

Il sistema GPS

3.1 Introduzione

Il dipartimento della difesa USA (DoD) sviluppò, negli anni '70, NAVSTAR GPS (**NAV**igation **S**atellite **T**iming **A**nd **R**anging **G**lobal **P**ositioning **S**ystem), un sistema di posizionamento basato sulla ricezione e successiva elaborazione di segnali provenienti da una costellazione di satelliti artificiali in orbita non geostazionaria; tale sistema consente di calcolare la posizione di stazioni riceventi in un sistema di riferimento tridimensionale geocentrico; come si vedrà più avanti nel dettaglio, le coordinate del punto di ricezione potranno essere espresse utilizzando, in alternativa, un classico sistema cartesiano (x,y,z) oppure uno specifico modello di riferimento (φ, λ, h) ellissoidico.

Il sistema GPS prevede molteplici utilizzi civili e militari: posizionamento di stazioni riceventi fisse o mobili, rilievi cinematici e topografici di alta precisione, fotogrammetria aerea, monitoring delle deformazioni crostali...; fornendo un duplice livello di accuratezza nella misura, è in grado di supportare applicazioni con requisiti diversi in termini di precisione richiesta. Il sistema di posizionamento GPS utilizza, infatti, due distinte metodologie operative, diverse sia per lo schema di principio sottostante che per il grado di precisione conseguibile; di seguito vengono presentate sinteticamente le due tecniche, rimandando ai paragrafi successivi un'analisi più dettagliata.

- **Standard GPS:** sfrutta la tecnica **Point Positioning** e si riferisce alla localizzazione assoluta di un punto singolo in un sistema di riferimento assegnato; in particolare, i segnali originati da diversi satelliti e demodulati da un ricevitore, vengono elaborati in modo indipendente per determinare la posizione del ricevitore stesso. Questo metodo può essere applicato agevolmente da hardware dedicato in tempo reale e si presta a stime di velocità e traiettoria per veicoli in moto; fornisce misure la cui precisione risulta dell'ordine della *10 metri* per i ricevitori in dotazione all'esercito statunitense e di circa *100 metri* per i ricevitori destinati all'utenza civile: la differenza di precisione è dovuta essenzialmente alla degradazione intenzionale introdotta sul segnale diretto all'utenza civile;

- **Differential GPS:** sfrutta **Differential Positioning**, una tecnica di posizionamento di natura differenziale: l'idea base è quella di utilizzare una stazione ricevente fissa in posizione nota $[(\varphi, \lambda, h)_{\text{REFERENCE}}]$ in grado di calcolare l'errore corrente di posizionamento che grava sul segnale ricevuto localmente e di assumere che tale errore si mantenga inalterato entro una certa distanza dalla stazione di riferimento; la stima dell'errore viene infine utilizzata dal ricevitore per correggere la posizione calcolata con la tecnica Point Positioning. Il metodo differenziale consente di ottenere una precisione dell'ordine del metro e anche migliore: un risultato notevole rispetto all'accuratezza del metodo di posizionamento assoluto ma lo svantaggio principale risiede nella necessità di stabilire un link di comunicazione con la stazione fissa per il posizionamento real time; per questo motivo, la tecnica differenziale viene utilizzata, solitamente, a posteriori [post – processing delle osservazioni raccolte durante un rilievo].

L'interesse da parte della comunità scientifica internazionale verso il sistema e gli studi condotti hanno consentito anche agli apparati riceventi non militari di utilizzare il servizio con una precisione notevolmente superiore a quella ipotizzata dai progettisti; per questa ragione le applicazioni pratiche nel settore civile sono diventate numerosissime:

- GPS è utilizzato come sistema base per il controllo della navigazione aerea e marittima;
- GPS è utilizzato dall'industria estrattiva per la localizzazione dei mezzi terrestri e navali impiegati nelle ricerche di giacimenti e per mantenere costantemente in posizione le piattaforme offshore;
- GPS è utilizzato per il monitoring di posizione e velocità per veicoli (servizi location based, antifurto, soccorso..);
- L'introduzione di varianti ancora più precise al sistema di posizionamento [DGPS] ne ha consentito l'uso in applicazioni cartografiche, geofisiche, idrografiche; molti laboratori scientifici utilizzano i segnali GPS per estrarre l'ora UTC (Universal Time Coordinated) con una precisione di $0,1 \mu\text{s}$.

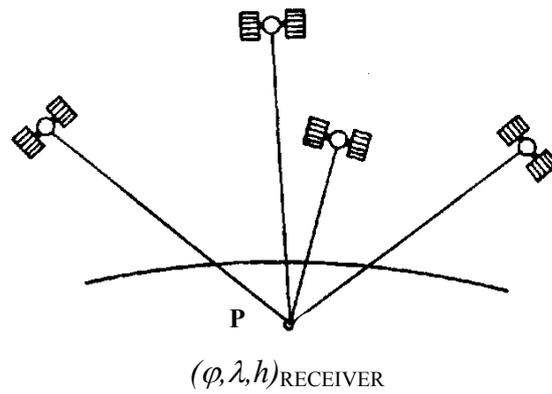


Figura 3.1 Point Positioning

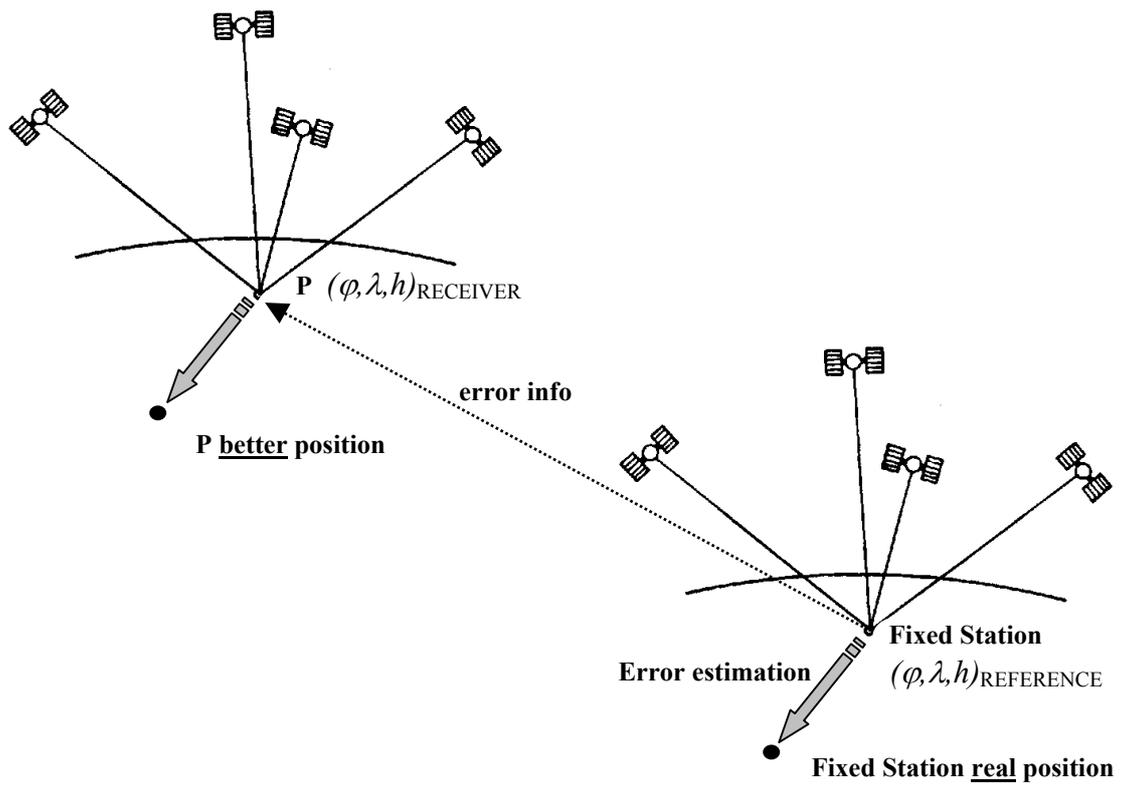


Figura 3.2 Differential Positioning

3.2 Architettura di sistema

L'architettura del sistema GPS risulta piuttosto complessa; tradizionalmente, viene illustrata suddividendola in tre segmenti: il **segmento spaziale**, il **segmento di controllo** ed il **segmento di utilizzo**

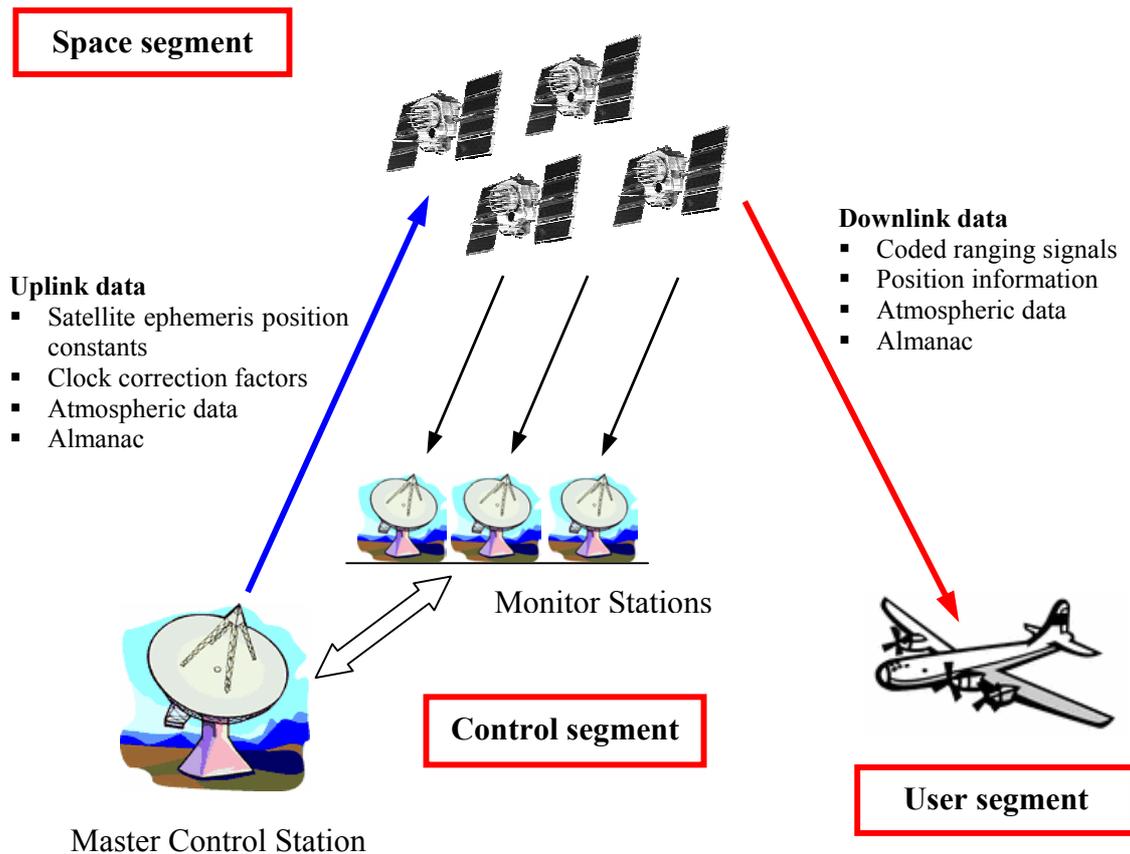


Figura 3.3 Segmento spaziale, segmento di controllo e segmento di utilizzo

3.2.1 Il segmento spaziale

Il segmento spaziale era stato inizialmente progettato in modo da realizzare una costellazione di 24 satelliti [21 attivi e 3 di riserva] distribuiti su 6 piani orbitali inclinati di 55° rispetto al piano equatoriale e posizionati ad intervalli di longitudine fra i nodi

ascendenti di 60° ; ciascun satellite trasmette, in broadcast, i codici di distanza a radiofrequenza ed i dati di navigazione; attualmente il numero di satelliti è variabile da un minimo di 24 ad un massimo di 32 a seconda delle rispettive condizioni di funzionamento; i 4 satelliti sono (nominalmente) distribuiti in modo simmetrico all'interno di ciascuno dei 6 piani orbitali; in realtà tale distribuzione, nel corso degli anni, è stata continuamente modificata per fornire una copertura adattativa rispetto alle esigenze di accuratezza nelle diverse zone del pianeta: la configurazione attuale, apparentemente irregolare, garantisce la copertura diffusiva del servizio di posizionamento con un'accuratezza commisurata alle esigenze medie locali. Inoltre, i 3 satelliti di riserva (spare) possono essere eventualmente spostati per ottenere copertura radio aggiuntiva in alcune zone del pianeta (essenzialmente a scopi militari).

Con riferimento alla costellazione attuale, ciascun piano orbitale viene etichettato con una lettera [da **A** ad **F**] e ciascun satellite [nell'ambito di uno stesso piano orbitale] con un numero da **1** a **4**; nel figure da 3.3 a 3.10 verranno mostrati i singoli piani orbitali ed una vista d'insieme della costellazione; nel tool che sarà descritto nel Capitolo 5, è stata realizzata una simulazione tridimensionale del sistema GPS utilizzando le API Java 3D.

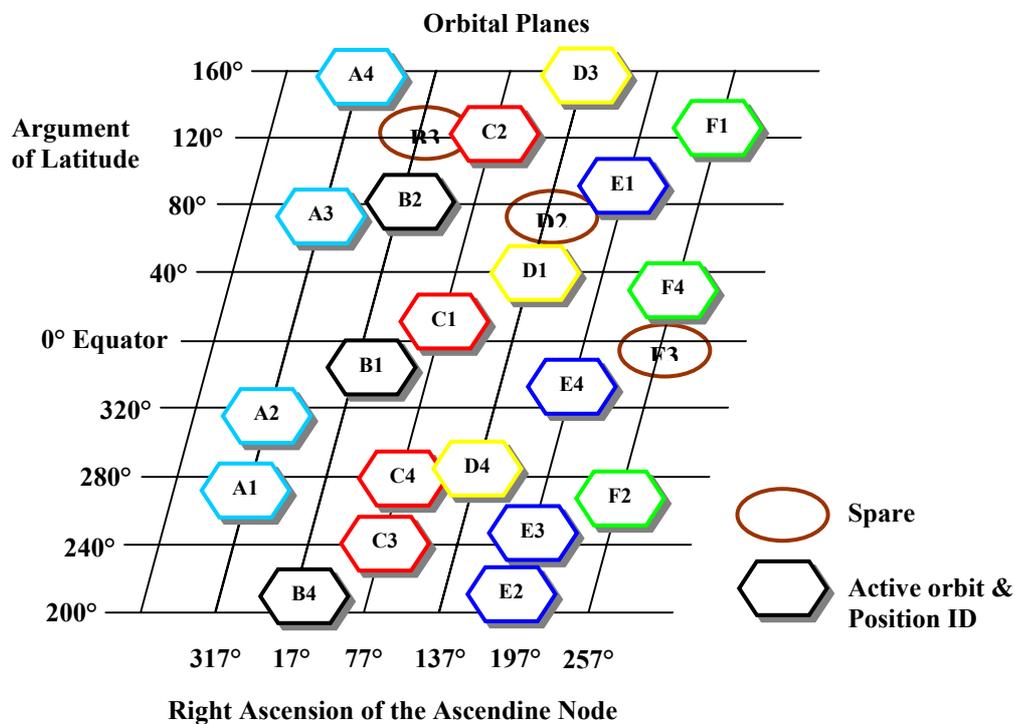


Figura 3.4 Rappresentazione semplificata della costellazione GPS [vedi figura 3.14]

