurare oltre che la distanza zenitale anche l'altezza strumentale h_s e l'altezza del prisma h_p : (fig. 3.1)

$$(3.1) \quad Q_B - Q_A = h_s - h_p + d \cot g z$$

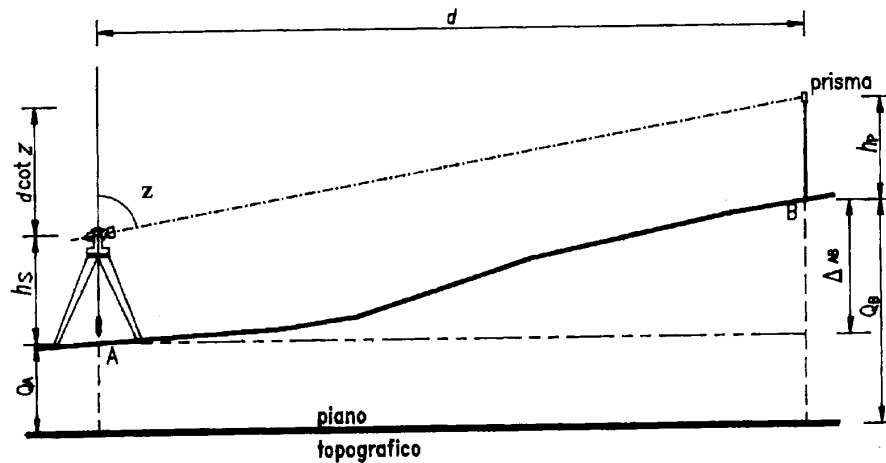


Fig. n. 3.1 - Livellazione distanziometrica

Per quanto riguarda la precisione, dalla relazione già vista si ricava (considerando h_s h_p):

$$\begin{aligned} \sigma_{\Delta}^2 &= \left(\frac{\partial f}{\partial d} \right)^2 \sigma_d^2 + \left(\frac{\partial f}{\partial z} \right)^2 \sigma_z^2 = \\ &= \cot^2 z \sigma_d^2 + \left(\frac{d}{\sin^2 z} \right)^2 \sigma_z^2 = \left[\left(\frac{\cot g \cdot z}{d} \right)^2 \sigma_d^2 + \frac{\sigma_z^2}{\sin^4 z} \right] d^2 \end{aligned}$$

rapidamente crescente con la distanza e con l'inclinazione delle visuali.

Per ottenere ottime precisioni (minori del cm/Km), anche con visuali molto inclinate, bisogna ricorrere a distanziometri di precisione e sistemi di centramento dello strumento e del prisma.

CAPITOLO X

METODI DI RILIEVO

TOPOGRAFICO

Capitolo X

METODI DI RILIEVO TOPOGRAFICO

1- Generalità sulle reti dei punti d'appoggio

Nel rilievo topografico si devono determinare le posizioni di un certo numero di punti, sempre elevato, ma molto variabile, per giungere poi alla rappresentazione del terreno per mezzo di una carta. I criteri operativi possono differenziarsi in modo notevole, pur conservando una certa unità di impostazione, in funzione di due elementi: l'estensione del rilievo e la scala della carta (elementi che non sono del tutto indipendenti tra loro).

La scala della carta influisce in due modi: sulla precisione con cui devono essere determinati i punti e sul dettaglio del rilievo affinché esso sia idoneo ad ottenere una rappresentazione sufficientemente fedele del terreno (si ricordi che un errore di posizionamento di un punto che sia ridotto in scala, inferiore all'errore di graficismo pari a ± 0.2 mm, si considera trascurabile).

In ogni caso, il rilievo è diviso in due parti nettamente distinte. La determinazione di una rete di appoggio ed il rilievo dei punti di dettaglio.

La rete di appoggio costituisce l'ossatura generale del rilievo cui si riferiscono, in seguito, le operazioni di dettaglio. Queste ultime possono essere eseguite con due procedimenti fondamentali differenti: quello topografico, effettuato con misure sul terreno (per rilievi di

precisione o grandi scale), quello fotogrammetrico, nel quale le misure vengono eseguite su immagini fotogrammetriche del terreno stesso (rilievi a media e piccola scala) che, comunque richiedono sempre un certo numero di punti direttamente rilevati sul terreno.

Quando il numero dei punti di appoggio è assai limitato (perché è piccola l'estensione del rilievo) la rete che li collega è assai semplice e può ridursi ad un poligono chiuso. Quando invece lo stesso numero è molto elevato, la rete diventa sempre più complessa e consta di diversi ordini, che infittiscono la rete fondamentale fino a raggiungere la densità desiderata.

La rete fondamentale investe nel suo insieme tutto il territorio da rilevare, mentre quelle di ordine inferiore hanno via via carattere più localizzato. Ciò significa che, mentre un qualunque errore di misura nella rete del 1° ordine influisce in qualche modo su tutta la rete, l'influenza degli errori nelle reti di ordine inferiore è più localizzato e quindi meno pericolosa.

Deriva da ciò che, mentre la rete fondamentale va rilevata con grande precisione, le misure relative a quelle degli altri ordini possono essere eseguite con precisione decrescente, senza che per questo la precisione dei punti diminuisca sensibilmente. I calcoli risultano assai facilitati, poiché solo la rete del 1° ordine dovrà essere compensata rigorosamente e calcolata nel suo insieme, mentre per gli altri punti saranno sufficienti compensazioni e calcoli locali.

2 - Reti trigonometriche

Ormai tutte le Nazioni dispongono di una rete fondamentale nazionale, con punti posti a meno di una decina di km di distanza l'uno dall'altro, rilevata con la massima precisione possibile; ed anche reti di infittimento che, almeno per i primi ordini, hanno precisione sufficiente per qualunque rilievo e che sono l'ossatura portante per

ogni rilievo topografico e tecnico¹.

Gli schemi geometrici dei collegamenti planimetrici di una serie di punti del terreno sono essenzialmente due. Il più semplice consiste in un poligono o in una serie di poligoni che congiungono tutti i punti; ciascun poligono è determinato planimetricamente se si misurano gli angoli nei vertici e le lunghezze dei lati. (fig. 2.1)

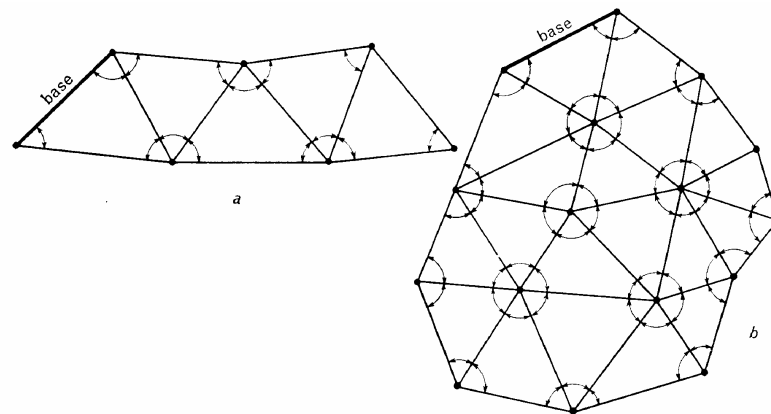


Fig. n. 2.1 - Schemi di triangolazione

Il secondo collegamento è realizzato invece per mezzo di triangoli, che hanno due a due un lato in comune (triangolazione); per determinare la rete è sufficiente misurare due elementi per ciascun triangolo (due angoli, due lati, un lato ed un angolo), ad eccezione di un triangolo in cui gli elementi misurati devono essere tre, di cui almeno un lato. In particolare, tutta la rete può essere determinata misurando un solo lato e almeno due angoli per ciascun triangolo.

Le “triangolazioni” sono state introdotte tre secoli fa proprio per utilizzare al massimo le misure angolari ed evitare quelle di distanze,

¹ Si noti la necessità di avere tutti i punti riferiti ad uno stesso sistema di riferimento, mentre in Italia, purtroppo, coesistono diversi sistemi (Gauss-Boaga, Cassini-Soldner, UTM).

allora estremamente delicate e lunghe. Oggi invece si è accertato il grande vantaggio in precisione che si ottiene misurando molti lati, tanto che in alcuni casi si arriva a misurarli tutti, effettuando quella che si chiama una “trilaterazione”. La rete trilaterata non avrebbe bisogno di misure angolari che, tuttavia, si effettuano per ottenere dati di misura largamente sovrabbondanti (compensazione rigorosa).

Le triangolazioni (o trilaterazioni) costituiscono una struttura assai più rigida delle poligonali e, se bene studiate e conformate, limitano considerevolmente la propagazione degli errori che, comunque, permangono per cui nelle grandi reti, oltre a misurare molti lati, è opportuno misurare anche un certo numero di azimut astronomici.

Nelle reti di appoggio è indispensabile in ogni caso misurare grandezze sovrabbondanti e procedere ad una compensazione rigorosa delle misure, ottenendo il duplice scopo di aumentare la precisione delle coordinate dei punti e di poter **calcolare e rappresentare l'ellisse di errore di ogni vertice**.

Le poligonali si trovano in svantaggio sotto questo punto di vista, poiché si vede facilmente che in un poligono chiuso in cui si misurano tutti i lati e tutti gli angoli, le misure esuberanti sono solo tre, indipendenti dal numero dei vertici.

Le poligonali hanno tuttavia, insieme agli inconvenienti enunciati, il grande vantaggio pratico, rispetto alle triangolazioni, di adattarsi meglio alle forme di terreno, richiedendo solo che da ogni vertice siano visibili quelli adiacenti.

3 - La rete nazionale - Triangolazioni

Il principio della triangolazione, introdotto dall'olandese Snellius, è il seguente: scelti i punti da collegare per determinare la posizione reciproca, si pensa di congiungerli a due a due in modo da costituire

una serie di triangoli aventi successivamente un lato in comune. Si misura un lato di un triangolo (base della triangolazione) e possibilmente tutti gli angoli interni dei triangoli. In questo modo, nota la posizione di un vertice di partenza e la direzione (ossia l'azimut) di un lato uscente da esso, risultano definite le posizioni di tutti i punti, i quali prendono il nome di “**vertici trigonometrici**”.

I due schemi di triangolazione più usati sono detti a catena e/o a rete; la differenza consiste nel fatto che, mentre in una catena il passaggio da un lato ad un altro successivo qualunque può avvenire per una sola strada, nella rete il passaggio può avvenire per più strade, cosicché esistono possibilità di controllo e di compensazione dovute alla conformazione stessa della rete (che comunque non devono mancare, anche per le catene). La triangolazione italiana del I ordine è a rete, con punti a distanza (e 24 basi) di 30-40 km; su di essa è stata eseguita una rigorosa compensazione di insieme, con elevata precisione (gli angoli sono stati misurati effettuando 24 reiterazioni) (fig. 3.1).

Alla rete del I ordine si collega quella del II ancora di pochi punti, situati, nello schema a rete, ciascuno all'incirca nel baricentro dei triangoli di I ordine e collegati ai vertici di questi ultimi ed agli altri vertici del II ordine con misure angolari un po' meno precise delle precedenti (12 reiterazioni); la compensazione delle reti del II° ordine è locale e si riferisce ai vertici di primo ordine supposti esatti.

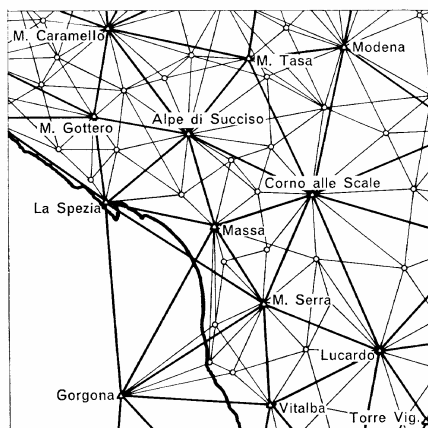


Fig. n. 3.1 - Stralcio della triangolazione italiana del 1° e 2° ordine.

Le triangolazioni di III (6 reiterazioni) e IV ordine (3 reiterazioni) servono ad intercalare nella rete di precisione sempre più numerosi punti di appoggio utili per le operazioni di rilievo topografico.

La precisione è sensibilmente inferiore alle precedenti, poiché i punti sono determinati appoggiandosi ai vertici di ordine superiore, misurando gli angoli con teodoliti di media grandezza e con meno cura. Per i vertici del III ordine si compensa ancora una triangolazione vera e propria; per quelli del IV ordine si adottano invece procedimenti di intersezione multipla in avanti ed anche all'indietro (compensazione per intersezione e/o a punti isolati).

All'interno di tali reti possono essere costruite le ulteriori poligonalali o triangolazioni per esigenze tecniche e topografiche fino a giungere ai dettagli (2 o 1 sola misura).

La scelta dei vertici si basa su alcuni criteri fondamentali: distanza reciproca il più grande possibile, nei limiti della visibilità consentita (30-40 km, riducibili a 10-20 km per i punti del II ordine); forma dei triangoli preferibilmente equilateri (il che comporta una minimizzazione dell'ellissi d'errore di ogni punto); posizionamento in

punti dominanti (perché sia possibile il loro collegamento col maggior numero di vertici vicini ed in modo che costituiscano ottimi punti di “**collimazione**”), possibilmente protetti (la distruzione dei segnali comporta una grave menomazione di tutta la triangolazione).

I vertici del I e del II ordine hanno in maggioranza segnali propri: raramente sono posti su edifici, per ottenere un maggior campo di visibilità ed una più grande stabilità.

Nella triangolazione fondamentale italiana, il segnale è costituito da un pilastrino con un centrino di fondo ed uno di superficie che individua effettivamente il vertice e che sovente si prolunga nel pilastrino. (fig. 3.2)

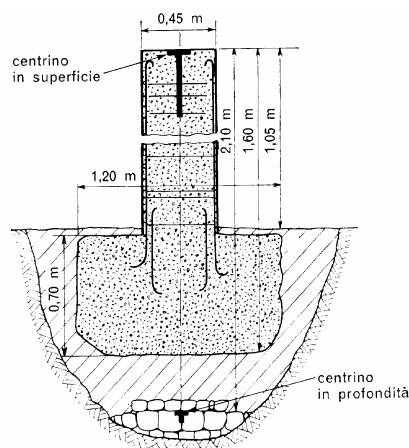


Fig. n. 3.2 - Materializzazione di un vertice di 1° e 2° ordine della rete geodetica italiana.

L'I.G.M. fornisce a richiesta le monografie dei vertici trigonometrici fondamentali² con le seguenti indicazioni: posizione del punto (accurata descrizione) e suo disegno (in corrispondenza di una certo punto di vista); coordinate N-E (al centimetro) riferite al

² Da richiedersi indicando il nome del punto e individuando la sua posizione nella carta 1:25000 (es. Laterza, foglio n.177, II SO).

sistema Gauss-Boaga; coordinate geografiche ellissoidiche (φ, λ); quota al “piano di paragone” (p-p) segnato nel disegno (al cm), che spesso deve però essere rapportata al suolo, con notevole diminuzione della precisione³ altimetrica (fig. 3.3).