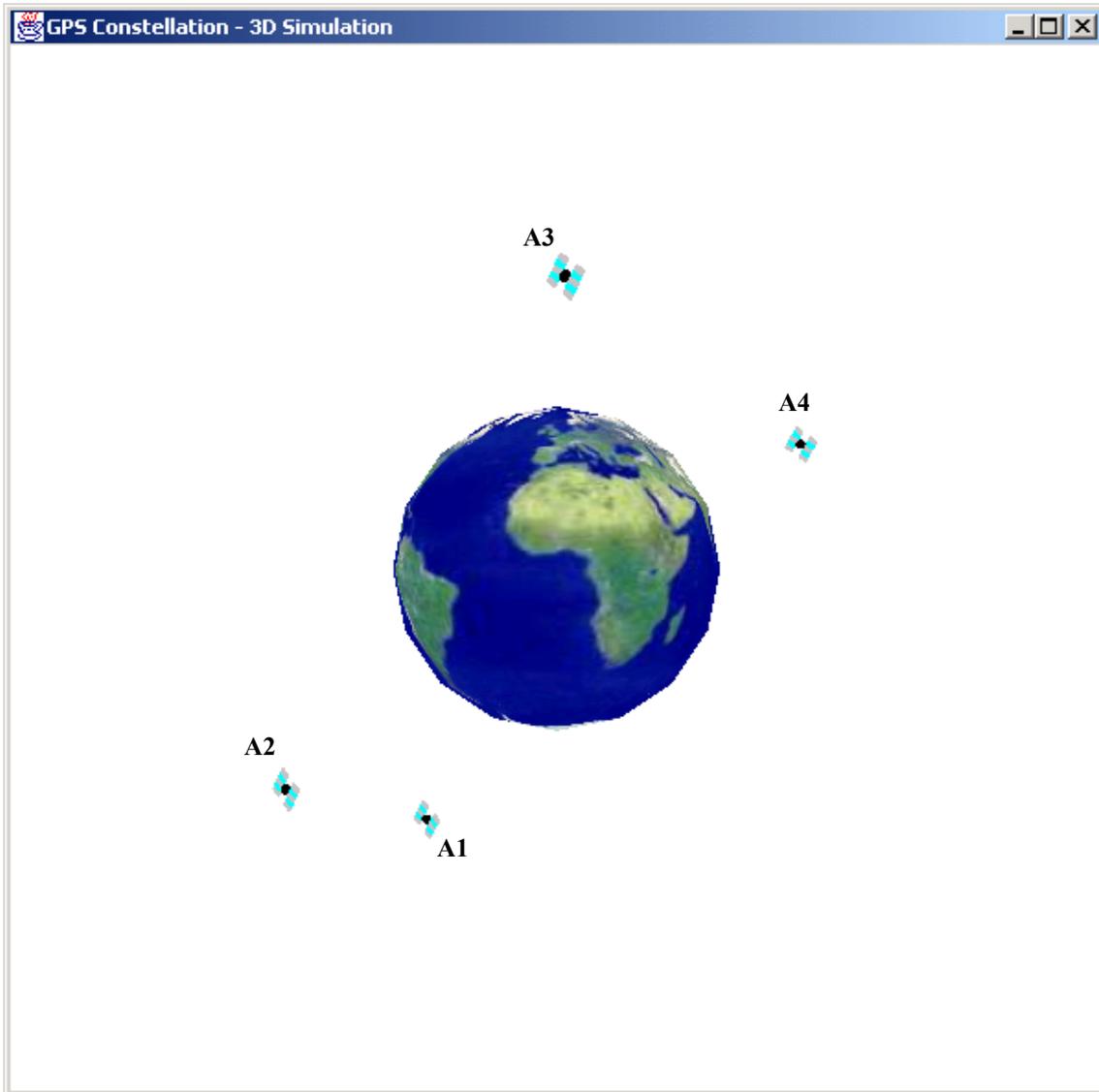


Figura 3.5 Orbital Plane A

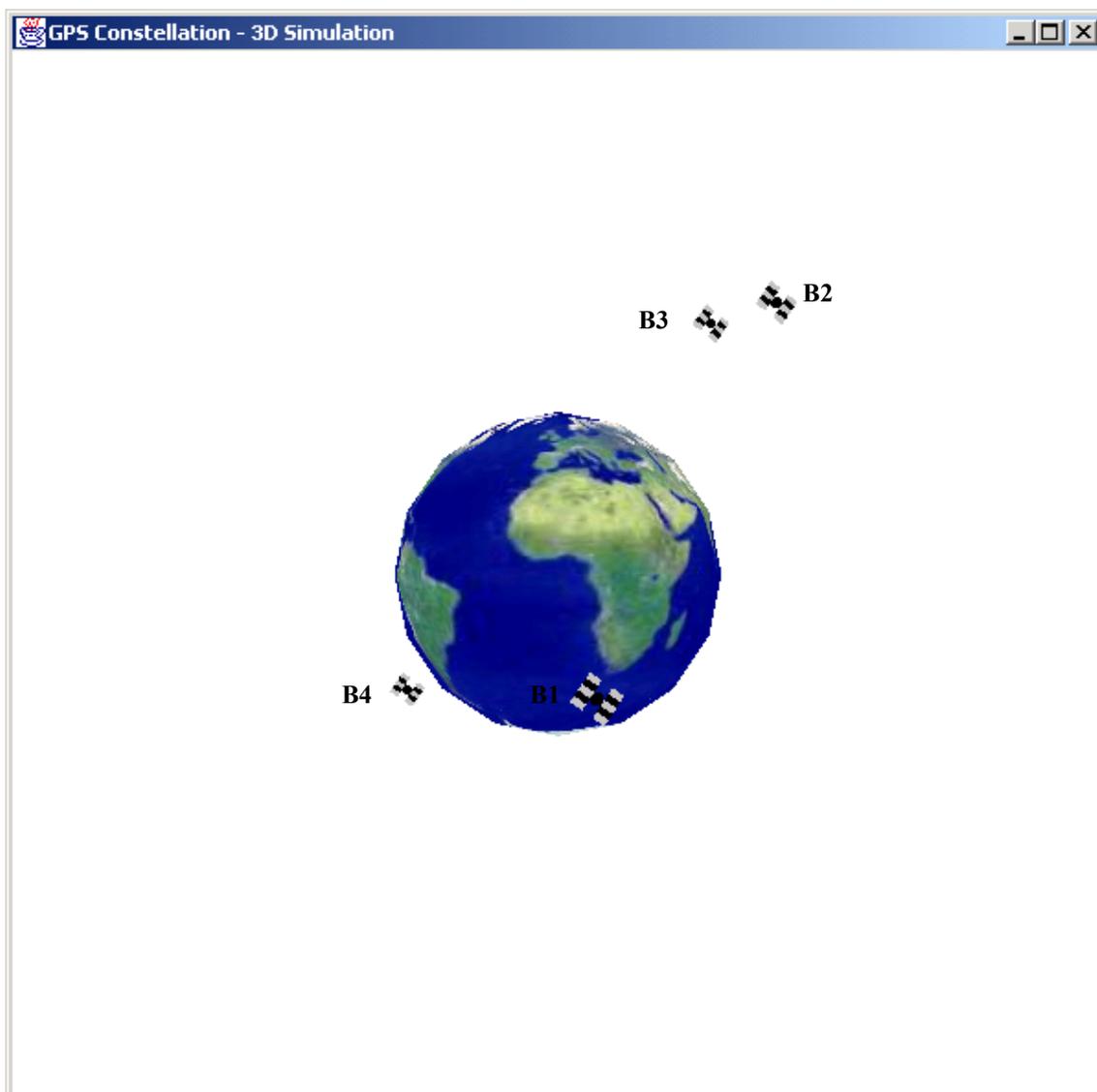


Figura 3.6 Orbital Plane B

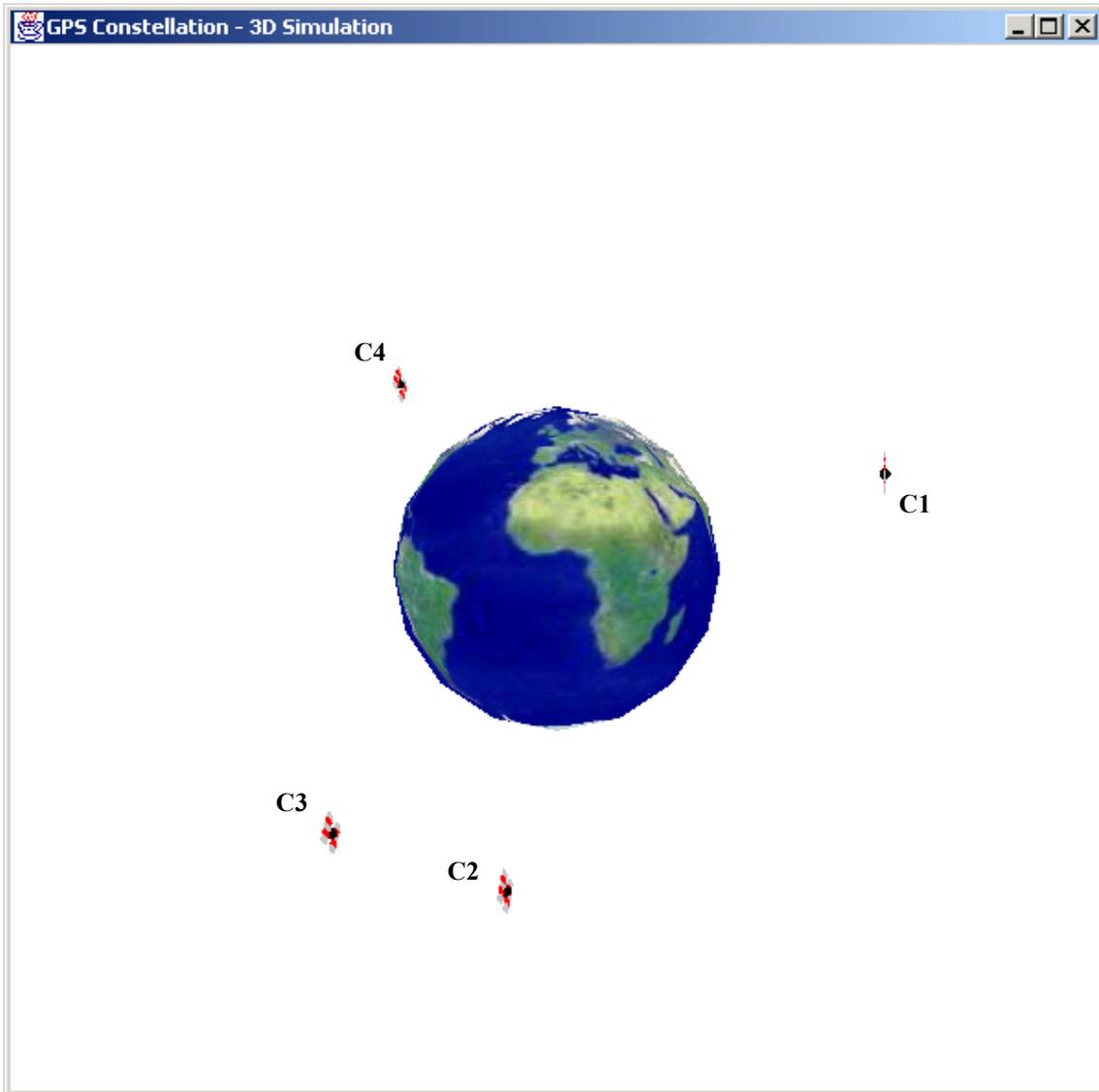


Figura 3.7 Orbital Plane C

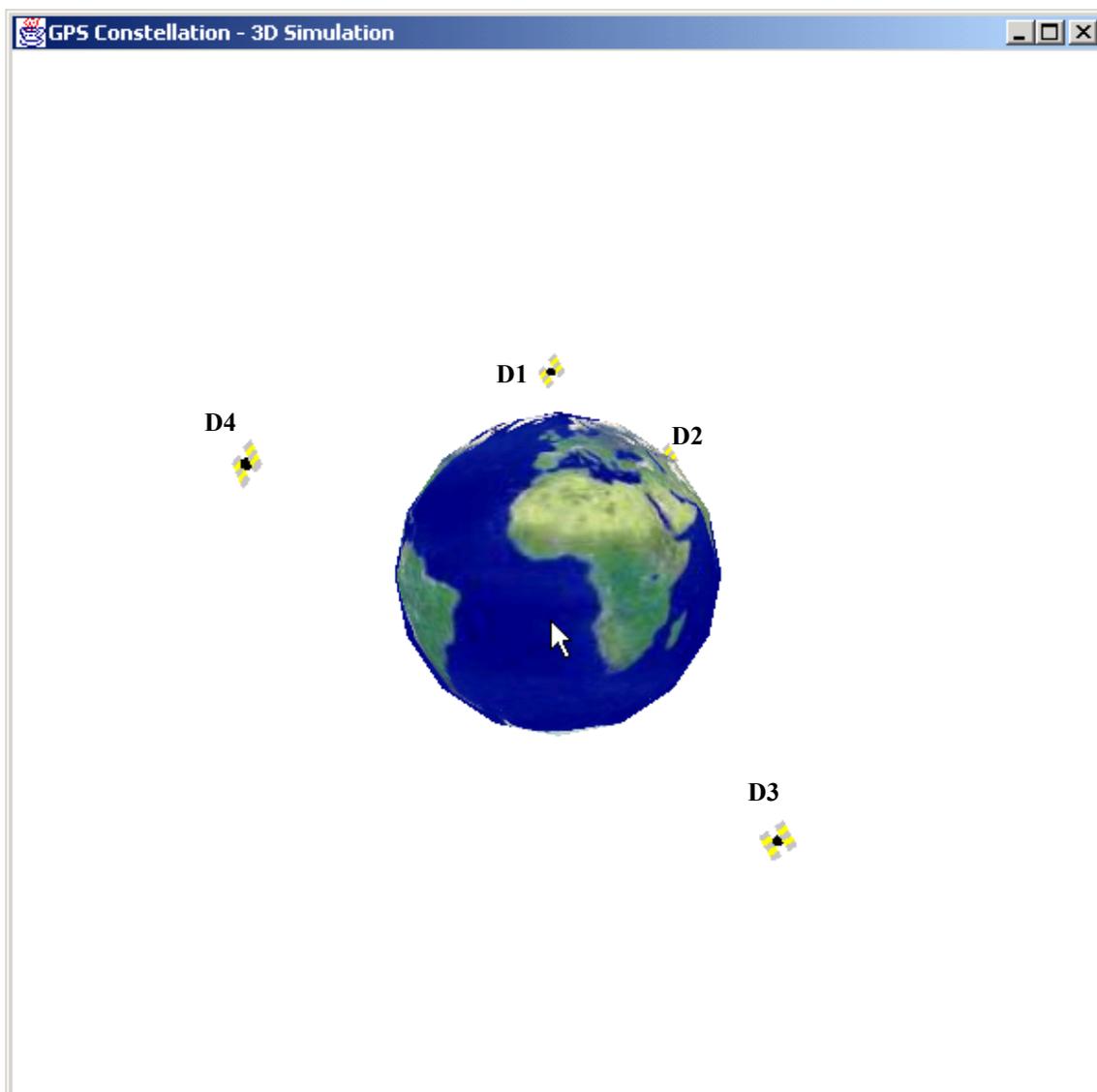


Figura 3.8 Orbital Plane D

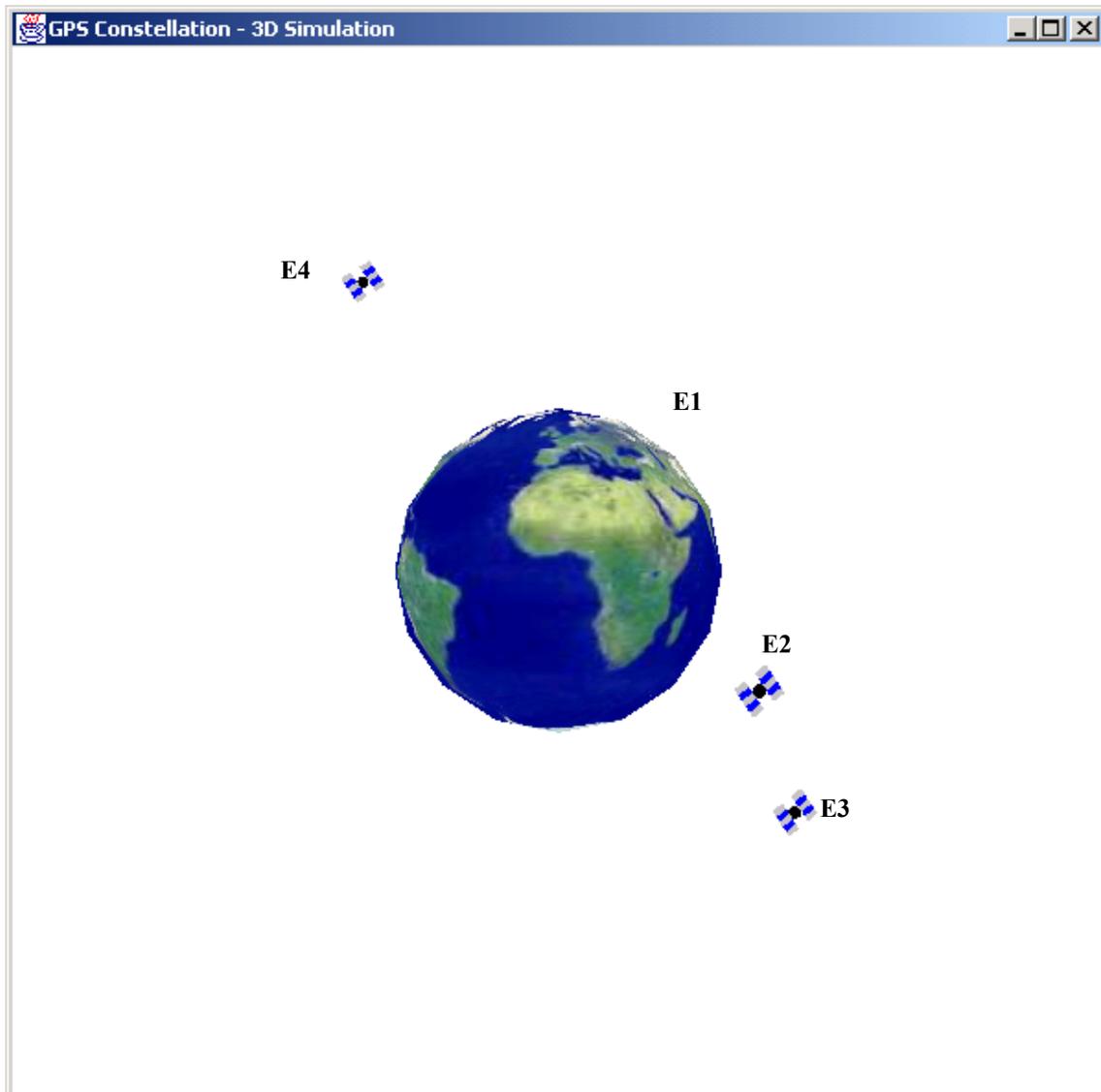


Figura 3.9 Orbital Plane E

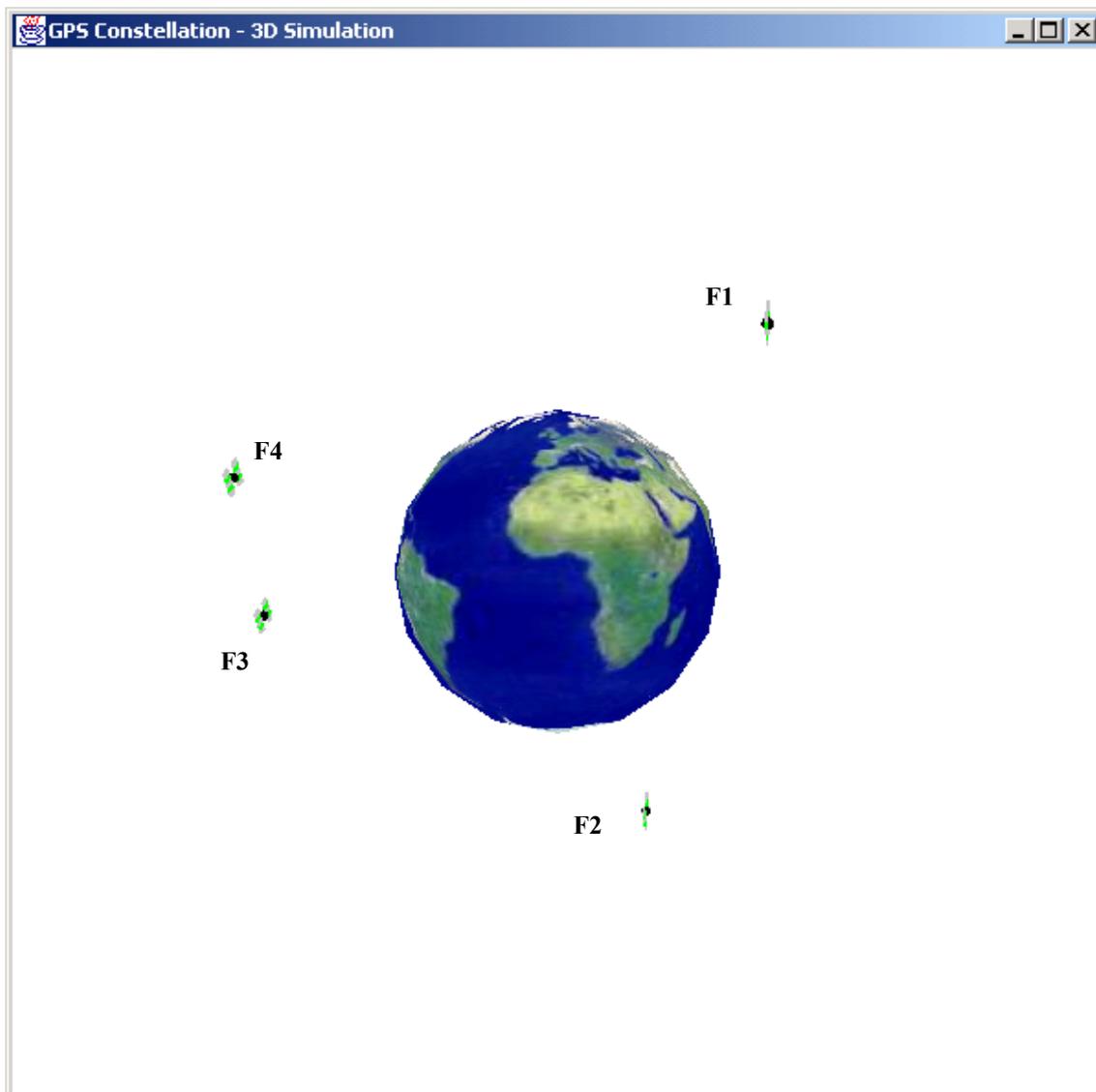


Figura 3.10 Orbital Plane F

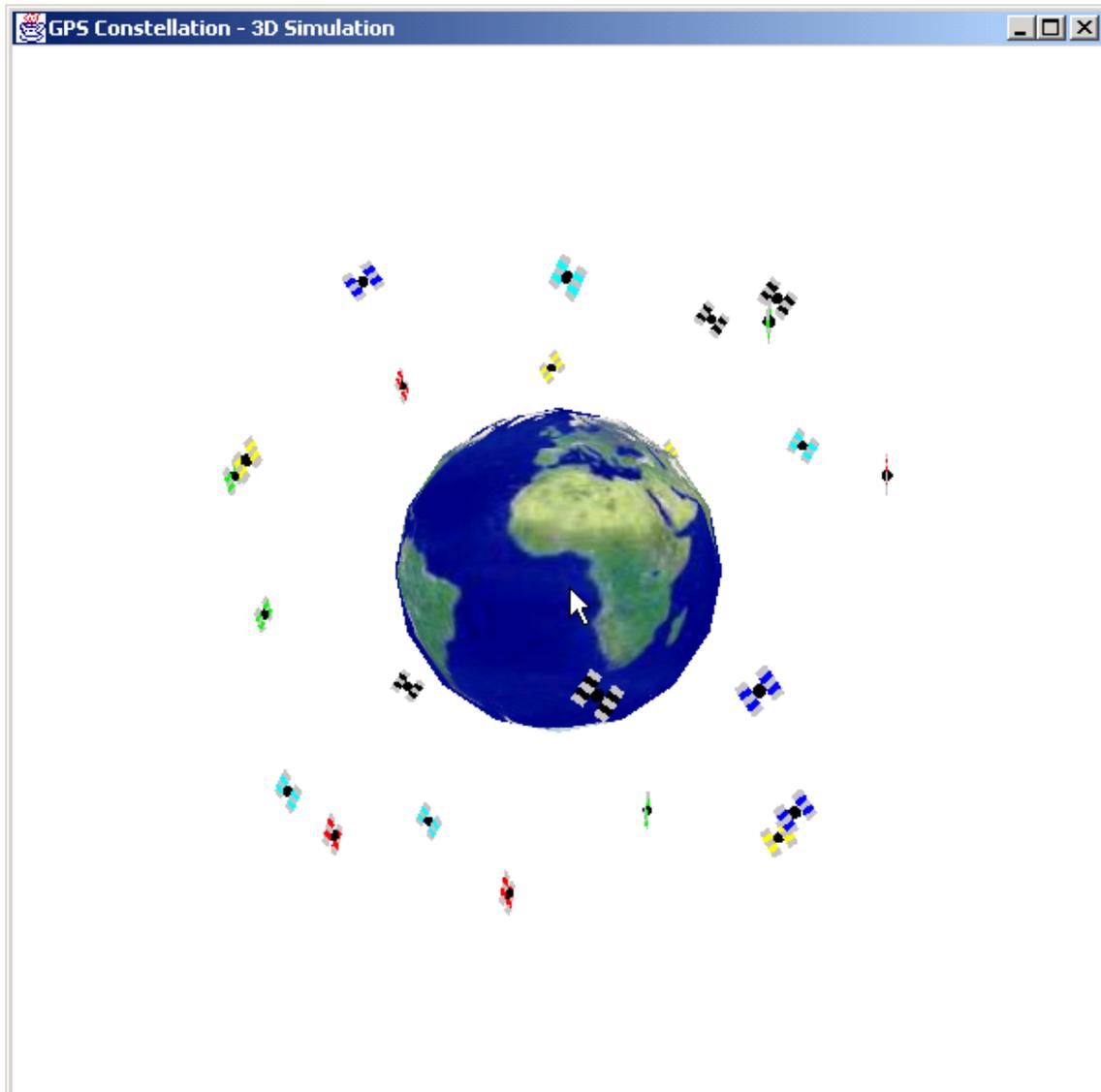


Figura 3.11 The GPS constellation

Nel tempo si sono alternate diverse generazioni di satelliti GPS, differenti fra loro per le componenti installate e la durata operativa; in particolare:

- **Satelliti del blocco I:** lanciati allo scopo di disporre di prototipi per lo sviluppo, dal 1978 al 1985 in numero di 11 e ormai tutti fuori uso; tempo di vita medio del singolo satellite: 5 anni;
- **Satelliti del blocco II [operativi]:** lanciati, in numero di 9, dal 1989 al 1990, sono stati i primi ad essere veramente operativi; tempo di vita medio del singolo satellite: 8 anni; sebbene la prima dichiarazione di inizio operatività del sistema risalga all'8 dicembre 1993, NAVSTAR GPS fu ufficialmente dichiarato operativo il 27 aprile 1995.
- **Satelliti del blocco IIA [operativi]:** lanciati, in numero di 15, dal 1990 al 1997; tempo di vita medio del singolo satellite: 8 anni;
- **Satelliti del blocco IIR [di sostituzione]:** lanciati, in numero di 20, dal 1997 ad oggi; tempo di vita medio del singolo satellite: 8 anni;
- **Satelliti del blocco IIF:** rappresenta la prossima generazione di satelliti GPS; i lanci sono iniziati nel 2002.

I satelliti GPS hanno massa di circa 800 Kg; a partire da quelli del blocco II, ciascun satellite è stato equipaggiato con 4 oscillatori, 2 al cesio e 2 al rubidio, allo scopo di garantire un segnale di tempo quanto più accurato possibile; l'approvvigionamento energetico è assicurato da pannelli solari la cui superficie si sviluppa per circa 7 m²; a partire dalla serie IIR, i satelliti effettuano anche **power control**: adattano la potenza trasmessa in funzione della loro posizione in modo da irradiare unicamente in un cono tangente alla superficie terrestre e con vertice nel satellite; una serie di retrorazzi viene utilizzata per apportare manovre correttive all'orbita del satellite; tutti i satelliti sono soggetti a periodi di inutilizzabilità a causa di manutenzioni o manovre: queste ultime vengono effettuate almeno una volta l'anno mentre gli orologi atomici al Cesio richiedono, mediamente 2 volte l'anno, una particolare manutenzione che dura circa 18 ore.

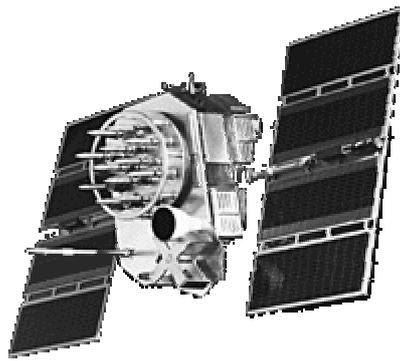


Figura 3.12 Un satellite GPS

Le orbite dei satelliti GPS, non geostazionarie, risultano approssimativamente circolari (eccentricità nulla) con raggio circa 26.560 Km e periodo nominale 12 ore siderali: ciascun satellite si muove pertanto ad una velocità dell'ordine di 4 Km/sec passando sullo stesso punto 2 volte in un giorno . La costellazione è stata progettata in modo tale che per un qualunque osservatore sulla superficie terrestre, ogni satellite risulti in vista per circa 5 di queste 12 ore: una tale configurazione garantisce in ogni momento la presenza sopra l'orizzonte di almeno 4 satelliti che inviano a Terra i dati necessari al posizionamento della stazione ricevente. In effetti, e questo è stato rilevato anche nelle nostre esperienze di misura, il numero di satelliti visibili resta compreso fra 7 e 9 per il 90% del tempo.

In figura 3.3 vengono riportate le distanze satellite – superficie terrestre per il sistema GPS e per sistemi geostazionari, questi ultimi utilizzati nelle applicazioni meteorologiche e multimediali; è opportuno ricordare che un satellite in orbita geostazionaria (**orbita sincrona**) possiede velocità relativa nulla rispetto alla terra: ad un generico osservatore statico a terra appare pertanto come un punto fisso; per questo motivo le orbite dei satelliti GPS vengono spesso indicate come **orbite semisincrone**.

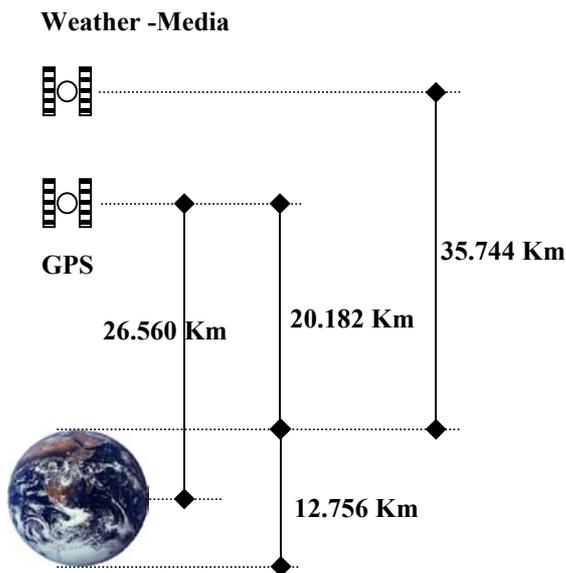


Figura 3.13 Distanze satellite – superficie terrestre

Di seguito vengono riportate le funzionalità principali realizzate dal generico satellite:

- mantenere e fornire un segnale di tempo molto accurato utilizzando gli oscillatori al cesio ed al rubidio (accuratezza compresa fra 10^{-13} e 10^{-14} sec.);
- trasmettere informazioni agli utilizzatori a Terra attraverso segnali di tipo opportuno [interfaccia con il segmento di utilizzo];
- ricevere, memorizzare ed elaborare le informazioni trasmesse da una rete di stazioni a terra [interfaccia con il segmento di controllo]; l'informazione di segnalazione trasportata può riguardare i parametri dell'architettura protocollare o eventuali comandi di manovra diretti ai retrorazzi allo scopo di apportare correzioni all'orbita al satellite;

Il metodo di posizionamento GPS presuppone che l'utente conosca o possa calcolare, riferendola ad una certo istante di misura nel passato, la posizione dei satelliti osservati; a causa della velocità relativa non nulla e dell'elevato ritardo di propagazione dei segnali, un satellite non può comunicare la propria posizione in modo istantaneo: per

questo motivo invia un insieme sintetico di parametri che permettono il calcolo della sua posizione (**effemeridi**); questa scelta trova giustificazione in quanto segue.

Qualora si considerasse la Terra perfettamente sferica ed omogenea e non si considerasse l'azione di Luna, Sole ed atmosfera, qualunque satellite artificiale si muoverebbe lungo orbite ellittiche e predicibili nel tempo (leggi di Keplero); la sua orbita geocentrica $\mathbf{r}(t)$ in uno spazio inerziale verrebbe descritta da una equazione differenziale del secondo ordine del tipo:

$$\ddot{\mathbf{r}} = -GM \frac{\mathbf{r}}{r^3}$$

dove GM indica il prodotto tra la costante gravitazionale e la massa della terra, \mathbf{r} il vettore congiungente il satellite con il centro della terra ed r il suo modulo; la posizione di un satellite GPS sarebbe pertanto facilmente calcolabile, per qualsiasi epoca del futuro, fornendo unicamente un numero assegnato di parametri iniziali.

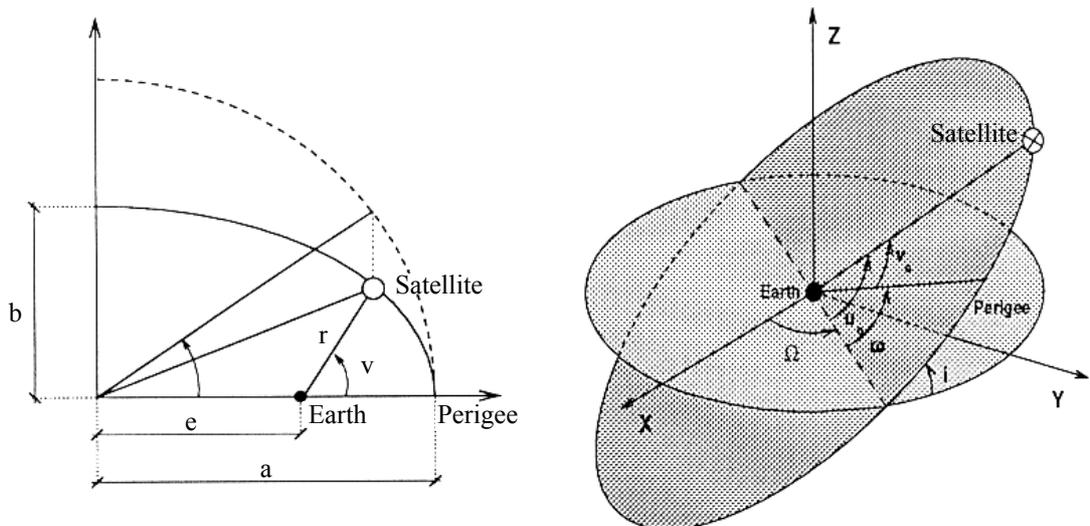


Figura 3.14 Parametri orbitali dei satelliti.

I 6 elementi orbitali caratteristici sono:

- il *semiasse maggiore* **a** dell'orbita: definisce l'ampiezza dell'orbita;
- l'*eccentricità* dell'orbita **e**: descrive lo schiacciamento dell'orbita;
- l'*inclinazione* **i** del piano orbitale rispetto al piano equatoriale;

- l'*ascensione retta* Ω del nodo ascendente, ossia l'angolo tra la direzione dell'equinozio di primavera (direzione X della figura) e la linea di intersezione del piano orbitale con il piano equatoriale;
- l'*argomento del perigeo* ω che rappresenta l'angolo (nel piano orbitale) tra il nodo ascendente e il perigeo (misurato nella direzione del moto del satellite);
- l'*argomento di latitudine* u_0 , cioè l'angolo tra il nodo ascendente e la posizione del satellite al tempo iniziale t_0 .

Gli elementi i e Ω rappresentano gli *angoli Euleriani* e definiscono l'orientazione del piano orbitale nel sistema equatoriale.

L'ipotesi di completa predicibilità dell'orbita, tuttavia, non risulta realistica a causa delle molteplici azioni perturbative che agiscono sul satellite e che determinano fluttuazioni periodiche o degradazioni nelle orbite; nella figura 3.15 viene riportato un elenco, non esaustivo, delle principali non idealità ed il corrispondente effetto.

Azione perturbativa o degradante	Effetto sull'orbita in 24 h (m)
Disomogeneità della Terra	10.000
Attrazione Lunare	3.000
Attrazione Solare	800
Effetti gravitazionali	2.000
Pressione di radiazione diretta	200
Y-bias	2
Effetti di maree	0.3

Figura 3.15 Principali forze perturbative e degradanti che agiscono sul satellite nell'arco di una giornata e corrispondente effetto.

Come si può facilmente osservare, la non perfetta sfericità della terra e gli effetti gravitazionali rappresentano le azioni perturbative che influiscono maggiormente sulle fluttuazioni periodiche delle orbite.

L'equazione reale del moto, che considera anche i parametri di disturbo dell'orbita, risulta:

$$\ddot{\mathbf{r}} = -GM \frac{\mathbf{r}}{r^3} + d(t, \mathbf{r}, \dot{\mathbf{r}}, f_0, f_1, f_2, \dots)$$

dove f_0, f_1, f_2, \dots indicano le forze perturbative stimate in fase di determinazione dell'orbita; ne consegue l'impossibilità pratica di descrivere il moto del satellite a lungo termine con un numero ragionevole di parametri. Per questo motivo, ciascun satellite GPS comunica una descrizione della propria orbita giornaliera mediante 24 blocchi di parametri; ogni blocco ha validità un'ora e contiene i parametri necessari a descrivere un arco d'orbita quasi ellittica. Il ricevitore acquisisce tali parametri e li utilizza per stimare la posizione del satellite; in questo modo, invece di calcolare l'orbita una volta per tutte cercando di tener conto delle innumerevoli azioni di disturbo, di cui molte aleatorie, si preferisce, come si illustrerà meglio nel prossimo paragrafo, effettuare il calcolo continuativamente inseguendo dinamicamente l'orbita reale del satellite.

3.2.2 Il segmento di controllo

Il segmento di controllo ha la funzione di controllare il moto orbitale dei satelliti (tracking), il funzionamento entro i limiti di tolleranza degli orologi e comunicare ai satelliti i messaggi di navigazione da trasmettere in broadcast. Il controllo viene effettuato mediante 5 stazioni a terra, distribuite lungo l'equatore in modo tale che il segnale di un qualunque satellite sia ricevuto da almeno una di queste stazioni; nel complesso, l'insieme delle 5 stazioni realizza una rete di controllo, in grado di ricevere continuamente i segnali emessi da tutti i satelliti, verificare le rispettive effemeridi (i dati descrittivi sulla posizione dei satelliti) e predire, per ciascuno satellite, l'orbita in un determinato intervallo di tempo futuro. Nella figura successiva vengono mostrate le posizioni delle 5 stazioni **monitor**; di queste, solo l'**MSC** (Master Station Control) situata a Colorado Springs, è localizzata sulla terra ferma.



Figura 3.16 Le stazioni di controllo.

I dati raccolti da ciascuna stazione monitor vengono inviati alla stazione Master la quale effettua una stima quotidiana dell'orbita e dell'offset d'orologio previsti per ciascun satellite nelle 24 ore successive; le orbite previste vengono parametrizzate ed i dati corrispondenti inviati al singolo satellite attraverso una stazione di **upload**; quest'ultimo provvederà a diffonderli verso gli utenti durante la giornata successiva. Le **Broadcast Ephemerides**, sebbene trasmesse dai satelliti in tempo reale, sono in realtà il risultato di previsioni effettuate dalla rete di controllo nelle 24 ore precedenti.

Le stazioni di upload (ULS) sono situate in corrispondenza delle 3 stazioni monitor Ascension Island, Diego Garcia e Kwajalein e sono dotate di antenne trasmittenti del diametro di circa 10 metri; l'upload avviene, mediamente, ogni 8 ore.

Come si vedrà meglio nel seguito, il principio di funzionamento del GPS è centrato sulla possibilità di misurare in modo accurato il tempo; per questo motivo la stazione Master provvede ad utilizzare una serie di orologi atomici che costituiscono il riferimento temporale comune per la componente terrestre e satellitare del intero sistema.