056129	Comando Regione Militare Nord-Ovest-Torino										
N°	Staz. CC. S. Benigno Canavese										
N°	Comune S. Benigno Canavese Prov. Torino										
056129	Proprietario										
mod. 281/GAR I.G.M.	MONOGRAFIA	(1954) Belvedere della cascina Asse geometrico del belvedere									
		Coordinate geografiche									
		<table border="1"> <tr> <td>φ</td> <td>ω</td> </tr> <tr> <td>45° 13' 30'', 919</td> <td>4° 41' 39'', 744</td> </tr> </table>		φ	ω	45° 13' 30'', 919	4° 41' 39'', 744				
		φ	ω								
		45° 13' 30'', 919	4° 41' 39'', 744								
Coordinate Gauss-Boaga											
<table border="1"> <tr> <td>N</td> <td>E</td> </tr> <tr> <td>5 008 817,781</td> <td>402' 488,09</td> </tr> </table>		N	E	5 008 817,781	402' 488,09						
N	E										
5 008 817,781	402' 488,09										
N°	Quota al PP H = 228,43										
056129	PP = Gronda tetto belvedere										
056129	<table border="1"> <tr> <td>TRIG. Nome</td> <td>Ord.</td> <td>F°</td> <td>N°</td> </tr> <tr> <td>CASCINA MORENTONE</td> <td>IV</td> <td>056</td> <td>129</td> </tr> </table>			TRIG. Nome	Ord.	F°	N°	CASCINA MORENTONE	IV	056	129
TRIG. Nome	Ord.	F°	N°								
CASCINA MORENTONE	IV	056	129								

056129
N°
056129
mod. 281/GAR I.G.M.
N°
056129
N°
056129
N°
056129

(1954) Belvedere della cascina
Asse geometrico del belvedere

Coordinate geografiche

φ	ω
45° 13' 30'', 919	4° 41' 39'', 744

 Coordinate Gauss-Boaga

N	E
5 008 817,781	402' 488,09

Nome Cascina Morentone
N° = 056129
O = IV

Fig. n. 3.3 - Monografia di un vertice di 4° ordine della rete geodetica italiana.

La compensazione di una rete si basa essenzialmente su vari tipi di equazioni:

- equazioni angolari, ad es. la verifica degli angoli interni di un triangolo:

$$\alpha + \beta + \gamma - (\pi + \epsilon) = 0$$

³ Si noti la differenza dai “capisaldi” di livellazione che non solo individuano con esattezza il “punto” di appoggio della stadia ma ne riportano la quota con l’accuratezza di $\pm 0.1\text{mm}$.

(in genere si otterrà $\alpha' + \beta' + \gamma' - (\pi + \epsilon) \neq 0$ e bisognerà determinare le correzioni $\alpha_0, \beta_0, \gamma_0$, da apportare alle misure:

$$\alpha = \alpha' + \alpha_0; \quad \beta = \beta' + \beta_0; \quad \gamma = \gamma' + \gamma_0.$$

- l'equazione d'orizzonte, riferita a tutti gli angoli facenti capo ad uno stesso vertice: $\sum \alpha_i - 2\pi = 0$
- l'equazione poligonale, riferita alla somma degli angoli interni di un poligono di n lati: $\sum \alpha_i - (n-2) \pi = 0$
- equazioni laterali, che impongono l'uguaglianza tra gruppi di angoli necessari per ricavare più misure dello stesso lato, ottenute seguendo percorsi diversi partendo dagli elementi noti (basi). Tali percorsi diversi possono scegliersi con riferimento ad una rete di appoggio o con riferimento a diverse basi misurate o calcolate (se si utilizza il metodo dello sviluppo di una base). (fig. 3.4)

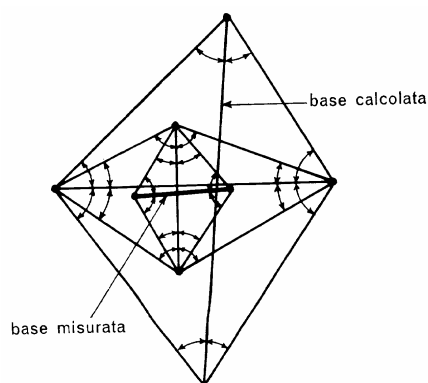


Fig. n. 3.4 - Rete di triangolazione per lo sviluppo di una base misurata.

Si determina con precisione una base di appoggio, dalla quale ci si riallaccia a quella calcolata mediante una piccola triangolazione locale (rete di sviluppo), eseguita però con grandissima precisione (36 reiterazioni), con notevole numero di equazioni esuberanti, in modo da collegare rigidamente la base misurata con quella calcolata.

- equazioni alle basi: il valore di una base si può ricavare dal calcolo mediante un'altra base:

$$(3.1) \quad \frac{\sin \alpha_1 \sin \alpha_2 \sin \alpha_3 \dots \sin \alpha_n}{\sin \beta_1 \sin \beta_2 \sin \beta_3 \dots \sin \beta_n} - \frac{b_1}{b_2} = 0$$

Riassumendo, una triangolazione comporta le seguenti operazioni:

- 1) Scelta e posizionamento dei vertici;
- 2) Misure di lati e di angoli della rete;
- 3) Compensazione rigorosa delle misure;
- 4) Determinazione delle coordinate piane (x, y) dei vertici e delle loro ellissi di errore;
- 5) Trasformazione delle coordinate (x, y) in coordinate Gauss-Boaga (N, E) (da effettuarsi tenendo conto del coefficiente di deformazione lineare nella zona considerata).

Per le triangolazioni tecniche, con i lati di qualche km, si usano procedimenti analoghi a quelli già visti, ma si semplificano notevolmente i metodi di rilievo e di calcolo, perché diminuiscono il numero dei vertici, le dimensioni dei triangoli e la propagazione degli errori. Le misure di distanza possono effettuarsi con distanziometri ad onde ovvero con il metodo G.P.S., mentre per gli angoli sono sufficienti 3-4 reiterazioni. La compensazione si effettua, solitamente, con metodi rigorosi utilizzando i numerosi programmi di calcolo esistenti in commercio.

4 - Metodi di intersezione

Si è già detto che per rilevare alcuni dei vertici trigonometrici del IV ordine l'IGM ha fatto ricorso a metodi di intersezione, sempre appoggiandosi a vertici trigonometrici noti. Anche in altre occasioni può presentarsi la necessità di determinare la posizione di punti isolati, per esempio quando per eseguire un rilievo di dettaglio non sono sufficienti i vertici trigonometrici disponibili nella zona.

a) Intersezione in avanti

Si applica quando si conosce la posizione planimetrica di due punti e si vuole determinare quella di un terzo punto, misurando gli angoli orizzontali formati dalla congiungente i punti noti con le direzioni che vanno al punto incognito.

Graficamente il punto P risulta definito dall'intersezione delle due direzioni formanti gli angoli α e β con AB, condotte rispettivamente da A e da B "in avanti" verso P (da cui il nome) (fig. 4.1).

Per il calcolo, si valutano anzitutto gli angoli di direzione⁴ (AB), (AP), (BP).

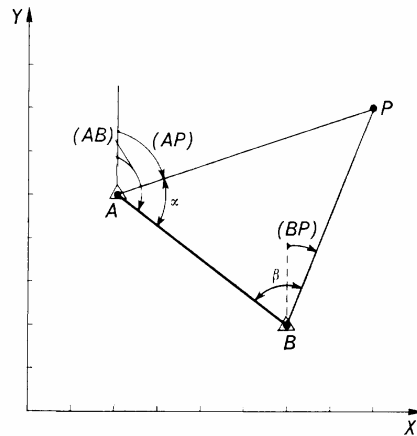


Fig. n. 4.1 - Intersezione in avanti.

$$\text{Posto}^5: \operatorname{tg}(AB) = \frac{[X_B - X_A]}{[Y_B - Y_A]}$$

⁴ Angolo di direzione di una retta (segmento orientato) è l'angolo destrorso di cui deve ruotare la direzione positiva dell'asse delle ordinate per sovrapporsi alla direzione positiva della retta (segmento).

⁵ Si noti che per l'indeterminazione della funzione tangente è necessario stabilire a quale quadrante appartiene il segmento AB e tenerne conto nei calcoli.

(l'angolo (AB) viene univocamente determinato considerando le posizioni reciproche dei punti A e B).

Essendo $(AP) = (AB) - \alpha$ e $(BP) = (AB) + \beta + \pi$ si ottiene il sistema:

$$\begin{cases} X_P - X_A = (Y_P - Y_A) \operatorname{tg}(AP) \\ X_P - X_B = (Y_P - Y_B) \operatorname{tg}(BP) \end{cases}$$

che si risolve rispettivamente alle incognite X_P, Y_P .

Ricavando AP e BP con il “teorema dei seni”, si può passare a formule più pratiche che consentono il calcolo delle coordinate di P in doppio modo per la verifica.

Si ha:

$$\begin{cases} X_P = X_A + AP \operatorname{sen}(AP) \\ Y_P = Y_A + AP \cos(AP) \end{cases}$$

oppure

$$\begin{cases} X_P = X_B + BP \operatorname{sen}(BP) \\ Y_P = Y_B + BP \cos(BP) \end{cases}$$

L'intersezione semplice in avanti così descritta non consente un controllo delle misure fatte, essendo queste in numero strettamente sufficiente perché il punto da rilevare sia univocamente definito⁶. Spesso perciò si ricorre all'*intersezione multipla in avanti*, che

⁶ Si noti che l'incertezza delle coordinate X_P e Y_P (ellisse di errore) dipenderà:

1. dalla incertezza delle misure angolari e, eventualmente, da quelle lineari;
2. dalla forma geometrica del triangolo ABP con semiasse maggiore dell'ellisse diretto lungo Y quando il rapporto base/distanza è piccolo e, viceversa, diretto lungo X quando tale rapporto è molto grande.

Ne consegue che la forma ottimale del triangolo è quella equilatera.

consiste nel determinare la posizione di un punto facendo stazione in più vertici P_1, P_2, \dots, P_n di coordinate note e misurando gli angoli $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$, come in figura 4.2.

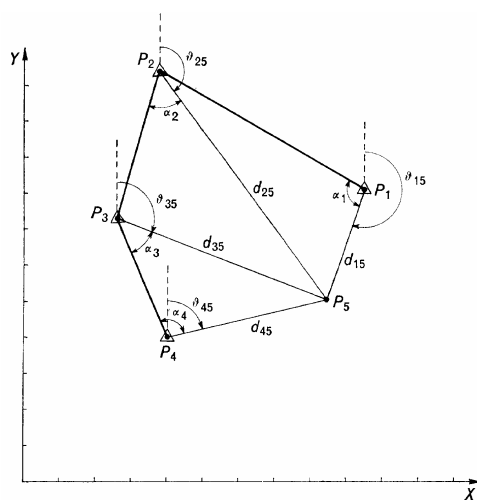


Fig. n. 4.2 - Intersezione multipla in avanti.

Siccome in tal modo le misure sono esuberanti, risulta possibile la compensazione rigorosa; per ottenere ulteriori misure sovrabbondanti possono, ovviamente, misurarsi anche i lati con distanziometri e ricondurre il calcolo di compensazione a quello già visto di una piccola rete trigonometrica.

b) Intersezione laterale

Si applica quando si conosce la posizione planimetrica di due punti e si vuole determinare quella di un terzo punto misurando gli angoli orizzontali nel punto incognito ed in uno dei punti noti. Il problema è del tutto analogo a quello dell'intersezione in avanti (misurati due angoli, resta definito il terzo) ed analoghi sono i procedimenti di calcolo.

c) Intersezione all'indietro o inversa (metodo di Pothenot)

Si applica quando si conosce la posizione planimetrica di tre punti e si vuole determinare la posizione di un 4° punto misurando da questo gli angoli formati dalle direzioni ai punti noti gli angoli si misurano dal punto incognito e non da quelli noti, da cui il nome di intersezione inversa (operazione che risulta più comoda potendo scegliere la più opportuna posizione della stazione ed utilizzare i vertici trigonometrici solo per le collimazioni). (fig. 4.3)

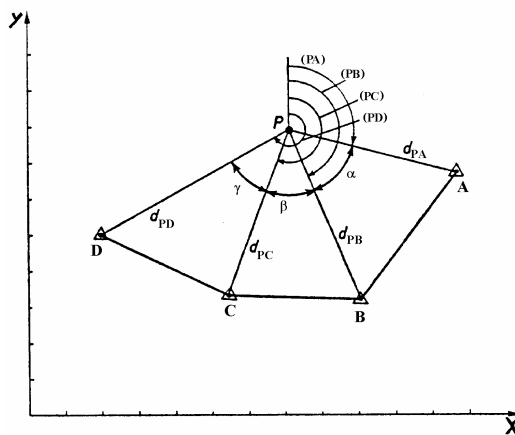


Fig. n. 4.3 - Intersezione inversa multipla.

Si considerano le relazioni:

$$tg(AP) = \frac{X_P - X_A}{Y_P - Y_A}; tg(BP) = \frac{X_P - X_B}{Y_P - Y_B}; tg(CP) = \frac{X_P - X_C}{Y_P - Y_C}; \text{ecc.}$$

nelle quali gli angoli di direzione (PB) e (PC) si possono esprimere in funzione di (PA):

1. $(PB) = (PA) + \alpha$
2. $(PC) = (PB) + \beta = (PA) + \alpha + \beta$;

sostituendo ed eliminando i denominatori si ottiene il sistema che si risolve rispettivamente alle incognite $tg(PA)$, X_P , Y_P :

$$[X_P - X_A = (Y_P - Y_A) tg(PA)]$$

$$\begin{cases} X_P - X_B = (Y_P - Y_B) \operatorname{tg} [(PA) + \alpha] \\ X_P - X_C = (Y_P - Y_C) \operatorname{tg} [(PA) + \alpha + \beta] \end{cases}$$

L'intersezione inversa semplice presenta dei casi di indeterminazione. Questo accade quando il punto P si trova sulla circonferenza(critica) passante per i tre punti noti A,B e C. Perché il problema sia determinato occorre allora che sia $\alpha + \beta + \gamma \neq \pi$. Anche nell'intersezione inversa si ricorre, spesso, a misure esuberanti di lati e/o di angoli per poter effettuare delle compensazioni (intersezione multipla).

Il cosiddetto metodo di Snellius ampliato è un'estensione dell'intersezione al caso in cui, nota la posizione di tre punti, si voglia determinare contemporaneamente quella di due o più punti (di cui sia nota la posizione reciproca rispetto ai punti noti), eseguendo da questi ultimi misure angolari.

d) Doppia intersezione inversa (Metodo di Hansen)

Questo metodo si applica quando si conosce la posizione planimetrica di due punti e si vuole determinare la posizione di due altri punti misurando da questi gli angoli formati dalle direzioni che vanno ai punti noti ed all'altro punto incognito. Come nei problemi precedenti, perché si abbia un'unica soluzione occorre che sia nota da quale parte si trovino i punti P_1 e P_2 rispetto ad AB. (fig. 4.4)

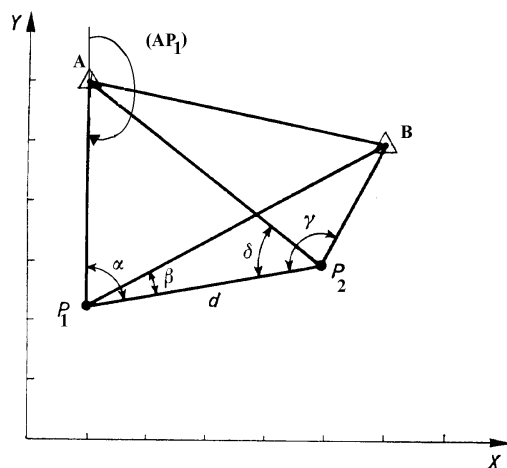


Fig. n. 4.4 - Problema di Hansen.