Un'alternativa alle effemeridi trasmesse sono le **Precise Ephemerides** : con un ritardo di circa 4-8 settimane dopo la raccolta dei dati da parte delle stazioni di controllo, il Naval Surface Weapon Center (NSWC), in collaborazione con il Defence Mapping Agency (DMA), calcola le effemeridi precise ottenute considerando le informazioni raccolte in 8 giorni da 10 stazioni sparse sul globo (le 5 stazioni di controllo più altre in

Australia, Ecuador, Inghilterra, Argentina e Bahrein); infine, attraverso l'utilizzo di un complesso algoritmo, viene effettuata una stima più accurata dell'orbita.

Le stime di orbite e offset degli orologi effettuate dalla rete di controllo e quindi comunicate in broadcast dai satelliti mediante i messaggi di navigazione, sono affette da errori medi rispettivamente di 3 m e 10 ns (tenendo conto della velocità della luce c , questo offset equivale a circa 3 m di errore nella stima della distanza fra satellite e ricevitore).

Attualmente diversi enti di ricerca associati all'IGS (**I**nternational **G**PS **S**ervice for **G**eodynamics) calcolano e rendono disponibili le effemeridi precise con una settimana di ritardo; l'accuratezza risulta migliore di circa due ordini di grandezza rispetto alle effemeridi broadcast.

3.2.3 Il segmento di utilizzo

Questa sezione è costituita dagli utenti (civili e militari) del servizio GPS; un ricevitore GPS demodula i segnali emessi dai satelliti GPS allo scopo di stimare, in tempo reale, la propria posizione in un opportuno sistema di riferimento tridimensionale; una funzionalità aggiuntiva, che consente di ottenere misure di posizione più accurate, prevede che il ricevitore possa raccogliere dati per una compensazione in tempo reale o a posteriori [differential GPS]. I ricevitori presenti sul mercato utilizzano diverse tecniche di decodifica del segnale ricevuto, alle quali corrispondono, in generale, diversi livelli nell'accuratezza della stima di posizione; i più moderni e diffusi utilizzano entrambe le portanti radio L1 ed L2 del segnale GPS (vedi paragrafo successivo).

3.3 La struttura del segnale GPS

3.3.1 Introduzione

NAVSTAR GPS è un **one-way ranging system**: i segnali utili, destinati agli utenti del servizio, sono originati dai satelliti e unidirezionali; il ricevitore si limita ad applicare a quanto ricevuto gli algoritmi necessari alla stima di posizione. Come si avrà

modo di illustrare nel seguito, la stima della distanza satellite - terminale GPS assume un ruolo centrale nella procedura di posizionamento; in realtà, questa stima può essere effettuata anche utilizzando un **two-way ranging system**, tipico dei distanziometri elettro-ottici; la scelta di utilizzare nel GPS sistemi ad una via, sebbene abbia reso più complesse le procedure, ha evitato di installare trasmettitori costosi ed ingombranti nei terminali GPS.

Nel seguito verrà analizzata nel dettaglio la struttura del segnale trasmesso da un generico satellite GPS.

Tutti i satelliti trasmettono utilizzando due diverse portanti nella banda L, entrambe multiple della frequenza fondamentale $f_0=10,23$ MHz generata dall'utilizzo concorrente degli oscillatori al cesio ed al rubidio installati sui satelliti; indicando con L_1 ed L_2 le 2 portanti e con f_1 ed f_2 le due frequenze corrispondenti, si ha:

$$f_1 = 154 \cdot f_0 = 1.575,42\text{MHz} \Leftrightarrow \lambda_1 \cong 19\text{cm};$$

$$f_2 = 120 \cdot f_0 = 1.227,60\text{MHz} \Leftrightarrow \lambda_2 \cong 24\text{cm};$$

dove $\lambda_i = \frac{c}{f_i}$ $i \in \{1,2\}$ indica la lunghezza d'onda della radiazione elettromagnetica di frequenza f_i .

Le portanti L_1 ed L_2 sono modulate dai seguenti segnali:

- **P code** (Protected or Precision);
- **C/A code** (Clear Access or Course Acquisition);
- **D code** (Navigation Data)

I tre codici corrispondono, in definitiva, a sequenze di bit; il codice D trasporta messaggi informativi mentre i codici P e C/A sono costituiti da sequenze pseudocasuali di bit, ossia sequenze generate senza alcuna logica apparente (come nel caso del rumore) ma in realtà perfettamente note; tali sequenze risultano addirittura periodiche con periodo 1 ms per il codice C/A e 7 giorni per il codice P. Per questi motivi i codici P e C/A vengono definiti codici di tipo **PRN** (Pseudo Noise Random); ogni satellite utilizza una propria sequenza PRN distintiva ed ortogonale alle altre; sono utilizzate 38 sequenze diverse numerate da 0 a 37. Si noti che ogni satellite utilizza le stesse

frequenze trasmissive: lo schema d'accesso multiplo adottato è la tecnica **CDMA** (Code Division Multiple Access) che sfrutta l'ortogonalità dei codici sopra evidenziata.

3.3.2 Il codice P (Protected or Precision)

È costituito da una sequenza di bit a frequenza 10,23 MHz che si ripete periodicamente dopo circa 38 settimane e che indicheremo **sequenza P**; l'intera sequenza non viene trasmessa da alcun satellite, in quanto viene suddivisa in segmenti di durata una settimana ed ognuno di essi è assegnato ad un satellite: il numero massimo di questi ultimi dovrebbe, pertanto, essere 38; in realtà 1 segmento non è utilizzato e 5 sono utilizzati dalle stazioni a terra: solo 32 segmenti sono quindi assegnati effettivamente ai satelliti.

Al termine di ogni settimana, su tutti i generatori di codice P nei satelliti viene effettuata la procedura di **reset** al corrispondente punto iniziale dell'intera sequenza P; se così non avvenisse, il satellite inizierebbe a generare il segmento successivo assumendo l'identità di un altro satellite; l'istante di reset coincide con l'istante UTC 0^h di ogni Domenica.



Figura 3.17 Esempio codice trasmesso da un satellite; nel caso del codice C/A la sequenza è formata da 1023 bit per una durata di 1 msec; nel caso del codice P la sequenza è di ~ 15 milioni di bit per una durata di una settimana.

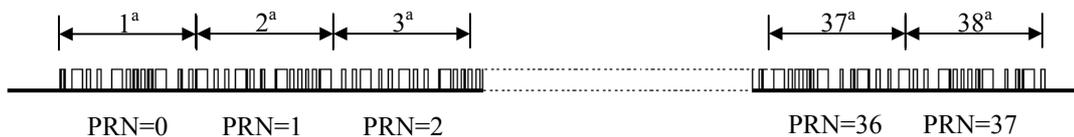


Figura 3.18 Sequenza P completa

Il codice P modula entrambe le portanti e consente di raggiungere la massima precisione nella procedura di posizionamento.

Inizialmente tale codice poteva essere utilizzato soltanto in appositi apparati militari e da utenti autorizzati; tale protezione fu tuttavia aggirata da vari ricercatori fin dall'inizio degli anni ottanta per cui è stato introdotto a bordo dei satelliti operativi (BLOCKII, IIA e IIR) la possibilità, su comando, di generare un codice Y criptato, in luogo del codice P, utilizzabile soltanto dagli apparati che dispongono degli opportuni decodificatori.

3.3.3 Il Codice C/A (Clear Access or Coarse Acquisition)

Ha una frequenza $\frac{f_0}{10} = 1,023\text{Mbps}$ ed un periodo di ripetizione pari a 1 msec; **il**

codice C/A modula soltanto la portante L1 e risulta di più facile ricezione essendo formato da soli 1023 bit; è unico per ciascun satellite, è fruibile da tutti i ricevitori e consente una precisione nettamente inferiore a quella ottenibile con il codice P [“acquisizione grossolana”]

3.3.4 Il codice D (Navigation Data)

Contiene il Messaggio di Navigazione, trasmesso a bassa velocità: 50 bps. **Il codice D modula entrambe le portanti** e contiene:

- le effemeridi del satellite ossia la posizione esatta per un dato istante ed i parametri utili per il calcolo delle sue successive posizioni; i dati vengono trasmessi con continuità e vengono aggiornati dal satellite ogni ora;
- l'istante di trasmissione del messaggio;
- le correzioni per l'orologio di bordo;
- le correzioni per i ritardi causati dalla rifrazione ionosferica e troposferica;
- varie informazioni sullo stato dei satelliti (il cosiddetto health);
- l'almanacco, ossia le effemeridi approssimate di tutti gli altri satelliti.

L'intero messaggio (**frame**), lungo 1500 bit e pertanto trasmesso in $\frac{1.500}{50} = 30$ sec, è

suddiviso in cinque blocchi di 300 bit (**subframe**) la cui struttura è mostrata nella figura

successiva. Ogni blocco, trasmesso in 6 sec, è preceduto da un header comune, generata dallo stesso satellite e costituita dai campi:

- **TLM (TeLeMetry word)**, utilizzato dal sistema di controllo terrestre, essenzialmente per verificare il corretto caricamento dei dati durante la fase di upload descritta in precedenza.
- **HOW (HandOver Word)**

Il payload trasporta dati generati dal sistema di controllo terrestre e scaricati nei buffer dei satelliti dalle stazioni di aggiornamento.

HOW è un numero progressivo che moltiplicato per 6 fornisce i secondi trascorsi all'inizio della settimana, ossia fornisce il **GPS time o tempo di sistema**.

L'ultimo subframe contiene i dati dell'**almanacco**, ossia i dati di posizione di tutti i satelliti presenti nel sistema; essi sono una versione ridotta e approssimata dei dati trasmessi da ogni singolo satellite nei subframes precedenti. L'almanacco consente al ricevitore di determinare velocemente la posizione di tutti i satelliti e quindi di sapere quali sono visibili e utilizzabili. Consente inoltre di selezionare fra i satelliti utili quali sono quelli meglio disposti da un punto di vista geometrico per avere la massima precisione nel **FIX**. La trasmissione dell'intero almanacco richiede più spazio di quella disponibile, pertanto esso è stato frazionato in 25 blocchi (**pagine**) trasmessi consecutivamente ad intervalli di 6 secondi. La trasmissione dell'intero almanacco richiede pertanto $25 \cdot 30 \text{ sec} = 12,5$ minuti di tempo. Ogni blocco viene trasmesso utilizzando l'ultimo subframe, tuttavia, essendo attualmente i satelliti 27, una parte dei dati vengono trasmessi anche nel penultimo subframe, che inizialmente era stato riservato alla trasmissione di messaggi in chiaro. Normalmente i ricevitori GPS mantengono in memoria l'intero almanacco anche quando vengono spenti, in modo da averlo già disponibile quando l'apparato viene acceso nuovamente. Dopo i calcoli precedentemente descritti, il ricevitore è in grado di sincronizzarsi velocemente sui segnali dei satelliti desiderati.

Nel caso di lunghi periodi di inattività o dopo la prima accensione (subito dopo l'installazione) o dopo una perdita di dati per qualche avaria, il ricevitore rimane inutilizzabile per diversi minuti fino alla completa ricezione dell'almanacco.

Fra gli altri dati, ogni “pagina” contiene il codice PRN del satellite cui essa si riferisce ed un indicatore, definito **Space Vehicle Health**, che fornisce all’utente informazioni preliminari sull’utilizzabilità del satellite prima che il ricevitore tenti di acquisirne i segnali (è tutto regolare se il valore 0, altrimenti esso può assumere diversi valori ad ognuno dei quali corrisponde un particolare malfunzionamento).

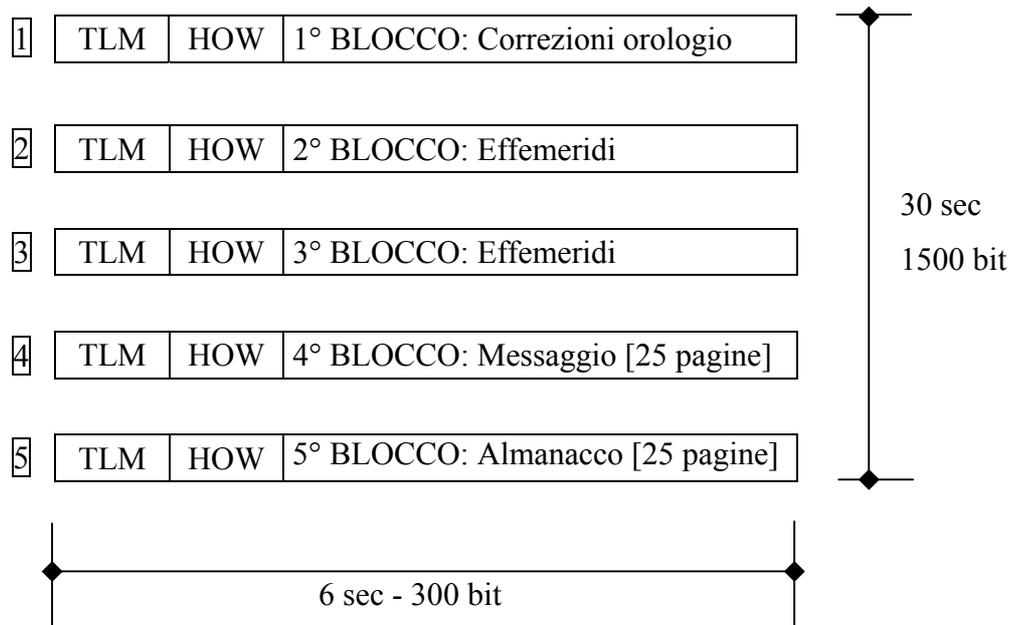


Figura 3.19 Struttura del messaggio di navigazione [codice D]

3.3.5 Modulazione

Come si è anticipato in precedenza, tutti i satelliti utilizzano le stesse frequenze per trasmettere simultaneamente i codici C/A, P e D. In particolare:

- la portante L1 è modulata dai codici C/A, P e D;
- la portante L2 è modulata soltanto dai codici P e D.

Il segnale risultante:

$$S(t) = S_{L1}(t) + S_{L2}(t)$$

dove sono esplicitamente indicate le 2 componenti a frequenza f_1 e f_2 :

$$S_{L1}(t) = P(t)D(t)\cos(2\pi f_1 t) + C(t)D(t)\sin(2\pi f_1 t)$$

$$S_{L2}(t) = P(t)D(t)\cos(2\pi f_2 t)$$

Il segnale trasmesso dal satellite risulta pertanto costituito da 3 componenti sovrapposte:

- la componente a frequenza f_1 che trasporta i codici P e D;
- la componente a frequenza f_1 e sfasata di 90° che trasporta i codici C/A e D;
- la componente a frequenza f_2 che trasporta i codici P e D.

Ciascuna componente viene modulata simultaneamente (ma separatamente da un punto di vista logico) ed i corrispondenti segnali modulati sommati in fase; il formato di modulazione numerica utilizzato è **BPSK (Binary Phase Shift Keying)**: la fase portante è funzione del bit da trasmettere e resta costante durante l'n-esimo intervallo di segnalazione; si tratta di uno schema di modulazione lineare senza memoria che consente l'utilizzo di modulatori e demodulatori semplici ma che risulta particolarmente sensibile alle non linearità dei componenti.

Indicando con:

- φ_k la fase associata al k-esimo bit da trasmettere
- $u_T(t) = \text{rect}_T(t)$ l'impulso base di durata T e ampiezza unitaria;
- $\varphi(t) = \sum_{k=1}^L u_T(t - kT) \cdot \varphi_k$ l'espressione della fase istantanea; L indica la lunghezza della sequenza da trasmettere.
- $v(t) = A \cos(2\pi ft + \varphi(t)) \quad t \in [0, LT]$ l'espressione del segnale modulato; A indica l'ampiezza della portante sinusoidale supposta a fase iniziale nulla

si ha, per la modulazione BPSK:

- $\varphi_k \in \{0, \pi\}$: Ogni volta che il segnale modulante PRN passa dallo stato logico "0" a "1" o viceversa, la fase della portante viene invertita di 180° ;
- $v(t) = \sum_{k=1}^L u_T(t - kT) \cdot \cos \varphi_k \cdot \cos(2\pi ft) = I(t) \cdot \cos(2\pi ft)$
- $I(t) = \sum_{k=1}^L x_k \cdot u_T(t - kT) \cdot \cos(2\pi ft)$ con $x_k \in \{+1, -1\}$ in funzione del bit da trasmettere

- La modulazione BPSK equivale ad una modulazione d'ampiezza con portante sinusoidale nella quale le 2 possibili ampiezze valgono ± 1 ; questa osservazione semplificherà notevolmente la costruzione di un demodulatore BPSK.

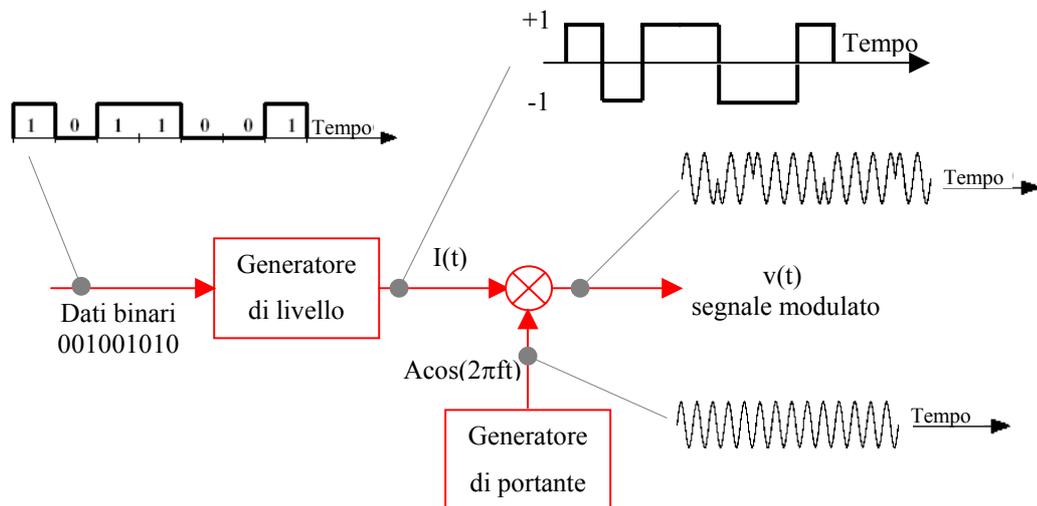


Figura 3.20 Modulatore BPSK

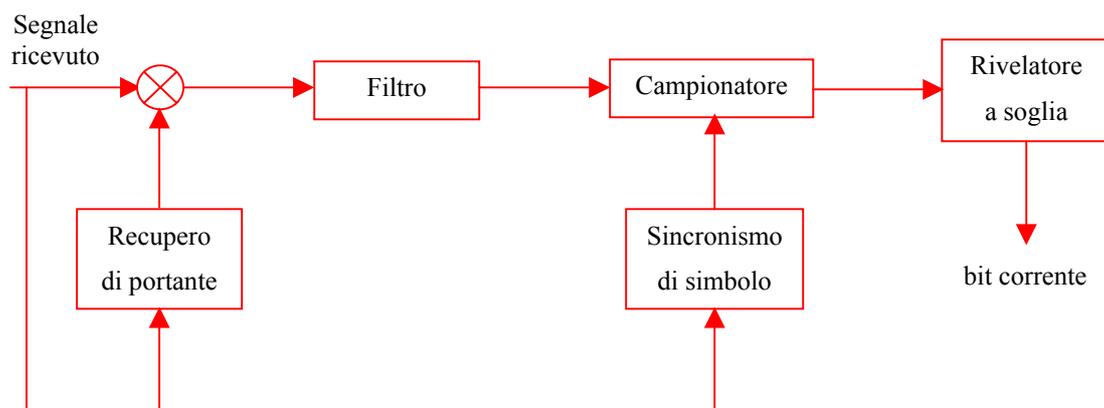


Figura 3.21 Demodulatore BPSK [a massima verosimiglianza]

Come già sottolineato in precedenza, lo schema d'accesso multiplo adottato per la risorsa radio è CDMA: sulla stessa frequenza trasmettono contemporaneamente tutti i satelliti senza che essi interferiscano fra di loro; nei vecchi sistemi di radionavigazione si evitavano le interferenze facendo trasmettere le stazioni terrestri su diverse frequenze (FDMA) oppure in istanti diversi (TDMA).