Si ha:

$$(BP_1) = (AP_1) + \alpha - \beta$$

$$(P_1P_2) = (AP_1) + \alpha - \pi$$

$$(AP_2) = (AP_1) + \alpha - \pi + \delta$$

$$(BP_2) = (AP_1) + \alpha - \beta + \gamma + \delta - \pi$$

Avendo espresso tutti gli angoli di direzione in funzione di (AP_1) e degli angoli misurati $\alpha, \beta, \gamma, \delta$ possiamo scrivere le relazioni:

$$X_{P1} - X_A = (Y_{P1} - Y_A) \operatorname{tg} (AP_1)$$

$$X_{P1} - X_B = (Y_{P1} - Y_B) \operatorname{tg} [(AP_1) + \alpha - \beta]$$

$$X_{P2} - X_{P1} = (Y_{P2} - Y_{P1}) \operatorname{tg} [(AP_1) + \alpha - \pi]$$

$$X_{P2} - X_A = (Y_{P2} - Y_A) \operatorname{tg} [(AP_1) + \alpha - \pi + \delta]$$

$$X_{P2} - X_B = (Y_{P2} - Y_B) \operatorname{tg} [(AP_1) + \alpha - 2\beta + \gamma + \delta - \pi]$$

nelle 5 incognite:

(AP_1)	X_{P1}	Y_{P1}	X_{P2}	Y_{P2}
1	2	3	4	5

Il metodo della doppia intersezione in avanti (detto anche problema della distanza inaccessibile), è analogo al metodo della

doppia intersezione inversa e si utilizza quando i punti P_1 e P_2 sono inaccessibili e gli angoli si misurano nei punti noti A e B .

Le posizioni dei due punti possono ricavarsi con un sistema analogo al precedente o separatamente, applicando due volte il metodo dell'intersezione in avanti.

Si vede subito che lo schema di quattro punti con le misure degli elementi noti (6 lati) e/o (angoli), dà luogo allo schema più semplice di rete trigonometrica (v. osservazioni indirette), per la determinazione non solo delle coordinate dei vertici ma anche, e soprattutto, delle loro incertezze (ellisse di errore).

5 - Poligonazioni

In una poligonale aperta, indicando con n il numero dei vertici, si misurano $n-1$ lati ed $n-2$ angoli; cioè si misurano $2n-3$ elementi, corrispondenti a quelli "strettamente sufficienti" per la determinazione della stessa poligonale. In una poligonale chiusa, che si riduce ad un poligono, se si misurano tutti gli angoli e tutti i lati, si hanno $2n$ elementi noti e quindi tre in più di quelli strettamente necessari e sufficienti, per cui diventa possibile eseguire una compensazione "speditiva" degli errori.

Anche per le poligonali aperte può essere possibile la compensazione speditiva; occorre a tal fine che siano note le coordinate del 1° e dell'ultimo vertice e che si misurino gli angoli che in questi punti il primo e l'ultimo lato della poligonale formano con le direzioni a due altri punti anch'essi di coordinate note. In questo caso la poligonale aperta si dice ad "estremi vincolati". (fig. 5.1)

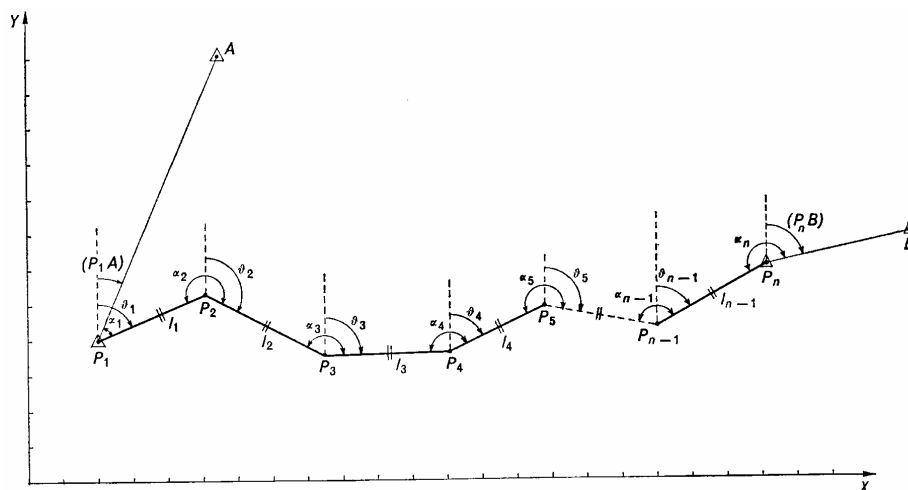


Fig. n. 5.1 - Poligonale aperta ad estremi vincolati.

Le poligonazioni possono essere topografiche o geodetiche, in relazione alla minore o maggiore lunghezza dei lati (le compensazioni delle poligonali geodetiche possono ottenersi con gli stessi metodi delle poligonali topografiche se si calcolano le coordinate dei vertici nella proiezione conforme di Gauss-Boaga). Naturalmente occorre tener conto che le misure si fanno sul terreno, cioè sull'ellissoide, mentre i calcoli si fanno sulla superficie deformata della proiezione conforme; occorre perciò apportare agli angoli ed ai lati misurati opportune correzioni in modo da avere gli stessi valori che si sarebbero ottenuti nel piano "gaussiano".

Per individuare punti di appoggio in un rilievo di non grande estensione si ricorre comunemente alle "poligonali chiuse", le quali possono essere "orientate", se si conoscono le coordinate di un vertice e l'azimut di un lato uscente da esso, oppure "non orientate", se sono riferite ad un sistema di assi cartesiani arbitrario, che può essere assunto con l'origine in un vertice e l'asse delle ascisse coincidente con

un lato. (fig. 5.2)⁷

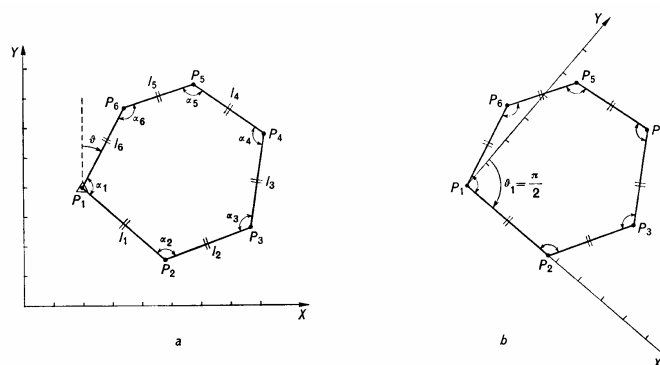


Fig. n. 5.2 - Poligonale chiusa (a = poligonale orientata; b = poligonale a sé stante).

5.1 - Rilevamento di una poligonale topografica

I vertici si stabiliscono su punti fissi e facilmente individuabili dai vertici precedente e seguente. I vertici vengono contraddistinti con una numerazione in ordine progressivo e di ciascuno di essi si esegue una monografia comprendente il particolare del vertice e le sue distanze da almeno tre punti circostanti (riferimenti).

Dell'intera poligonale, infine, si esegue l'“Eidotipo”(fig. 5.3), piccolo disegno planimetrico contenente le numerazioni dei punti, per definire l'ordine di successione dei lati e la posizione della poligonale rispetto a tutta la zona circostante.

⁷ L'operazione di “inquadramento” di una rete “locale” nel sistema di riferimento nazionale (G-B) richiede una rototraslazione con variazione di scala e, quindi, l'esecuzione di misure di collegamento alla rete nazionale con gli schemi già visti (intersezioni, reti, ecc.), per poter ricavare le coordinate di almeno due “punti doppi”, noti cioè, in entrambi i sistemi.

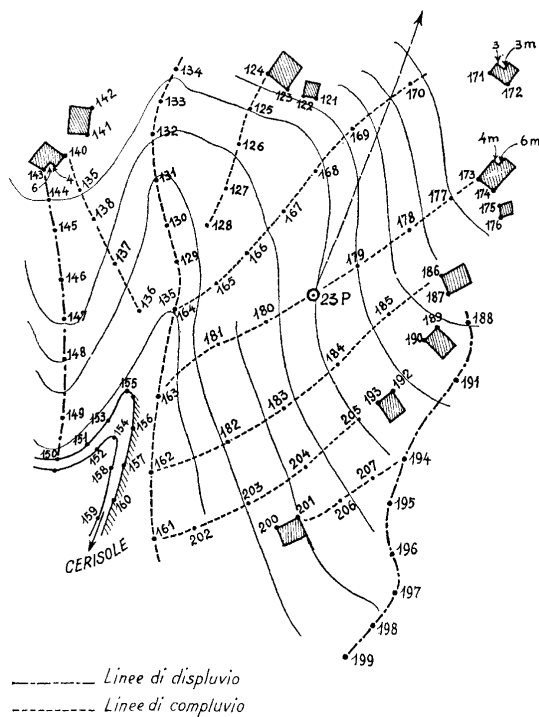


Fig. n. 5.3 - Eidotipo

I lati sono solitamente di lunghezze intermedie (da centinaia di metri a qualche km) e si misurano solitamente due volte, in andata e in ritorno; confrontate le due misure di uno stesso lato, se la loro differenza non è superiore alla tolleranza si assume come lunghezza del lato la media aritmetica delle misure eseguite. Per le misure angolari si utilizzano teodoliti elettronici dell'approssimazione di $0^{\circ}.1 \div 1^{\circ} \cong (3'' \div 30'')$ e in generale ciascun angolo si misura almeno due volte, nelle posizioni coniugate dello strumento, con l'applicazione della regola di Bessel.