Da quanto esposto appare evidente la complessità della struttura di segnale GPS; le ragioni sono molteplici:

- GPS è un sistema studiato per supportare un numero virtualmente illimitato di utenti; per questo motivo un sistema che presuppone comunicazione tra ricevitore e satellite sarebbe troppo complicato e difficilmente scalabile; non bisogna poi dimenticare che il metodo è nato come metodo militare e quindi motivi di sicurezza impongono che l'utilizzatore non comunichi la sua posizione inviando segnali al satellite; d'altra parte un metodo one-way richiede l'impiego di due orologi: uno a bordo del satellite, l'altro nella stazione ricevente;
- l'utilizzo di più satelliti contemporaneamente è fondamentale per ottenere il posizionamento in tempo reale; è necessario pertanto poter identificare ogni satellite attraverso il segnale ricevuto senza ambiguità e ciò può avvenire solo mediante l'assegnazione di un diverso codice ad ogni satellite;
- la richiesta di precisioni inferiori ai 10 metri per il posizionamento in tempo reale suggerisce l'utilizzo di un codice ad alta frequenza come il codice P; utilizzando codici a bassa frequenza la lunghezza d'onda del segnale potrebbe non fornire la risoluzione necessaria;
- l'esistenza di un gruppo di utenti non militari o appartenenti ad eserciti non alleati richiede l'esistenza di due diversi tipi di segnale, in modo da consentire precisioni differenti nel posizionamento: i codici Y e C/A;
- per compensare l'effetto della ionosfera sul segnale, poiché questo è selettivo in frequenza, si utilizzano due diverse portanti [diversità in frequenza].

3.4 Sistema di riferimento temporale [GPS time]

La **Master Station Control** (MSC) di Colorado Spring, come si è già detto precedentemente, ha la funzione di assicurare un riferimento temporale comune attraverso una serie di orologi atomici; il grado di precisione assicurato è tale da consentire un disallineamento al più di 1 secondo in 32 secoli.

Il riferimento temporale è definito **tempo del sistema GPS (GPS Time)** e differisce dall'UTC (Universal Coordinated Time) di una quantità nota. L'ora di sistema deve essere di tipo uniforme e non può quindi essere uguale all'UTC che viene "ritoccata" una o due volte all'anno aggiungendo o sottraendo 1 secondo; tale correzione comporterebbe notevoli problemi nella regolarità di funzionamento del sistema per gli scopi della navigazione. Tempo e frequenza a bordo dei satelliti sono forniti da tre o quattro orologi atomici (al Cesio ed al Rubidio) sincronizzati giornalmente dal centro di controllo terrestre a causa della lenta deriva.

3.5 Degradazione artificiale della precisione

La precisione offerta dal sistema GPS, utilizzando anche il solo codice C/A, può risultare nettamente superiore a quella prevista dai progettisti del sistema. Sulla base delle specifiche del Federal Radionavigation Plan del governo USA, sono previsti due livelli di utilizzo:

- **Standard Positioning Service** (SPS) utilizza il codice C/A sulla portante L1 ed è caratterizzato da un'accuratezza di 100 m nel piano orizzontale;
- **Precise Positioning Service** (PPS) utilizza il codice P su entrambe le portanti L1, L2 con un'accuratezza di 18 m nel piano orizzontale;

Allo scopo di fornire il servizio più accurato soltanto ai terminali autorizzati e per proteggere gli utilizzatori militari dalla possibilità di essere ingannati da false trasmissioni con segnali simili a quelli del GPS, il DoD (Departement of Defence degli USA) introdusse, a bordo dei satelliti del secondo blocco, dispositivi in grado di

produrre una degradazione sui segnali, in modo da alterare significativamente la precisione ottenibile nel posizionamento; questi possono essere attivati o disattivati dal Centro di Controllo su richiesta del DoD e senza alcuna preventiva informazione agli utilizzatori; il protocollo utilizzato assicura che gli errori introdotti artificialmente sul segnale risultino noti soltanto agli utenti autorizzati e che comunque non producano, per gli altri utenti, errori superiori ai 100 m. L'utilizzo del GPS differenziale ha comunque consentito di aggirare anche queste ultime protezioni, per cui si presume che in futuro il GPS perderà il suo carattere militare per diventare un sistema interamente libero.

Le due tecniche utilizzate con questa finalità sono:

- **Selective Availability (SA)** o “Disponibilità selettiva”, consiste nell'introduzione di appositi errori sulle effemeridi dei satelliti da parte della Stazione di controllo terrestre; soltanto gli utenti autorizzati sono in grado di prevederne, istante dopo istante, i valori e quindi eliminarne le conseguenze. La SA produce un'oscillazione complessiva della posizione indicata dal ricevitore di circa 100 m in un periodo variabile da 3-4 min a 60 min; nel 2000, con sei anni di anticipo rispetto al programma, gli Stati Uniti hanno rimosso la SA sul segnale GPS con notevoli implicazioni tecniche e di mercato.
- **Anti-Spoofing (AS)** o “Anti-Imbroglione”, agisce direttamente sul codice P e su entrambe le portanti L1 e L2; consiste nell'introduzione, in luogo del codice P, di un nuovo codice criptato denominato **codice Y**. Tale limitazione ha effetto soltanto sugli utenti, civili o militari, autorizzati a servirsi del codice P, che dovranno munirsi di un nuovo ricevitore o dovranno sostituire i programmi di decifrazione del codice P con quelli del codice Y.

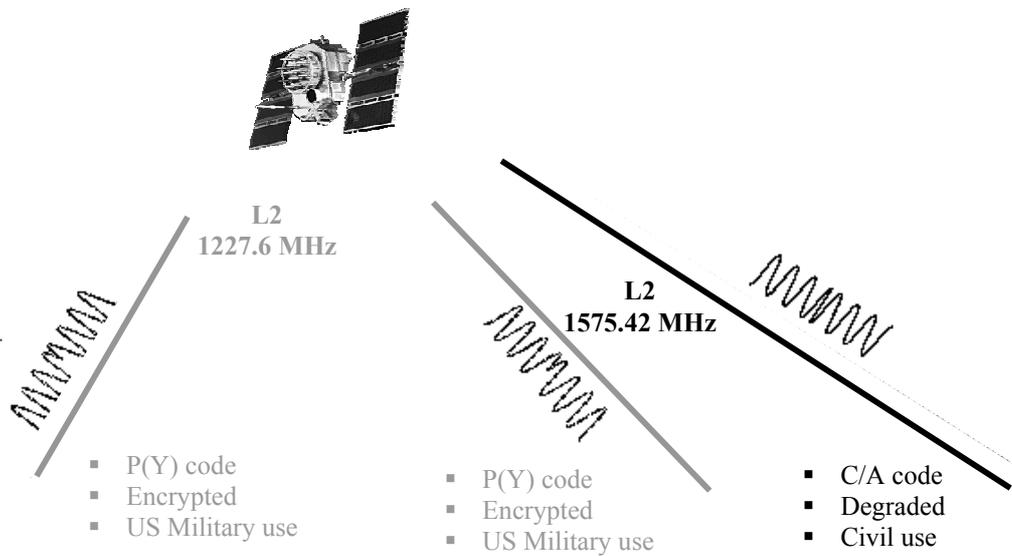


Figura 3.22 Segnali trasmessi da un satellite GPS: il segnale portante nella banda L1 che trasporta il codice C/A è destinato all'utenza civile; l'accesso alla banda L2 ed al codice P(Y) nella banda L1 è consentito solo agli utenti autorizzati dal DoD a causa del processo di cifratura

3.6 Ricezione ed utilizzo dei segnali

Un ricevitore GPS è costituito essenzialmente da quattro componenti: l'antenna, il ricevitore, il sistema di calcolo ed i dispositivi per l'introduzione e la visualizzazione dei dati. Le funzioni fondamentali realizzate da un ricevitore possono essere così sintetizzate:

- Selezionare quattro satelliti in base ai dati contenuti nell'Almanacco in modo che risultino visibili e distribuiti opportunamente in azimuth al fine di consentire la migliore determinazione della posizione;
- Individuare i segnali dei satelliti selezionati decodificandoli attraverso il rispettivo codice PRN;

- Decodificare i dati di navigazione contenuti nel codice D e memorizzarli nei buffer di memoria;
- Misurare i ritardi nei tempi d'arrivo dei segnali dai satelliti e calcolare le corrispondenti stime di distanza;
- Calcolare la posizione del ricevitore e l'ora del sistema (GPS Time).

Le ultime due funzionalità saranno oggetto d'analisi nei paragrafi successivi.

Le suddette funzioni possono essere realizzate in diversi modi dipendenti dall'uso cui è destinato il ricevitore e dagli obiettivi di precisione prefissati; in alcuni scenari (es. armi) un ricevitore dovrà essere in grado effettuare calcoli di alta precisione e con continuità; in altri un ricevitore potrà avere una precisione inferiore ed un aggiornamento dei dati di posizione molto meno frequente; nel primo caso si dice che il ricevitore deve avere un'**alta dinamica**, nel secondo caso una **bassa dinamica**.

Per questo motivo i ricevitori GPS presenti sul mercato possono essere di tre tipi:

- **Ricevitori sequenziali**, dotati fisicamente di un solo canale che segue i quattro satelliti selezionati per la determinazione della posizione, uno dopo l'altro e per un periodo di tempo relativamente lungo ($1\div 2$ sec); le misure di distanza sono quindi intervallate, per cui il calcolo della posizione si effettua combinando una distanza effettivamente misurata con tre distanze stimate. Tali tipi di ricevitori sono piuttosto economici, ma, considerato il tempo piuttosto lungo ($4\div 8$ sec) richiesto per l'osservazione dei quattro satelliti, non sono utilizzabili nei casi in cui la velocità di elaborazione è cruciale.
- **Ricevitori multicanali**. Sono dotati di più canali ($10\div 12$) che seguono contemporaneamente altrettanti satelliti; ogni canale "aggancia" un satellite, demodula il segnale ed effettua la misura della distanza. Un microprocessore provvede successivamente a combinare tutte le osservazioni (simultanee) e a fornire la posizione ed altri dati istantanei come velocità e direzione; tali ricevitori sono più costosi e si prestano ad essere usati in tutti i casi dove è richiesta un'alta dinamica. Il ricevitore GPS utilizzato nelle esperienze di misura per questo lavoro di tesi è un ricevitore multichannel a 12 canali.

- **Ricevitori multiplexed.** Sono dotati di un solo canale e seguono sequenzialmente i satelliti selezionati effettuando però una loro completa scansione in meno di 20 msec che è la durata di un bit nel messaggio di navigazione (trasmesso a 50 bps). Con questi ricevitori, analogamente a quelli multicanali, i dati di navigazione sono ricevuti con continuità, senza le interruzioni tipiche dei ricevitori sequenziali. Le caratteristiche sono intermedie a quelle dei due precedenti tipi e con costi piuttosto contenuti.

3.7 Il calcolo delle distanze

Il codice C/A, generato dal satellite, è sincronizzato con il GPS Time e si ripete periodicamente ogni millisecondo; arriva all'utilizzatore con un ritardo $t = R/c$ dove R è la distanza fra il satellite ed il ricevitore e $c = 300.000$ Km/sec è la velocità di propagazione delle onde elettromagnetiche. Questa è l'osservazione chiave sulla quale si sviluppa l'intero sistema GPS: lo **pseudorange** è una misura della distanza tra satellite ed antenna del ricevitore, riferita alle epoche di trasmissione e ricezione del segnale

Il ricevitore è dotato di un generatore locale di codice C/A analogo a quello installato nei satelliti e come questi dovrebbe risultare sincronizzato con il GPS Time; tuttavia tale sincronismo non è perfetto a causa essenzialmente degli orologi utilizzati nei terminali, dotati di normali oscillatori al quarzo e caratterizzati da una deriva di cui occorre tenere conto nella misura dei tempi; le misure di distanza (**range**) fatte dal ricevitore sono quindi affette dall'offset dell'orologio del ricevitore stesso: sono pertanto note come misure di **pseudorange**.

I due segnali, quello proveniente dal satellite e quello generato internamente al ricevitore, vengono confrontati mediante **autocorrelazione**; il dispositivo elettronico utilizzato è il **correlatore** e produce in uscita un segnale $V(\tau)$ che è funzione dello sfasamento fra i due segnali in ingresso; più precisamente $V(\tau)$ assume il valore massimo con i due segnali perfettamente allineati ed un valore nullo o comunque molto piccolo nel caso di segnali sfasati (disallineati) fra di loro; nella figura successiva si può

osservare come la tensione in uscita dal correlatore comanda il dispositivo che genera il ritardo stesso in modo da ottenere automaticamente il massimo di $V(\tau)$. Tale configurazione ad anello consente il continuo inseguimento (**tracking**) del segnale; successivamente alla **fase di aggancio (lock on)**, infatti, in corrispondenza del massimo di $V(\tau)$, **qualsiasi variazione di distanza prodotta dal moto relativo fra satellite e ricevitore produce uno sfasamento nei codici in ingresso al correlatore che viene annullato immediatamente dal circuito a controreazione**. In tal modo la misura del ritardo, e quindi della distanza, è continua. I ritardi temporali misurati rappresentano lo sfasamento che il ricevitore deve introdurre nel proprio generatore di codice C/A per ottenere un perfetto sincronismo con i codici provenienti dai satelliti, tali ritardi sarebbero corretti se tutti gli orologi (dei satelliti e del ricevitore) fossero in perfetto sincronismo, in realtà ciò non avviene per cui è necessario sommare agli Pseudo Time (i Tempi misurati) gli errori negli orologi dei satelliti e del ricevitore. I primi sono determinati dalle stazioni di controllo terrestre e fanno parte del messaggio di navigazione (codice D), i secondi vanno determinati analiticamente considerandoli un'incognita del sistema.

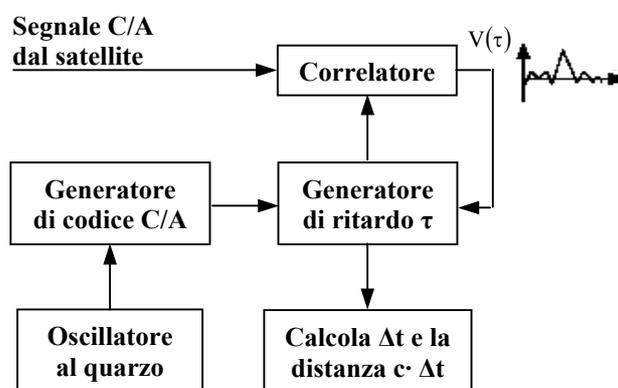


Figura 3.23 Misura del ritardo con correlatore a controreazione. Il codice C/A generato internamente al ricevitore viene confrontato con quello ricevuto dal satellite; il segnale interno viene ritardato fino a che i due segnali non sono perfettamente allineati. L'allineamento si ha quando la tensione in uscita dal correlatore raggiunge il massimo, il ritardo τ introdotto consente di calcolare il tempo di viaggio t dei segnali e quindi la distanza del satellite pari a $c \cdot t$.

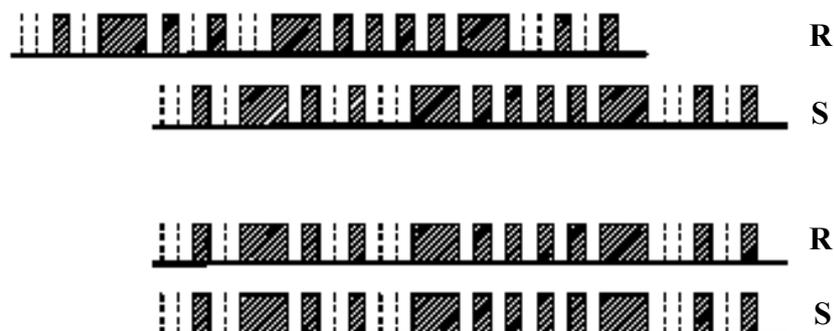


Figura 3.24 I segnali R sono i codici C/A generati internamente al ricevitore, mentre i segnali S sono quelli in arrivo dal satellite. Nel primo caso i due segnali non sono allineati, pertanto in uscita dal correlatore il segnale è minimo; nel secondo caso i segnali sono perfettamente allineati ed il segnale $V(\tau)$ risulta massimo.

3.8 Principio di posizionamento attraverso misure di pseudorange

3.8.1 Generalità

Nei paragrafi precedenti si è illustrata la struttura del segnale GPS; in sintesi, si è visto che tale segnale trasporta i dati orbitali per il calcolo della posizione del satellite ed implicitamente un'informazione di ritardo temporale per la determinazione dell'istante di trasmissione del segnale stesso; in base a questo ritardo il ricevitore è in grado di stimare la distanza dal satellite.

Resta pertanto univocamente determinato il **luogo di posizione**, ossia una **sfera** con il centro nella posizione occupata dal satellite nell'istante di emissione del segnale ed il raggio pari alla distanza calcolata; tale luogo interseca la superficie terrestre individuando una circonferenza luogo dei punti nei quali può trovarsi l'utilizzatore terrestre.