6 - Calcolo, verifica e compensazione di una poligonale topografica aperta ad estremi vincolati

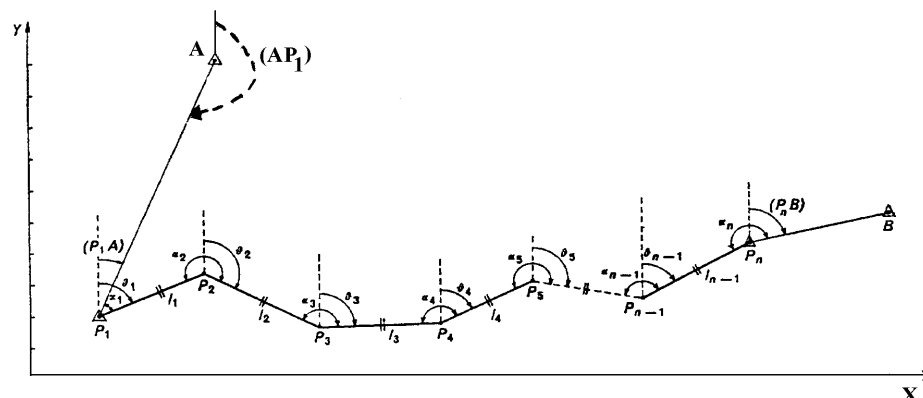


Fig. n. 6.1 - Poligonale aperta ad estremi vincolati

Dagli estremi della precedente poligonale siano visibili rispettivamente due altri punti trigonometrici A e B di coordinate note (x_A, y_A) e (x_B, y_B), siano stati misurati conseguentemente gli angoli $\alpha_1, \alpha_2, \dots, \alpha_n$ ed i lati l_1, l_2, \dots, l_n . Per mezzo delle coordinate di A e di P_1 si determina l'angolo di direzione (AP_1) (angolo di apertura):

$$(6.1) \quad \operatorname{tg}(AP_1) = \frac{X_{P_1} - X_A}{Y_{P_1} - Y_A} \quad \text{e} \quad \operatorname{tg}(P_n B) = \frac{X_B - X_n}{Y_B - Y_n}$$

Indi si calcolano gli angoli di direzione di tutti gli altri lati della poligonale mediante la formula per il trasporto:

$$\theta_i^I = \theta_{i-1}^I + \alpha_i \pm \pi$$

La verifica angolare si esegue notando che l'angolo θ_n^I , calcolato con la formula precedente per l'ultimo lato, deve essere uguale a quello (θ_n) che si ricava dalla (6.1);

dalla loro differenza si ricava:

$$\Delta_{\theta} = \theta_n^I - \theta_n$$

si effettua quindi la verifica:

$$\Delta_{\theta} \leq t_{\alpha}$$

In pratica, la tolleranza angolare si determina con la formula:

$$(6.2) \quad t_{\alpha} = c \sqrt{n}$$

dove il Catasto Italiano ha assunto per "c" i valori:

$$c \text{ variabile tra } 1^C \div 2.5^C.$$

Ovvero si può assumere anche $c = 3\sigma_{\alpha}$, utilizzando lo s.q.m. dello strumento.

Se risulta $\Delta\theta < t_{\alpha}$ si compensa ogni angolo misurato di $-\Delta\theta/n$. È preferibile però compensare direttamente gli angoli di direzione, nel qual caso la correzione va fatta proporzionalmente al numero d'ordine dei vertici, ponendo $\theta_i = \theta'_i - i \Delta\theta/n$.

E' evidente che dovrà risultare, infine, $\theta_n = \theta'_n \pm \Delta\theta = (P_n B)$, pari all'angolo ricavato dal calcolo.

Con gli angoli di direzione compensati si calcolano le coordinate cartografiche parziali X_i' e Y_i' , con le formule:

$$\begin{aligned} \{X_2' &= X_1 + l_1' \sin \theta_1 ; \\ \{X_3' &= X_2 + l_2' \sin \theta_2 ; \dots \{X_n' = X_{n-1} + l_{n-1}' \sin \theta_{n-1} \\ \{Y_2 &= Y_1 + l_1' \cos \theta_1 ; \dots \{Y_n = Y_{n-1} + l_{n-1}' \cos \theta_{n-1} . \end{aligned}$$

La verifica laterale si esegue notando che deve essere:

$$\sum_{i=1}^n l_i' \sin \theta_i = X_n - X_1 \quad \text{ed anche} \quad \sum_{i=1}^n l_i' \cos \theta_i = Y_n - Y_1 .$$

Diversamente si hanno gli errori di chiusura in ascissa e ordinata, dati rispettivamente da:

$$\Delta_x = (X_n - X_1) - \sum_{i=1}^n l_i' \sin \theta_i \quad \text{e} \quad \Delta_y = (Y_n - Y_1) - \sum_{i=1}^n l_i' \cos \theta_i .$$

Si determina allora l'errore di chiusura laterale:

$$\Delta l = \sqrt{\Delta x^2 + \Delta y^2} .$$

Anche questo errore si confronta con la tolleranza lineare t_l .

Il Catasto consente:

$t_l \leq 0.025 \sqrt{L}$, mentre per le poligoni chiuse

$$t_l \leq \left(\frac{1}{6000} \sqrt{L} + 0.5 \right) m .$$

Se risulta $\Delta l = < t_l$ si procede alla compensazione riportando sulle coordinate parziali gli errori Δx e Δy proporzionalmente alla lunghezza dei lati. Le correzioni da apportare alle ascisse e alle ordinate calcolate risultano rispettivamente:

$$(6.3) \quad C_{x_i} = \frac{-\Delta X}{\sum |l_i \sin \vartheta_i|} \cdot l_i \sin \theta_i \text{ e. } C_{y_i} = \frac{-\Delta Y}{\sum |l_i \cos \vartheta_i|} \cdot l_i \cos \theta_i \text{ e.}$$

Disponendo anche delle quote q_1 e q_n , ovvero del dislivello teorico $\Delta_{1n} = q_n - q_1$, la verifica altimetrica approssimata consiste nel confrontare Δ_{1n} con Δ'_{1n} misurato e, quindi, ripartire l'errore proporzionalmente alle lunghezze dei lati.

Se l'errore altimetrico $\Delta_q = \Delta_{1n} - \Delta'_{1n} < t_q$, esso si ripartisce correggendo ciascun dislivello della quantità $-\Delta_{q/n}$ (o meglio proporzionalmente alle distanze). Come valore più probabile del dislivello $\Delta_{i,i+1}$ si assume la media aritmetica dei due dislivelli ottenuti in andata e in ritorno .

7 - Calcolo, verifica e compensazione approssimata di una poligonale topografica chiusa

Il procedimento è analogo a quello esposto, tenendo però presente che dalla verifica angolare si determina l'errore di chiusura angolare:

$$(7.1) \quad \Delta \alpha = (n-2)\pi - \sum \alpha_i$$

da confrontare con la tolleranza e ripartire equamente tra i vari angoli, ricavando successivamente gli angoli di direzione dei lati già corretti.