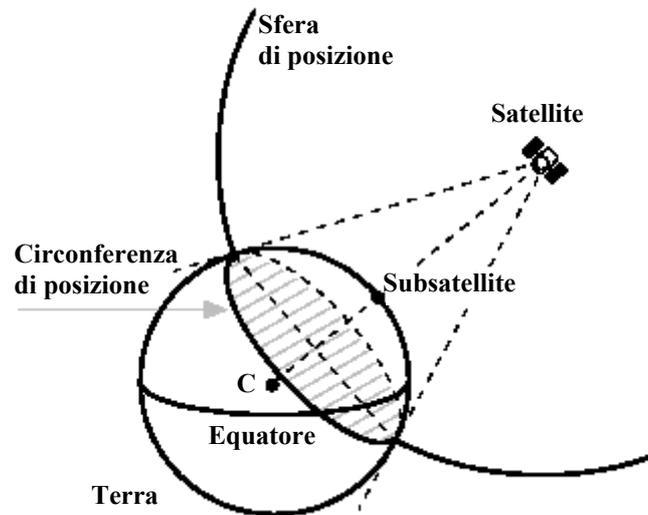


Figura 3.25 Il luogo di posizione individuato da una misura di distanza è un sfera

Due misure di distanza, disponibili utilizzando due satelliti, individuano due circonferenze che si intersecano in due punti di cui uno è certamente la posizione dell'osservatore; l'ambiguità fra i due punti può essere eliminata con la posizione stimata del ricevitore (di solito uno dei due punti è inverosimile); considerando come ulteriore incognita anche la quota sono necessarie tre osservazioni che individuano tre sfere, la cui intersezione individua un volume entro il quale si trova il ricevitore. Nel caso bidimensionale, ossia nel caso in cui la quota è nota, le tre misure individuano un triangolo, praticamente piano, la cui area è funzione degli errori di misura e della distribuzione in azimut dei satelliti.

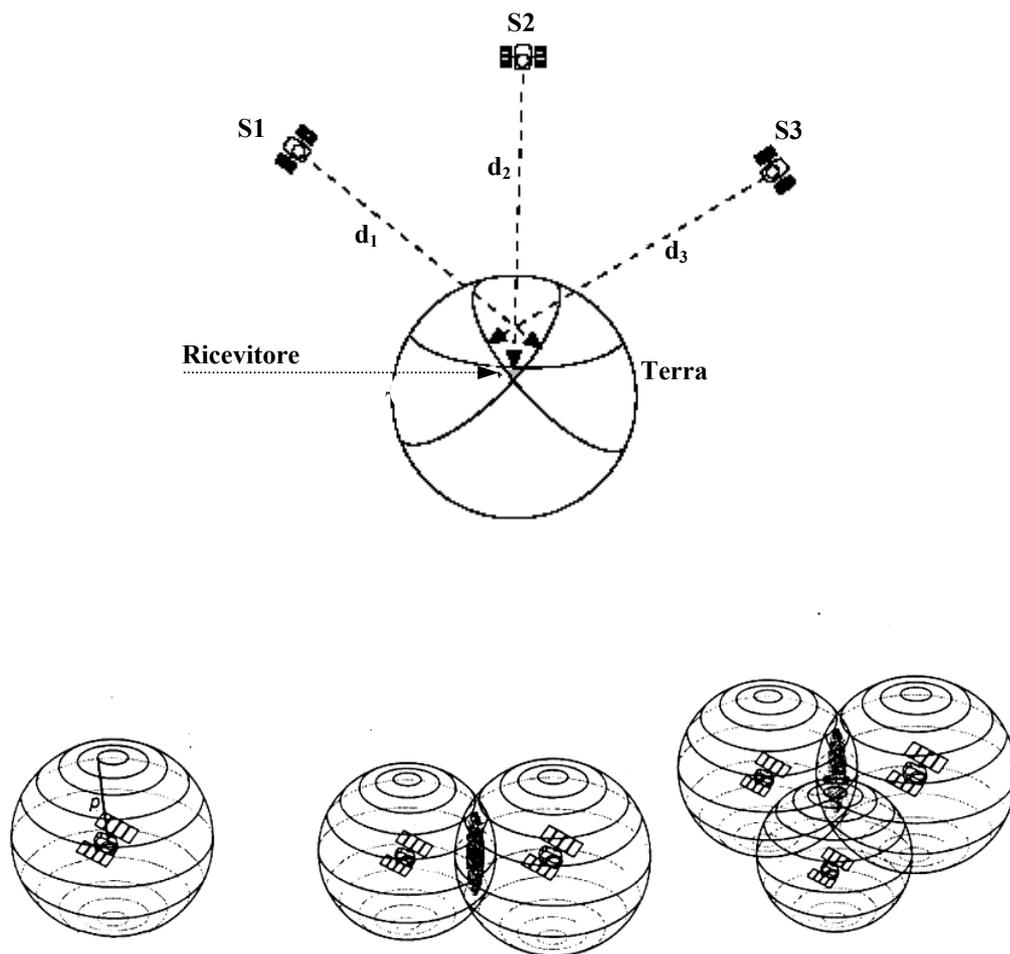


Figura 3.26 Principio di posizionamento attraverso tre misure di distanza

Vale appena la pena di osservare che un errore di soli $0,1 \mu\text{sec}$ nella misura del ritardo temporale equivale ad un errore di $3 \cdot 10^8 \cdot 0,1 \cdot 10^{-6} = 30$ metri nella misura delle distanze; questo a riprova della già evidenziata necessità di disporre di orologi di grande precisione nell'ambito dell'intero sistema.

Per limitare i costi degli apparati riceventi terrestri, questi sono equipaggiati con orologi di precisione molto inferiore rispetto a quelli installati nei satelliti; quanto segue, tuttavia, mostrerà che in effetti tale precisione può essere recuperata cercando di determinare l'errore rispetto al segnale di tempo fornito dai satelliti; l'errore

dell'orologio rappresenta pertanto un'ulteriore incognita che può essere determinata con una quarta osservazione; si risolve in definitiva un sistema di quattro equazioni con le quattro incognite: latitudine, longitudine, quota ed offset dell'orologio del ricevitore.

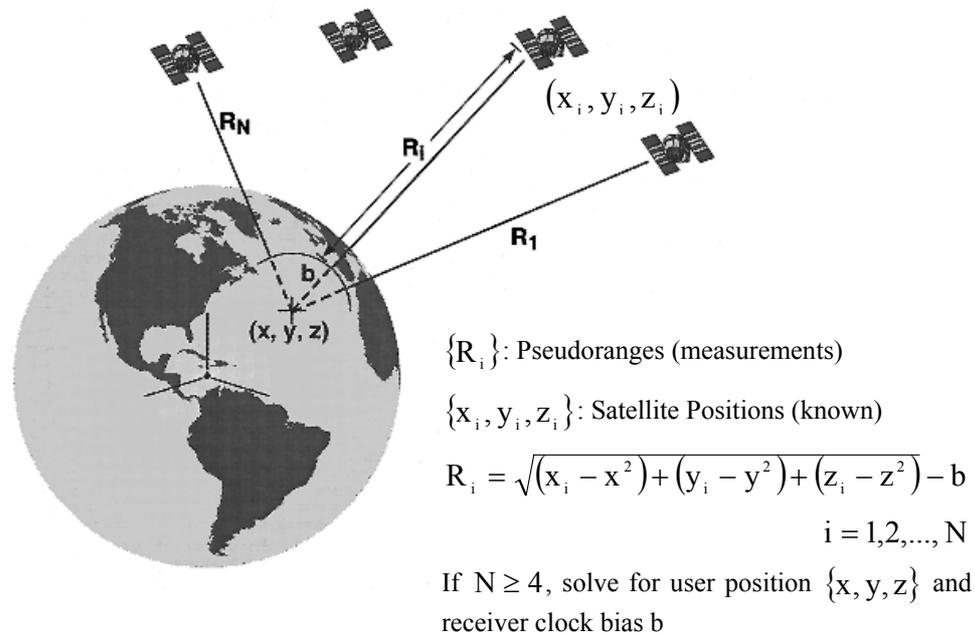


Figura 3.27 Le misure di pseudorange effettuate da uno stesso ricevitore sono affette da un valore comune di offset temporale che si traduce in un errore comune sulle distanze stimate R_i (bias b)

3.8.2 Algoritmo di calcolo

A partire dal messaggio di navigazione (Codice D) il ricevitore ricava i parametri orbitali dei satelliti; calcola quindi la posizione istantanea dei satelliti rispetto ad un sistema riferimento convenzionale terrestre (**CTRF**) [cartesiano - vedi paragrafo 3.9]; indicando con X_s , Y_s e Z_s le coordinate del satellite e con X , Y , Z le coordinate di un generico osservatore O di quota Q , si ottiene:

$$(X - X_s)^2 + (Y - Y_s)^2 + (Z - Z_s)^2 = R_v$$

che rappresenta l'equazione della **sfera di posizione** con $R_v = R + c \cdot e$; R indica la misura di pseudorange (distanza osservata ricevitore - satellite), e l'offset temporale, c la velocità della luce.

Disponendo di quattro misure di distanza si imposta un sistema di quattro equazioni in quattro incognite la cui soluzione fornisce X , Y , Z , e ; successivamente è possibile trasformare le coordinate cartesiane in coordinate geografiche: indicando con ρ la distanza \overline{CO} del ricevitore dal centro della Terra e con riferimento alla figura seguente si hanno:

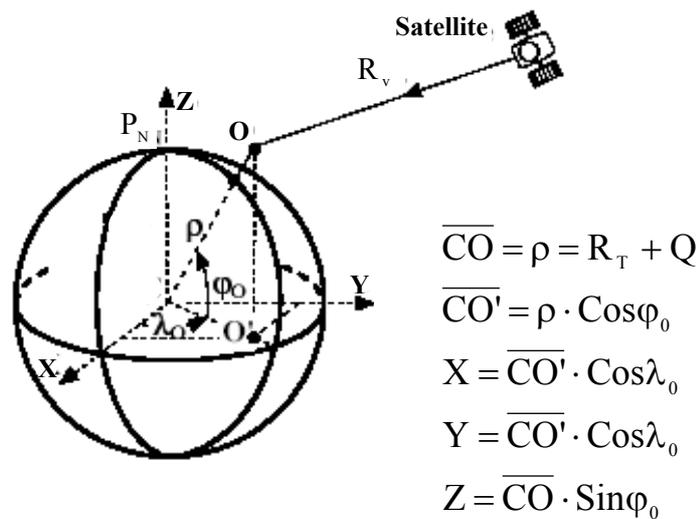


Figura 3.28 Posizione relativa del satellite e del ricevitore

- $X = \rho \cdot \cos \phi_0 \cdot \cos \lambda_0$
- $Y = \rho \cdot \cos \phi_0 \cdot \sin \lambda_0$
- $Z = \rho \cdot \sin \phi_0$
- $\rho = \sqrt{X^2 + Y^2 + Z^2}$
- $\sin \phi_0 = \frac{Z}{\rho}$
- $\tan \lambda_0 = \frac{Y}{X}$

La distanza ρ calcolata è comprensiva anche della quota Q del ricevitore; nell'ipotesi di terra sferica con raggio R_T si ottiene:

$$\square \quad Q = \rho - R_T$$

3.9 Errori

3.9.1 Generalità

Il sistema GPS è soggetto a differenti tipi di errori che influenzano la distanza misurata e quindi la posizione stimata; molti di essi possono essere ridotti d'ampiezza utilizzando delle procedure compensative ottenute dopo un'attenta valutazione delle cause e dopo una serie di rilievi effettuati in particolare durante la fase sperimentale del sistema.

Fondamentalmente si tratta di:

- \square errori dipendenti dai satelliti, fra cui anche la Selective Availability e l'Anti Spoofing;
- \square errori dipendenti dal ricevitore;
- \square errori prodotti dalle riflessioni multiple e dalla propagazione dei segnali nella ionosfera e nella troposfera;
- \square errori introdotti dal Sistema di Controllo terrestre nella determinazione delle orbite, nelle correzioni degli orologi...

Nella tabella che segue sono riportati gli errori residui dopo l'applicazione delle varie compensazioni e correzioni; alcuni di essi sono comuni a tutti gli utilizzatori e quindi possono essere ulteriormente ridotti con tecniche differenziali (Differential GPS). Altri errori (rumore del ricevitore e multipath) sono invece dipendenti dal singolo utilizzatore e non possono pertanto essere eliminati con accorgimenti di questo tipo.

Tipo di errore	Errore quadratico medio residuo [m]
Orologi satellite	3
Effemeridi	3
SA	30
Ritardi ionosferici	9
Ritardi troposferici	2
Rumore ricevitore	10
Multipath	3

L'errore quadratico medio complessivo diventa di circa 33 m, mentre nei ricevitori a doppia frequenza, utilizzando il codice P, si riduce a circa 6 m.

I suddetti errori rendono incerti i luoghi di posizione individuati e quindi la posizione finale del ricevitore come in qualsiasi altro sistema di posizionamento usato nella navigazione.

Un aspetto importante è il seguente: l'errore sulla posizione dipende, oltre che dagli errori nelle misure, da un fattore scalare legato alla geometria del sistema, ossia alla distribuzione in azimut ed in elevazione dei satelliti utilizzati nelle misure; tale fattore è denominato **Fattore d'espansione dell'errore** o **PDOP (Position Dilution Of Precision)**; in altri termini si può scrivere:

$$\boxed{\text{Errore sulla posizione} = \text{Errore sulla distanza} \cdot \text{PDOP}}$$

Un errore di 30 m nella misura delle distanze con un PDOP = 2,5 determina un errore di $2,5 \cdot 30 = 75$ m nella posizione; la posizione esatta si trova internamente ad una sfera di raggio 75 m e centro nella posizione individuata [sfera di incertezza] (in realtà questa affermazione va intesa in termini statistici); nel caso della navigazione marittima o terrestre interessa soltanto la posizione, e quindi la precisione, nel piano orizzontale: PDOP viene sostituito da **HDOP (Horizontal DOP)**.

Con riferimento alla dinamica del fattore PDOP, questa è fortemente influenzata dalla distribuzione in azimut dei satelliti utilizzati per le misure di pseudorange; può assumere un valore molto basso (valore minimo 1,6) nel caso di una distribuzione

ottimale; può assumere valori consistenti, anche superiori a 10, nel caso di satelliti tutti concentrati in un ristretto arco di orizzonte; **valori fra 2 e 4 sono considerati normali**.

I ricevitori, utilizzando i dati contenuti nell'Almanacco, sono in grado di calcolare preventivamente il PDOP o l'HDOP per le varie geometrie di satelliti in vista in modo da scegliere la combinazione con il fattore di spreading più basso.

Di seguito è riportata molto sinteticamente un'analisi dei principali tipi di errori.

3.9.2 Errori dei satelliti

Tali errori sono già stati descritti, in gran parte, nei paragrafi precedenti:

- La SA e l'AS generate dai satelliti su comando del Sistema di Controllo terrestre per degradare la precisione del sistema;
- Gli errori sugli orologi atomici di bordo;
- Gli errori prodotti dalla non perfetta conoscenza dei parametri orbitali dei satelliti.

3.9.3 Errori relativi alla propagazione nella ionosfera e troposfera

Nel sistema GPS si considera come **strato troposferico** la porzione di atmosfera che va da 0 a 70 Km e come **strato ionosferico** quello che va da 70 a 100 Km.

Nell'attraversamento di tali strati, le onde elettromagnetiche subiscono rifrazioni che producono un allungamento dei percorsi rispetto a quelli lineari previsti fra i satelliti ed il ricevitore. La loro completa eliminazione richiederebbe l'uso di due frequenze (tecniche in diversità); è sufficiente, qualora possibile, usare il codice P.

3.9.4 Errori prodotti dal Multipath troposferico

Gli errori prodotti dai percorsi multipli o multipath derivano principalmente dalla combinazione dei segnali diretti con quelli riflessi dalle superfici circostanti, in particolare dalla superficie marina. Tali errori sono quindi dipendenti dalla natura e dalla localizzazione delle superfici riflettenti, per cui non è possibile normalmente prevederne l'ampiezza, ma è possibile ridurli, o eliminarli, con un'opportuna collocazione e progettazione dell'antenna ed adottando opportune tecniche nei ricevitori. Senza adottare gli opportuni provvedimenti, l'errore prodotto può superare

anche i 50 m. Uno degli accorgimenti più evidenti è quello di schermare inferiormente le antenne riceventi.

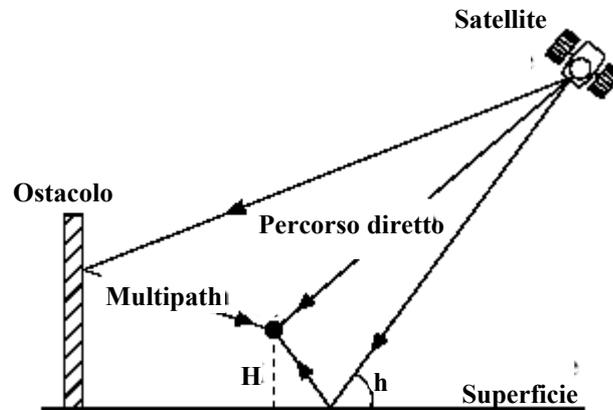


Figura 3.29 Multipath: interferenza tra raggio diretto e raggio riflesso

3.9.5 Errori del ricevitore

Ogni ricevitore genera degli errori legati al rumore interno, alla precisione con cui lavora il correlatore, ai ritardi prodotti sia dai vari dispositivi elettronici che dal software che effettua l'elaborazione dei dati. Gli effetti di tali errori possono essere consistenti nel caso di ricevitori utilizzati su veicoli molto veloci; una progettazione accurata del ricevitore, in rapporto all'uso cui è destinato, può pertanto renderli molto piccoli.

3.9.6 Errori sulla posizione

Incertezza sul luogo di posizione terrestre

La misura della distanza R , come si è visto, definisce sulla sfera terrestre una circonferenza che ha per centro il punto subsatellite S' e raggio sferico $S'A$. Un eventuale errore nella misura della distanza comporta una variazione del raggio $S'A$ di una quantità:

$$\overline{BA} = \frac{R' - R}{\cosh h}$$

dove $e_m = R' - R$ è l'errore nella misura delle distanze ed h è l'altezza apparente o topocentrica del satellite. La quantità \overline{BA} costituisce l'incertezza sulla linea di posizione, per cui si ha

$$e_1 = e_m \cdot \sec h = e_m \cdot F$$

La formula mostra che l'incertezza e_1 sulla l.d.p. (linea di posizione) aumenta all'aumentare dell'altezza del satellite secondo un fattore $F = \sec h$ (**fattore d'espansione degli errori**).

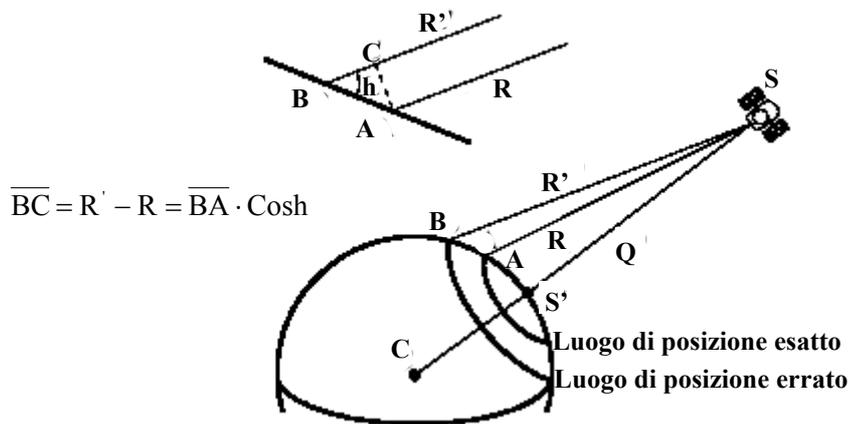


Figura 3.30 Errore sul luogo di posizione

Errore di posizione nel piano orizzontale

Nel caso di due soli luoghi di posizione, l'errore sul punto nave, analogamente alla navigazione astronomica, dipende dalle incertezze sulle due l.d.p. e dall'angolo da esse formato; dalla figura seguente si può facilmente concludere che il suddetto errore è dato dalla semidiagonale maggiore del parallelogramma individuato dalle due strisce di ampiezza $2 \cdot e_1 \cdot F_1$.

L'entità dell'errore si ottiene applicando il teorema di Carnot:

$$e_n = e_m \cdot \frac{\sqrt{F_1^2 + F_2^2 + 2 \cdot F_1 \cdot F_2 \cdot \cos \alpha}}{\sin \alpha}$$

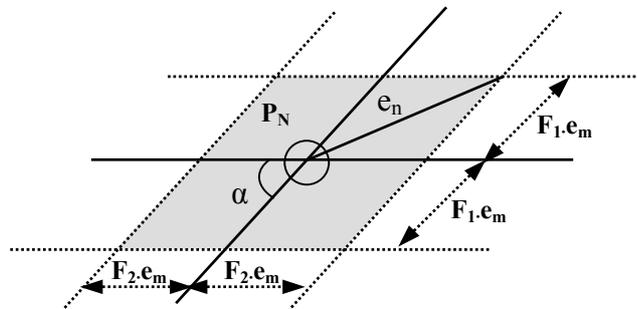


Figura 3.31 Punto nave con due linee di posizione; l'errore minimo si ha per $\alpha=90^\circ$

La migliore combinazione si ha per satelliti bassi sull'orizzonte (minime incertezze F_i) ed una differenza in azimuth di 90° . La precedente formula si può anche scrivere:

$$e_n = e_m \cdot \text{HDOP}$$

dove HDOP (Horizontal Dilution Of Position) è un fattore che dipende unicamente della distribuzione in azimuth ed in elevazione dei satelliti utilizzati per fare il punto nave.

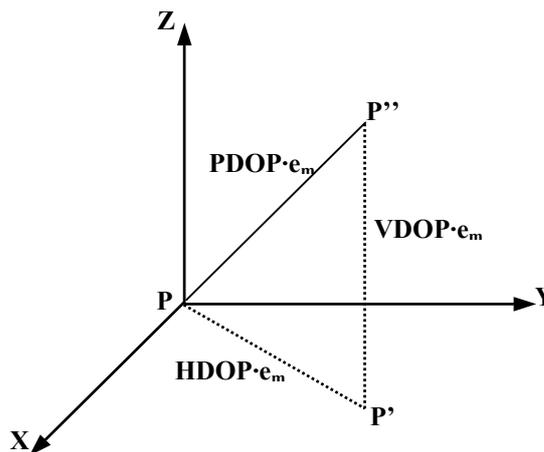


Figura 3.32 Errore di posizione nelle tre dimensioni; in genere interessa soltanto l'errore in direzione orizzontale, individuato dall'HDOP.

Nel caso di tre satelliti si arriva ad un'analogia conclusione, ma con i satelliti distribuiti uniformemente in azimut a 120° l'uno dall'altro. Nel caso di tre o più satelliti il calcolo dell'HDOP è più complesso ed i ricevitori GPS lo determinano con continuità sia per lo scopo di selezionare i satelliti che di fornire un indice sulla bontà del punto nave.

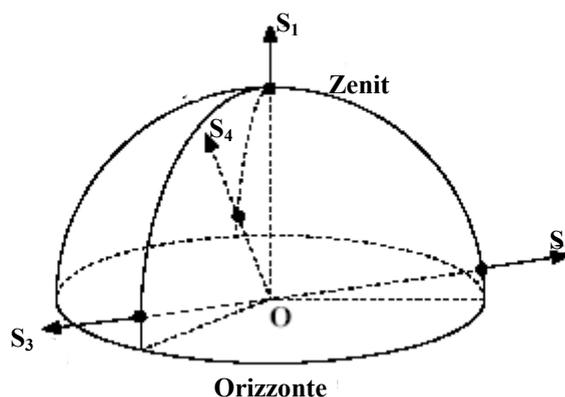


Figura 3.33 La migliore condizione si ha quando tre satelliti sono sull'orizzonte e con una differenza d'azimut pari a 120° ed il quarto satellite è allo zenit. L'altezza tuttavia deve essere sempre superiore a quella di mascheramento.

Tradizionalmente si è soliti considerare gli errori come dei valori massimi, per cui il punto esatto si trova “con certezza” dentro il parallelogramma individuato con le due strisce di posizione. Tale ipotesi semplifica la trattazione, ma è in contrasto con la teoria degli errori per la quale è sempre necessario associare ad ogni errore una probabilità.

In base alla teoria degli errori si può calcolare l'**errore quadratico medio** (o deviazione standard) indicato con σ , al quale è associata una probabilità 67%. Questo significa che l'errore reale dell'apparato è minore o uguale di σ nel 67% dei casi, supera σ nel restante 33%. Considerando un errore quadratico doppio (2σ) o triplo (3σ) la probabilità aumenta, rispettivamente al 95% e al 99,7%. Nel GPS viene indicata una precisione di 100 m al 95%: l'errore risulta minore o uguale a 100 m nel 95% dei casi. Nel restante 5%, gli errori possono superare i 100 m e diventare enormemente grandi anche se con probabilità molto basse.

Per tale ragione l'errore di posizione viene più giustamente individuato considerando un cerchio di precisione con centro in P_N e raggio pari all'errore, associandolo ad una certa

probabilità; in particolare si definisce il **CEP (Circular Error Probable)** corrispondente ad una probabilità del 50%; se uno strumento di navigazione ha un CEP = 100 m, significa che, mediamente, il 50% delle osservazioni sono all'interno di un cerchio di raggio 100 m, mentre le altre sono esterne e disseminate da 100 m all'infinito, con una probabilità prossima a zero che si verifichi quest'ultimo evento. Possono essere definiti altri cerchi di precisione di raggio 1σ , 2σ o 3σ ed ai quali corrispondono le probabilità già indicate.

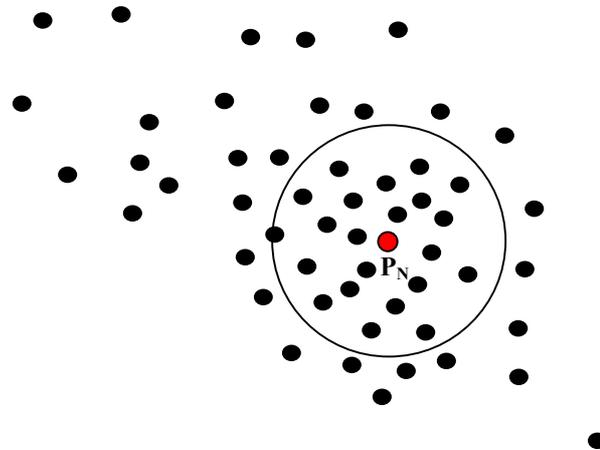


Figura 3.34 CEP[Circular Error Probable]

Errore di posizione nello spazio

Il GPS fornisce la posizione del ricevitore nelle tre dimensioni, pertanto occorre considerare l'errore di posizione nello spazio che, analogamente al caso bidimensionale, si ottiene moltiplicando l'errore e_m per il fattore **PDOP (Position DOP)** come mostrato in precedenza; in modo analogamente si può definire l'errore di posizione nella sola direzione verticale utilizzando il **VDOP (Vertical DOP)**; la relazione fra i vari fattori è intuitiva:

$$\text{PDOP} = \sqrt{\text{HDOP}^2 + \text{VDOP}^2}$$

Normalmente tutti i ricevitori GPS forniscono l'HDOP; in qualche caso forniscono anche il VDOP; tali fattori sono calcolati con continuità e si riferiscono ovviamente alla

configurazione di satelliti usata per fare il punto nave e rappresentano un indice per valutare l'accuratezza della geometria satellitare correntemente utilizzata.

I ricevitori calcolano regolarmente il PDOP di tutte le possibili combinazioni di quattro satelliti disponibili fra quelli visibili utilizzando i dati dell'Almanacco; lo scopo è quello di poter utilizzare sempre la combinazione con il PDOP minore. Il calcolo del PDOP e delle sue componenti è alquanto complesso; si dimostra che la migliore condizione si ottiene quando un satellite è allo zenit e gli altri tre sono sull'orizzonte separati in azimut da un angolo di 120° . La peggiore condizione si ha invece quando le quattro direzioni d'osservazione sono localizzate sulla superficie di un cono di semiapertura qualsiasi.

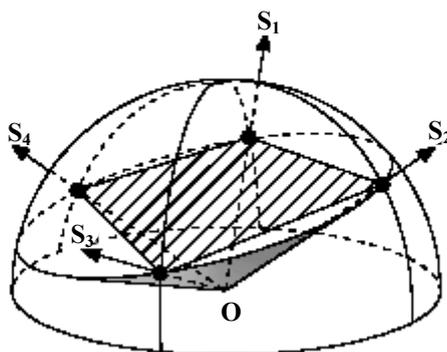


Figura 3.35 La peggiore situazione, da un punto di vista teorico, si ha quando le direzioni d'osservazione dei quattro satelliti sono sulla superficie di un cono con vertice nell'osservatore. In generale le peggiori condizioni si hanno anche quando i satelliti sono tutti concentrati dalla stessa parte dell'orizzonte.

3.10 Il sistema di riferimento utilizzato dal GPS

La posizione del terminale GPS ricevente viene espressa in termini di coordinate cartesiane geocentriche (x,y,z) o ellissoidiche (φ,λ,h) in un sistema di riferimento mondiale unificato denominato **WGS84 (World Geodetic System 1984)**.

WGS84 è modello matematico della Terra da un punto di vista geometrico, geodetico e gravitazionale ed è stato sviluppato sulla base delle misure e delle conoscenze scientifiche disponibili al 1984; può anche essere inteso come una realizzazione di un sistema di riferimento convenzionale terrestre (**CTRF**) in quanto possiede le seguenti caratteristiche (*Betti et al., 1989*):

- origine nel centro di massa della terra;
- asse Z passante per il polo convenzionale terrestre (**CTP**);
- asse X definito dall'intersezione del piano meridiano passante per Greenwich con il piano equatoriale terrestre;
- asse Y univocamente determinato in modo da ottenere una terna destrorsa di assi cartesiani ortogonali.

Stabilito il sistema di riferimento, il generico punto P risulta univocamente determinato dalla terna (X_p, Y_p, Z_p)

In un CTRF le coordinate cartesiane risultano in genere, per osservatori a terra, inadatte per la descrizione di posizioni e traiettorie; si preferisce utilizzare, in genere, una superficie di riferimento descritta da alcuni parametri (locali) e che ben approssimi il profilo terrestre; in particolare, la forma della terra può essere descritta attraverso una particolare superficie, il **geoide**, intesa come la superficie equipotenziale che assumerebbero gli oceani se sulla Terra agisse la sola forza gravitazionale e se questi potessero prolungarsi sotto le terre emerse.

Il solido di rotazione che meglio approssima il geoide è lo **sferoide** definito, in due dimensioni, dall'equazione : $R = a(1 - \alpha \sin \varphi)$; R indica la curvatura dello sferoide

mentre $\alpha = \frac{a-c}{a}$ rappresenta un parametro che dipende dall'asse maggiore a e dall'asse

minore c ; l'angolo φ varia tra 0 e π ; il solido viene generato dalla rotazione della curva generatrice intorno all'asse Z.

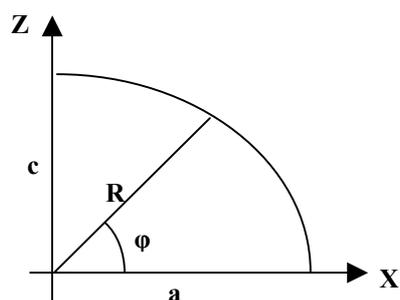
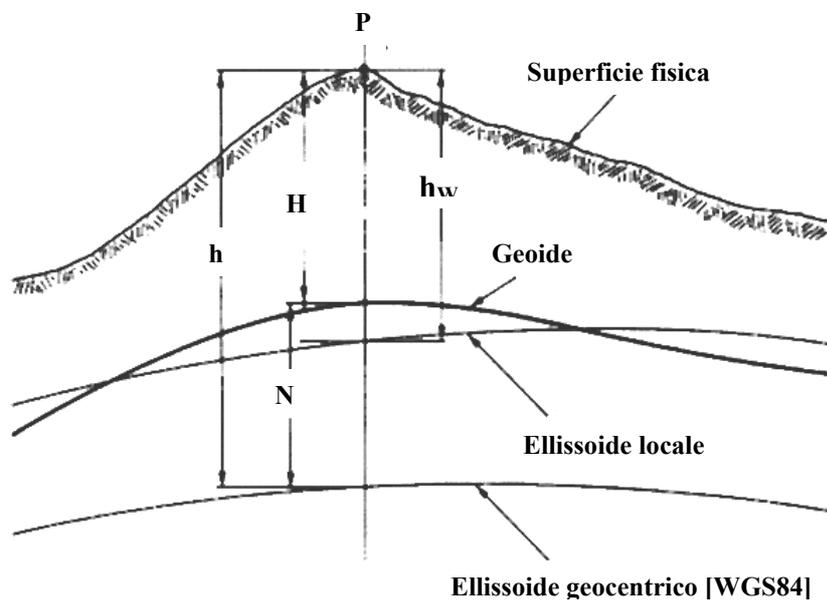


Figura 3.36 Curva generatrice dello sferoide [la parte simmetrica intorno all'asse X non è indicata]

La definizione di coordinate geografiche sullo sferoide risulta piuttosto complessa; per questo motivo il geoide viene approssimato attraverso un **ellissoide di rotazione** di forma e dimensione opportune, detto **ellissoide**; il massimo scostamento fra geoide ed ellissoide terrestre è circa 100 m.



- h indica l'altezza di un punto della superficie rispetto all'ellissoide di riferimento;
- H indica l'altezza di un punto della superficie rispetto al geoide;
- N rappresenta l'ondulazione del geoide rispetto all'ellissoide di riferimento;

Figura 3.37 Superficie terrestre, geoide, ellissoidi locale e terrestre, quote.

Si definisce ellissoide di rotazione con centro nell'origine degli assi, il luogo dei punti x,y,z dello spazio che soddisfano l'equazione:

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1$$

Il parametro a indica il semiasse maggiore o equatoriale dell'ellissoide, il parametro b il semiasse minore o polare; poiché la forma dell'ellissoide può essere modulata agendo sui valori assunti dai due parametri, possono essere utilizzati, allo scopo di migliorare l'approssimazione della superficie terrestre, ellissoidi locali in luogo di un unico ellissoide terrestre; in realtà si preferisce riferirsi ad altri due parametri di forma derivati dai due semiassi:

□ **schacciamento** $\alpha = \frac{a - b}{a}$

□ **eccentricità** $e^2 = \frac{a^2 - b^2}{a^2}$

Nella tabella successiva vengono mostrati i parametri caratterizzanti i principali ellissoidi utilizzati nel corso degli anni;

Nome (anno)	a [m]	b [m]	α	e^2
<i>Bessel (1841)</i>	6377397.155	6356078.963	1/299.152813	$6.67437223 \cdot 10^{-3}$
<i>Hayford (1909)</i>	6378388.000	6356911.946	1/297.000000	$6.72267002 \cdot 10^{-3}$
<i>WGS-84 (1984)</i>	6378137.000	6356752.314	1/298.257224	$6.69437990 \cdot 10^{-3}$

Figura 3.38 Paramteri dei principali ellissoidi

Il sistema di riferimento utilizzato dal sistema GPS e quindi anche in questo lavoro di tesi è il WGS84.

Fissato il sistema di riferimento e l'ellissoide associato, è possibile definire le **coordinate geografiche o ellissoidiche** di un punto P: **latitudine** φ , **longitudine** λ , **quota ellissoidica** h :

- latitudine φ : angolo compreso tra la normale all'ellissoide condotta per P ed il piano equatoriale XY; $\varphi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$: in alternativa può essere considerato il solo valore assoluto dell'angolo indicando con latitudini **NORD** le positive e con latitudini **SUD** quelle negative;
- longitudine λ : angolo compreso tra il piano meridiano passante per P ed un piano meridiano assunto come origine [meridiano di Greenwich]; $\lambda \in [-\pi, \pi]$: in alternativa può essere considerato il solo valore assoluto dell'angolo indicando con longitudini **OVEST** le positive e con longitudini **EST** quelle negative;
- quota ellissoidica h : distanza, misurata lungo la normale all'ellissoide condotta per P, tra l'ellissoide ed il punto P stesso; naturalmente i valori di quota possono risultare anche negativi.

Intuitivamente, la terna $(\varphi_P, \lambda_P, h_P)$ risulta in corrispondenza biunivoca con le coordinate cartesiane (X_P, Y_P, Z_P) e pertanto individua senza ambiguità il punto stesso; nel seguito verrà dimostrata formalmente questa proprietà.

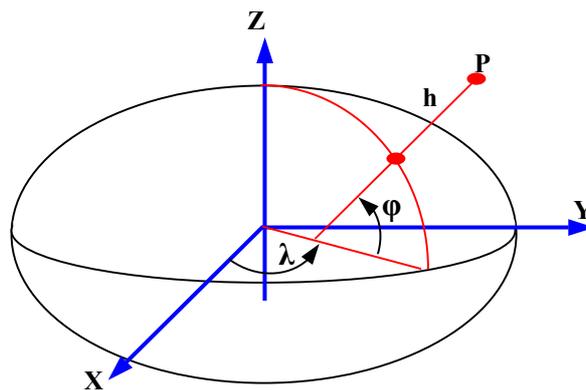


Figura 3.39 Ellissoide terrestre: coordinate geografiche o ellissoidiche.