Dalla verifica laterale si determinano invece direttamente gli errori di chiusura in ascissa e in ordinata $\Delta_x = \sum l_i \sin \theta_i$ e $\Delta_y = \sum l_i \cos \theta_i$, e quindi l'errore :

$$\Delta l = \sqrt{\Delta x^2 + \Delta y^2}$$

di chiusura laterale, da confrontarsi con la tolleranza (valgono quelle catastali); se $\Delta l < t_1$ si apportano le correzioni alle coordinate parziali, date dalle formule (6.3) già viste.

Anche in questo caso si possono ricavare le quote compensate riportando l'errore di chiusura altimetrico con la stessa procedura.

8 - Rilievo dei particolari, sezioni

Consideriamo dapprima il metodo delle “sezioni”, impiegato come metodo principale di rilievo in zone molto estese in una direzione e poco nell'altra, ed in particolare per il progetto di una strada. Riferendosi alla rete di poligonali, stabilita come trama del rilievo, si rileva il terreno in modo preciso e dettagliato secondo tanti profili o sezioni eseguiti con piani normali, che partono sempre o dai vertici o da punti intermedi dei lati poligonali (le loro direzioni possono essere qualunque, soprattutto per il rilevamento delle zone più accidentate). Il rilievo delle sezioni procede come segue. (fig. 8.1)

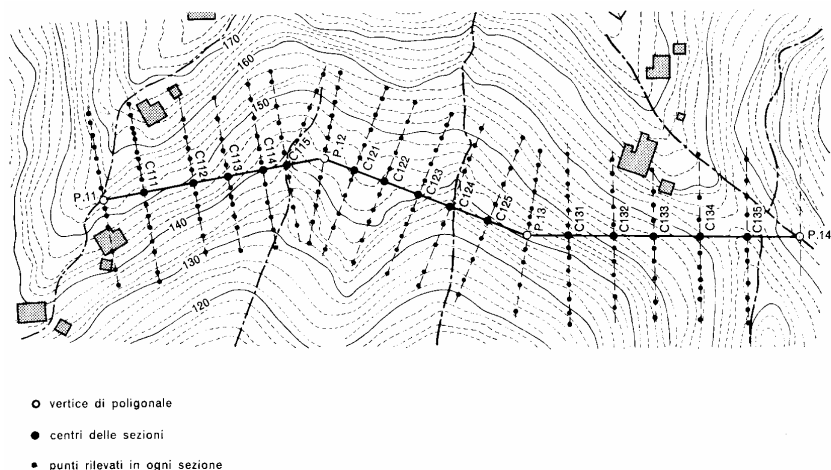


Fig. n. 8.1 - Rilievo di dettaglio con il metodo delle sezioni.

Occorre anzitutto determinare la posizione dei punti di partenza, sui quali si metteranno dei picchetti; bisogna quindi tracciare sul terreno gli allineamenti secondo cui le sezioni devono essere eseguite (con uno strumento ottico) e porre lungo di essi alcune “paline” in punti caratteristici dove il terreno cambia pendenza.

Disegnato sul “quaderno di campagna” uno schizzo della sezione, tracciando con tanti segmenti i tratti tra palina e palina, si determinano distanza e dislivello di ciascun punto battuto. Rilevate le sezioni per ottenere il piano a curve di livello si procede come segue.

Si disegna un profilo della sezione rilevata (con scale falsate fra distanze e dislivelli) con riferimento ad un conveniente piano (di quota alquanto inferiore alla quota più bassa del terreno) e si segnalano i punti corrispondenti alle quote a cui si desidera tracciare le curve di livello, nonché le loro proiezioni sulla traccia del piano di riferimento: sono queste ultime che servono per il disegno topografico. (fig. 8.2)

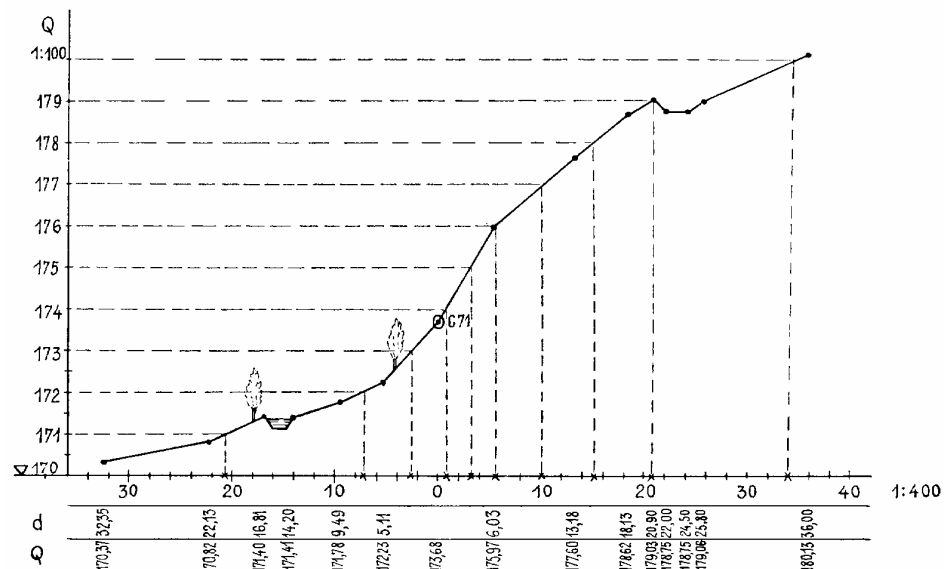


Fig. n. 8.2 - Montaggio di una sezione rilevata e deduzione delle intersezioni delle curve di livello.

Infatti, sulla carta, tracciate le poligonali e stabilite le direzioni delle sezioni (in base alle posizioni dei picchetti di sezione) si riportano sulle tracce dei piani di sezione i punti determinati dal profilo. Segnate così tutte le sezioni sulla planimetria, si congiungono i punti di ugual quota, costruendo le curve di livello. La distanza delle sezioni dipende dalla scala della carta e dalla natura del terreno, restando fissa la regola generale secondo cui la distanza fra due punti consecutivi determinati deve essere tale che tra l'uno e l'altro la pendenza del terreno si deve poter ritenere uniforme.

In base a tale regola può effettuarsi la trasformazione approssimata del "piano quotato" in "piano a curve di livello".

La celerimensura può considerarsi il metodo più generale e più duttile del rilievo dei particolari.

Viene eseguita impiegando una “stazione totale” (teodolite e distanziometro) con uno o più “porta prismi”. I punti sono determinati per mezzo delle loro coordinate cilindriche, e cioè: scelto un punto noto (quello di stazione) come polo e stabilita come direzione di riferimento quella di una stazione vicina, si determinano dei vari punti che interessano, l’angolo azimutale rispetto alla direzione scelta come riferimento, la distanza e il dislivello rispetto al punto di stazione.

Il rilievo celerimetrico presuppone naturalmente una rete di appoggio (poligonale aperta o chiusa) che deve essere rilevata con cura, sia planimetricamente che altimetricamente (la distanza tra i vertici può essere da qualche centinaia di metri fino ai km).

Lo strumento viene messo in stazione, rettificato, e si misura l’altezza strumentale h_s ; successivamente, posta la palina con “prisma retroriflettente” sul punto da battere si collima il centro del prisma annotando, quindi, i tre valori fondamentali H_A , angolo orizzontale, H_V , angolo verticale, e d_i , distanza inclinata (non è necessario applicare la regola di Bessel, è sufficiente una sola lettura, poiché si tratta di determinare punti abbastanza vicini, spigoli di fabbricati, punti caratteristici, ecc.).

I dati letti si registrano in una tabella, se non si dispone di un registratore magnetico, nella quale, oltre alle letture si devono memorizzare il nome e la quota della stazione, il vertice di riferimento, l’altezza strumentale e tutte le osservazioni che si ritengono utili per il disegno della carta; la seconda parte della tabella è dedicata ai calcoli della distanza e del dislivello. Lo schizzo, detto anche “Eidotipo” (vedi fig.12), deve contenere: il punto di stazione, tutti i particolari planimetrici del terreno interessanti il rilievo, i punti da battere e l’andamento approssimativo delle curve di livello.

I punti vengono contrassegnati con numeri progressivi, continuando la numerazione anche nelle successive stazioni (per evitare confusioni) e devono essere ben distinti.

Quando sono rilevati direttamente sul terreno vanno scelti con cura, in modo tale da poter ritenere costante la pendenza tra due punti vicini (più fitti dove maggiori sono le accidentalità), ipotesi che permetterà le interpolazioni lineari per il tracciamento delle curve di livello.

L'uso dei distanziometri ad onde da unirsi ad un teodolite, e dei teodoliti con il distanziometro incorporato, consente di effettuare una celerimensura di alta precisione e di raggio anche superiore al Km. Calcolate le coordinate cilindriche dei punti, si deve passare al piano quotato, riportando planimetricamente i punti per mezzo delle loro coordinate polari e scrivendo vicino a ciascuno la propria quota.

Per tracciare le curve di livello bisogna interpolare linearmente tra i punti quotati, la cui rappresentazione prende il nome di "PIANO QUOTATO". (fig. 8.3)

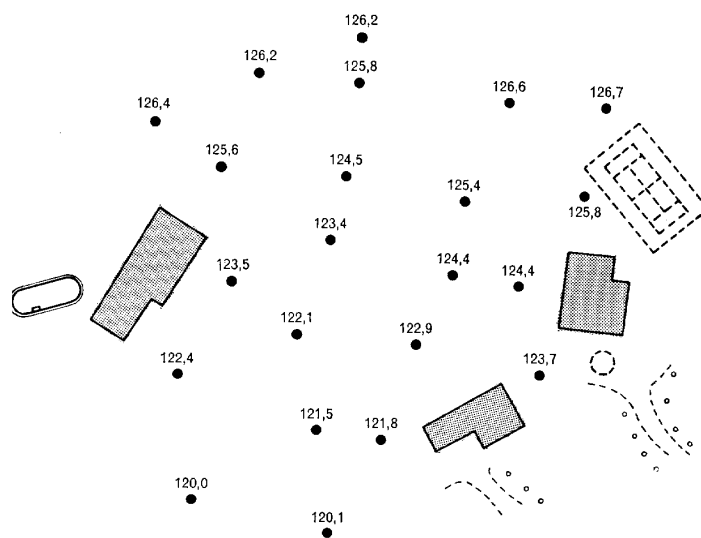


Fig. n. 8.3 - Piano quotato.

Se i dati di campagna sono registrati su registratore magnetico e se i calcoli sono effettuati con un calcolatore fornito degli opportuni programmi, i risultati possono essere trasferiti immediatamente al plotter, il quale disegna automaticamente il piano quotato e, con lo stesso programma, anche le curve di livello. (fig. 8.4)

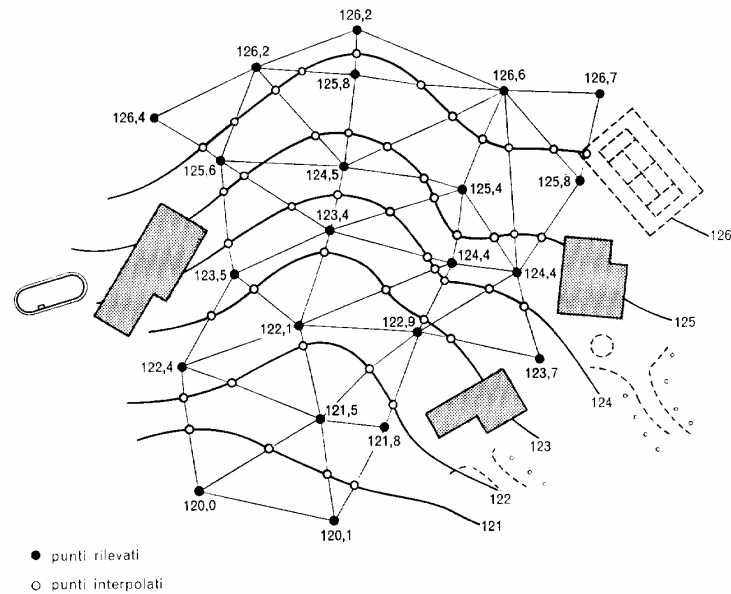


Fig. n. 8.4 - Curve di livello ottenute per interpolazione del piano quotato.

9 - Considerazioni riassuntive sui rilievi topografici

Dopo un primo sopralluogo sulla zona da rilevare, regola fondamentale sia per la scelta della metodologia da adottare sia degli strumenti più opportuni da impiegare è quella di adattare i

procedimenti di rilievo alle caratteristiche del rilievo stesso e del terreno.

Le caratteristiche del rilievo dipendono dallo scopo per cui esso va eseguito e dalla scala che si adotterà, la quale dipende a sua volta dall'uso che della carta si farà.

La precisione necessaria, che dovrà comunque far sì che l'errore medio del punto sia inferiore al prodotto di n (denominatore di scala) per ϵ (errore di graficismo $\cong 0.2$ mm) riportato alla scala del disegno, è assai elevata per le grandi scale, diviene grossolana alle piccole scale (e permette così di adoperare strumenti semplificati e metodi speditivi di rilievo). D'altra parte, la precisione altimetrica deve essere in armonia con quella planimetrica, ricordando l'importanza dei dislivelli nei confronti della pendenza (**l'acqua scende sempre dal punto più alto a quello più basso**). Anche la densità dei punti da rilevare dipende in modo essenziale dalla scala della carta.

Altro elemento da considerare attentamente è l'estensione del terreno, perché un rilievo di una vasta area porta con sé problemi piuttosto complessi qualunque sia la scala della carta. Infatti, tenendo presente che i punti di appoggio devono essere determinati con precisione tanto più grande quanto più elevato è il loro ordine, accade che il reticolato dei punti fondamentali deve avere una precisione tanto maggiore quanto più estesa è la zona da rilevare, qualunque sia la scala della carta. Pertanto, può accadere che la rete fondamentale di un grande reticolato destinato ad un rilievo in piccola scala debba essere molto più precisa della piccola poligonale necessaria per rilevare un modesto appezzamento, anche se nel secondo caso la carta avrà scala molto grande.

Lo studio del programma del rilievo comprende la rete di appoggio e le operazioni di dettaglio. Per la rete di appoggio bisogna stabilire se ricorrere ad una rete trigonometrica o solo a poligonali. Effettuata la scelta, si passa alla individuazione dei vertici in modo da seguire i percorsi più opportuni per distribuire sul terreno i punti di appoggio delle misure di dettaglio e in modo che ogni vertice sia ben

visibile da quelli adiacenti. Infine, per il rilievo di dettaglio si hanno a disposizione i due metodi: celerimetrico e per sezioni.

Il primo è di gran lunga il più usato, salvo che per rilevare strisce di modesta larghezza (studio dei tracciati di opere tecniche). Per determinare le quote si ricorre quasi sempre alla livellazione distanziometrica sia per i vertici delle reti che per le poligonali ed il rilievo di dettaglio. Il ricorso alla livellazione geometrica si giustifica quando è necessario raggiungere una maggiore accuratezza.

Individuati i metodi del rilievo diviene abbastanza semplice la scelta degli strumenti più opportuni.

Il segreto finale per la buona esecuzione di un rilievo è, l'ordinato svolgimento delle operazioni, che significa ordine e cura nell'esecuzione delle misure, nella registrazione dei dati, nel disegno degli schizzi e di tutto il materiale che costituisce l'eidotipo, che servirà per la redazione definitiva della carta.

Il rilievo topografico non presenta delle vere difficoltà concettuali e non bisogna spaventarsi dinanzi a condizioni difficili, ma ricordare che i metodi topografici sono stati applicati nelle condizioni più diverse e disparate e, quindi, offrono una soluzione in qualunque caso. In un rilievo non si devono operare invenzioni, ma applicare razionalmente procedimenti già sperimentati; si tratta solo di applicarli bene.