Si definisce **Gran Normale** un particolare raggio di curvatura dell'ellissoide in P:

$$N = \frac{a}{\sqrt{(1 - e^2 \sin^2 \varphi_p)}}$$

Le relazioni tra coordinate ellissoidiche e cartesiane del punto P sono le seguenti:

$$X_p = (N + h_p) \cdot \cos \varphi_p \cdot \cos \lambda_p$$

$$Y_p = (N + h_p) \cdot \cos \varphi_p \cdot \sin \lambda_p$$

$$Z_p = [N(1 - e^2) + h_p] \cdot \sin \varphi_p$$

Si noti che il passaggio coordinate ellissoidiche \Rightarrow coordinate cartesiane è risolubile in forma chiusa; il passaggio inverso deve invece essere risolto mediante un opportuno algoritmo iterativo:

$$\lambda_p = \operatorname{arctg} \frac{Y}{X}$$

Nella prima iterazione, si pone $h=0$ e si ricavano le prime stime approssimate di:

$$\varphi_0 = \operatorname{arctg} \left(\frac{z}{\sqrt{x^2 + y^2}} \frac{1}{1 - e^2} \right)$$

$$N_0 = a \sqrt{(1 - e^2 \sin^2 \varphi_0)}$$

$$h_0 = \frac{\sqrt{x^2 + y^2}}{\cos \varphi_0} - N_0$$

Da questi primi valori, iterando, si ottengono le stime delle coordinate ellissoidiche; in genere l'algoritmo converge in meno di 5 iterazioni.

3.11 DGPS o GPS Differenziale

3.11.1 Principio di funzionamento

Come già evidenziato nella introduzione a questo capitolo, allo scopo di superare i limiti di precisione imposti per gli utenti normali, è stato introdotto un sistema basato su tecniche differenziali; il concetto di base è semplice: un ricevitore, posto in una "**stazione di riferimento**" di coordinate geografiche "perfettamente" note, confronta con continuità la sua posizione con quella calcolata tramite il sistema GPS come farebbe un normale ricevitore; dal confronto delle due posizioni si possono determinare gli errori a carattere **sistematico** che si ritiene comuni a tutti i ricevitori situati entro una determinata area attorno alla stazione. Tali correzioni vengono poi trasmesse via radio con continuità e con modalità diverse a seconda della particolare stazione DGPS; per poterle utilizzare queste informazioni, un ricevitore GPS deve essere opportunamente predisposto ed in particolare dotato di un ricevitore supplementare funzionante nella banda di trasmissione della stazione di riferimento.

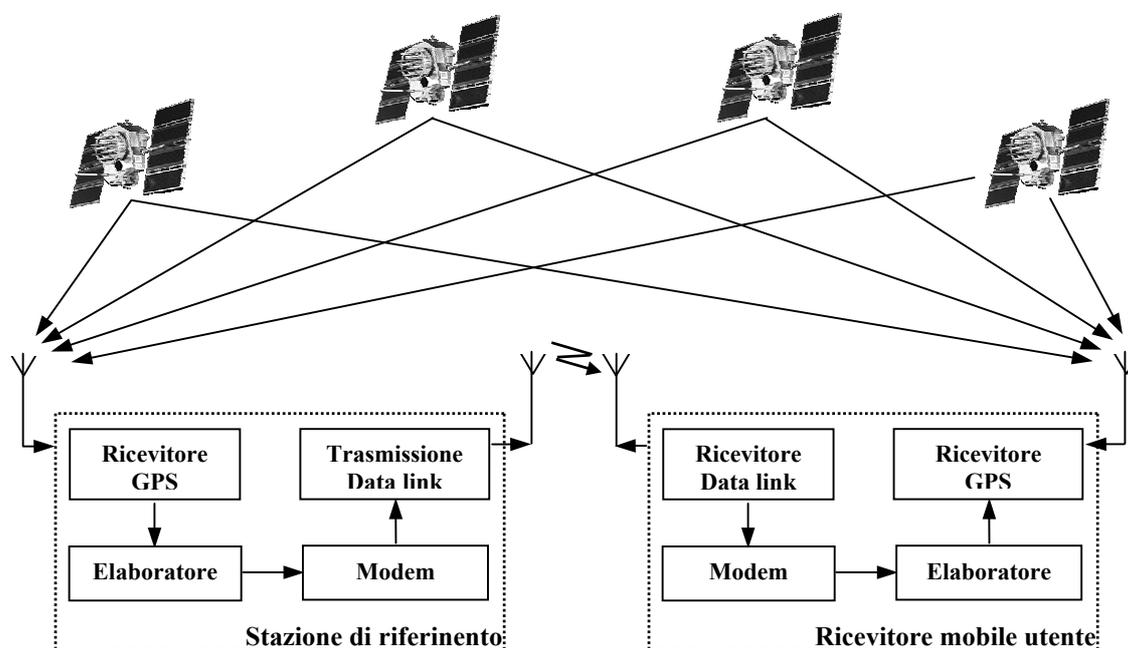


Figura 3.40 DGPS: Schema di principio

Nella tabella successiva sono confrontate le entità degli errori in modalità normale ed in modalità DGPS; si può notare come vengano completamente annullati gli errori dovuti direttamente ai satelliti (effemeridi, SA e orologi); gli errori causati dalla propagazione ionosferica e troposferica si riducono notevolmente anche se in misura minore all'aumentare della distanza dalla stazione di riferimento. La tabella contiene anche la precisione sulla posizione orizzontale, nell'ipotesi di HDOP=1,5 e di un errore sulla distanza misurata pari a 2σ (probabilità del 95%).

Errore	GPS Standard [m]	GPS Differenziale (errore in funzione della distanza dalla stazione DGPS) [m]		
		0 m	100 m	1000 m
Orologi satellite	3	0	0	0
Effemeridi	3	0	~0	1
SA	30	0	0	0
Ritardi ionosferici	9	0	~2	7
Ritardi troposferici	2	0	2	2
Rumore ricevitore	10	1	1	1
Multipath	3	0	0	0
Errore σ sulla distanza	33	1	~3	7
Errore 2σ sulla posizione orizzontale [HDOP=1,6]	100	3	10	21

Figura 3.41 Errori residui nel sistema differenziale

3.11.2 Stazioni di riferimento ed utilizzo del DGPS

Una stazione di riferimento differenziale è formata da:

- ❑ un ricevitore GPS a 10-12 canali paralleli;
- ❑ un microprocessore per calcolare le correzioni e preparare il messaggio da irradiare agli utilizzatori (**data links**)
- ❑ un modem
- ❑ un trasmettitore a radio frequenza

Negli apparati più sofisticati, i primi tre dispositivi sono tutti concentrati nel ricevitore, per cui la stazione si presenta molto compatta e facile da spostare, come ad esempio nelle campagne idrografiche per rilievi di fondali marini o di lineamenti di coste.

L'esatta posizione della stazione (ϕ , λ e h) deve essere introdotta manualmente durante la fase di inizializzazione, oppure, negli **apparati autoposizionanti**, viene calcolata direttamente dal ricevitore con particolari tecniche (necessariamente molto precise).

La stazione calcola le correzioni alle distanze relativamente a tutti i satelliti visibili e le trasmette agli utilizzatori; questi sommano le correzioni alle Pseudo Distanze misurate e calcolano le coordinate geografiche con una accuratezza maggiore.

I messaggi trasmessi dalle stazioni di riferimento, detti anche **data links**, sono stati standardizzati in modo da evitare che ogni stazione utilizzasse un protocollo proprietario; tale standard risale al 1987 ed il relativo protocollo è denominato RTCM - SC 104 (Radio Technical Commission for Maritime Services, Special Committee 104 - RTCM - SC 104). Per la precisione del sistema è necessario che le correzioni vengano aggiornate abbastanza rapidamente soprattutto nel caso in cui sia attiva la SA; in generale i messaggi vengono aggiornati ad intervalli variabili, a secondo della portata della stazione: ~ 5 sec per distanze fino a 100 m ed anche 20 sec per distanze maggiori di 700 m.

3.11.3 Organizzazione delle stazioni DGPS

A fronte delle ottime prestazioni offerte dal metodo differenziale, negli ultimi anni hanno cominciato a diffondersi le stazioni DGPS in particolare nelle aree di grande traffico marittimo o di particolare interesse per alcune applicazioni (geodetico o idrografico). Allo scopo di evitare sovrapposizioni fra le aree coperte da stazioni DGPS differenti e per aumentare l'efficacia del sistema che tende a diminuire con l'allontanarsi dalla stazione di riferimento, in molti paesi (USA e Nord Europa) si è cercato di organizzare le stazioni distribuendole uniformemente nell'area di interesse e coordinandole con una **stazione di controllo**.

Le correzioni sono calcolate dalla stazione di controllo utilizzando i dati misurati dalle varie stazioni che formano la rete. L'area coperta da tale sistema prende anche il nome di **Wide Area DGPS o WADGPS**. In molti paesi le stazioni sono collocate nelle strutture dei radiofari e trasmettono nella medesima banda con frequenze intorno ai 300

Khz; è previsto che le correzioni possano essere trasmesse utilizzando anche i satelliti del sistema di comunicazioni INMARSAT, assicurando così una copertura molto ampia (metodo SKYFIX).

3.12 Interfacciamento con ricevitori GPS

3.12.1 Introduzione

Un ricevitore GPS di solito implementa anche funzionalità di tipo applicativo: tracciamento di rotte, bussola digitale, marking di punti nave ed eventuali servizi location based; è dotato in genere di un display che realizza l'interfaccia utente oltre a visualizzare velocità e posizione.

In molte applicazioni è però necessario poter disporre dei dati raccolti dal ricevitore allo scopo di elaborarli tramite software opportuni su sistemi di calcolo decisamente più potenti del semplice terminale; da qui l'esigenza di interfacciare un ricevitore GPS con un calcolatore.

Diversi sono i protocolli utilizzati su questa interfaccia, quasi sempre di tipo seriale; l'aspetto negativo è che ciascun costruttore sfrutta tipicamente il protocollo (proprietario) studiato per l'upgrade del firmware dei suoi dispositivi per trasportare anche i dati GPS; se da una parte questa soluzione rende più efficiente il trasferimento, dall'altra crea evidenti problemi di compatibilità tra software e apparecchiature; fortunatamente i ricevitori consentono, in generale, di selezionare protocolli alternativi a quelli proprietari, meno efficienti ma di uso generale; i più diffusi sono:

- NMEA OUT
- TEXT OUT
- RTCM IN
- RTCM / NMEA
- RTCM / TEXT

In particolare, il protocollo NMEA OUT è, di fatto, il più utilizzato per questo tipo di applicazioni; poiché dovrà essere utilizzato dal tool di acquisizione dati presentato nel capitolo 5, nel paragrafo successivo ne verranno analizzati gli aspetti significativi.

3.12.2 Lo standard NMEA 0183

NMEA 0183 è uno standard di uso generale per l'interfacciamento di apparecchiature digitali. Il sistema ha origini e finalità prevalentemente nautiche: viene utilizzato, per esempio, nei sistemi di autopilota per imbarcazioni.

Per supportare le molteplici applicazioni marine e l'interfacciamento di un'ampia gamma di apparecchiature nautiche, i formati di messaggio utilizzati del sistema NMEA sono numerosi e solo una piccola parte ha rilevanza nell'ambito GPS, dove il protocollo viene prevalentemente impiegato per la trasmissione dati da un ricevitore GPS verso un calcolatore.

Relativamente alle applicazioni GPS del protocollo, un primo aspetto importante da sottolineare è l'unidirezionalità: non esiste uno schema richiesta – risposta in quanto i messaggi vengono trasferiti ciclicamente attraverso l'interfaccia dal ricevitore GPS verso il calcolatore; in altri termini NMEA non prevede alcun flusso dati o di controllo nella direzione opposta; come si avrà modo di illustrare nel capitolo 5, questa osservazione condiziona fortemente l'architettura software del tool di acquisizione dati.

Si è già accennato al fatto che esistono altri protocolli utilizzati da un ricevitore GPS: in effetti ogni costruttore ne ha prodotto uno: Trimble, SiRF, Garmin, Rockwell; spesso questi protocolli proprietari usano formati binari compressi, studiati per realizzare un trasferimento dati efficiente e veloce; il vantaggio principale dei protocolli proprietari non è tanto la minore dimensione in byte del messaggio (aspetto non più rilevante date le attuali capacità di trasmissione e memorizzazione) ma soprattutto la possibilità di trasportare dati relativi ai segnali radio delle portanti dei satelliti, la loro fase ed altre informazioni potenzialmente utilizzabili per produrre una accuratezza superiore; queste informazioni non sono disponibili nel protocollo NMEA.

NMEA è un protocollo orientato al carattere: i suoi messaggi risultano intelligibili all'occhio umano e facilmente decodificabili dal software di un elaboratore.

L'ente che gestisce lo standard (e che ne detiene il copyright) si chiama National Marine Electronics Association; si tratta di un'associazione composta da produttori, distributori, rivenditori, scuole ed altri parti che si occupano dell'elettronica in campo nautico. NMEA definisce una serie di standard, fra cui NMEA 0183.

Lo standard definisce i segnali elettrici, il protocollo, la temporizzazione della trasmissione ed i specifici formati di dati - detti **sentences** (frasi) - operanti su un **bus dati seriale** (tipicamente RS232, ma anche sue semplificazioni fino all'essenziale RX-TX-GND) a **4800 bps**. Lo standard prevede che ogni bus dati può avere un unico elemento trasmittente (**talker**) e molti riceventi (**listeners**). Alcuni ricevitori sono in grado di emettere dati a velocità di trasmissione maggiori, fermo restando che gli aggiornamenti vengono prodotti **una volta al secondo**. Sul mercato esistono ricevitori che, oltre all'interfaccia RS232, offrono l'interfaccia USB, PCMCIA/CompactFlash, infrarossi e Bluetooth.

Il ricevitore utilizzato nel esperienze di misura per questo lavoro di tesi è il **GARMIN eTREX 12 Channel GPS**, dotato unicamente di interfaccia RS232.

Le applicazioni per personal computer e palmari necessitano, per il corretto funzionamento, soltanto di pochissime frasi NMEA.

La versione 2.3 dello standard ha introdotto alcune significative varianti al protocollo:

- La frase **RMC**;
- La frase **VTC**;
- La frase **GLL**;
- Un campo prima della checksum che può valere:
 - A = Autonomous;
 - D = Differential;
 - E = Estimated;
 - N = Non Valid Data.

Il tool di acquisizione dati presentato nel capitolo 5 utilizzerà NMEA 0183 2.3.

3.12.3 Struttura delle frasi NMEA

Tutte le frasi NMEA sono costituite da una sequenza di caratteri ASCII con la seguente struttura:

- ❑ **prefisso**
- ❑ **una serie di campi**
- ❑ **una checksum finale.**

L'inizio della frase è delimitato da un carattere "\$", la fine da una sequenza CR LF ed i singoli campi sono separati da virgole.

\$PREFISSO, campo, campo, ..., campo, *checksumCRLF

Una frase NMEA può contenere fino a 82 caratteri compresi "\$" e CR LF; qualora i dati per un campo non fossero disponibili, il campo viene omissso ma le virgole che lo delimitano sono comunque aggiunte, senza spazi tra di loro.

3.12.4 I prefissi NMEA

I dispositivi che utilizzano NMEA sono molteplici: autopiloti, apparati di comunicazione, sensori di direzione (bussole e giroscopi), sensori di velocità, trasduttori...

Tutte le frasi NMEA iniziano con un prefisso costituito da un carattere "\$", seguito da due caratteri che identificano l'entità che le genera e le trasmette sul bus dati (il **talker**); nell'ambito GPS, tutte le frasi iniziano con "\$GP", ad indicare, per l'appunto, che la frase è generata da un dispositivo GPS. I successivi tre caratteri del prefisso della frase indicano il tipo di frase; a titolo di esempio, il prefisso \$GPGLL è inviato da un dispositivo di tipo GPS (GP) ed in questo ambito il suo tipo è Geographic position, Longitude and Latitude (GLL).

3.12.5 Il formato dei campi

La frase, dopo il prefisso, prosegue con una serie di campi, separati da virgole. Si noti che tra le varie versioni (e spesso tra diverse implementazioni e costruttori) le lunghezze dei campi possono variare, ed essere o meno presenti. È opportuno che il software di

decodifica consideri i campi di lunghezza variabile utilizzando, per riconoscerli, le virgole; alcune implementazioni includono gli zeri davanti ai numeri, altre li omettono e lo stesso numero di cifre decimali varia tra diverse implementazioni.

I formati utilizzati dalle frasi NMEA GPS sono:

- **A**: campo alfanumerico di lunghezza fissa (un carattere);
- **A-A**: campo alfanumerico di lunghezza variabile;
- **lll.ll**: campo latitudine (es: 4531.47 = 45°31.47')
- **yyyy.yy**: campo longitudine (es: 00917.21 = 9°17.21')
- **x**: campo numerico con n cifre intere;
- **x.x**: campo numerico con n cifre intere e n cifre decimali (es: 123.45)
- **n**: campo numerico, singola cifra;
- **nn**: campo numerico, due cifre;
- **nnnn**: campo numerico, quattro cifre;
- gli angoli sono misurati in gradi sessagesimali;
- riferimenti a UTC si intendono relativi all'ora universale, in pratica GMT, detta anche ora Zulu. In effetti esistono piccoli scostamenti (costanti e accuratamente monitorati) tra l'ora del sistema GPS e l'ora GMT, normalmente usata per gli orologi; l'offset attuale è intorno ai 13 secondi.

3.12.6 La checksum

Al termine di ciascuna frase NMEA viene posta, dopo un asterisco, una checksum per poter individuare eventuali errori sulla linea di trasmissione tra il talker ed il listener.

Per le attuali tecnologie, una trasmissione seriale a 4.800 bps su una interfaccia RS232 è, a dir poco, affidabile, ma qualora la trasmissione dovesse avvenire mediante un sistema più complesso (ad esempio a radiofrequenza - infrarosso, Bluetooth, GSM/GPRS-) la checksum può essere facilmente utilizzata per verificare l'integrità della trasmissione.

La checksum è l'OR esclusivo a 8-bit (senza i bit di start o stop) di tutti i caratteri della frase NMEA, compresi le virgole, ma escludendo i delimitatori "\$" e "*". Si considera poi l'equivalente esadecimale dei 4 bit più significativi e meno significativi del risultato ottenendo 2 caratteri ASCII (0-9, A-F); questi vengono aggiunti in coda alla frase

NMEA trasmettendo per primo il carattere più significativo. La frase viene terminata da una sequenza CR LF.

3.12.7 NMEA sentences

Il gruppo \$GP è costituito da circa 60 frasi; le più utilizzate sono:

\$GPRMC	\$GPGSV	\$GPBWC
\$GPRMB	\$GPRTE	\$GPAAM
\$GPGGA	\$GPBOD	\$GPXTE
\$GPGSA	\$GPWPL	\$GPMSS
\$GPGLL	\$GPVTG	

A queste si aggiungono alcune frasi speciali come \$HCHDG (per la direzione ricavata da sensore magnetico) e frasi proprietarie, specifiche di ciascun produttore di GPS.

Per un trattamento completo delle frasi NMEA è necessario far riferimento alla documentazione dello standard; tuttavia è opportuno riportare la decodifica delle frasi più comuni e comunque di quelle utilizzate nel tool di acquisizione presentato nel capitolo 5.

3.12.8 \$GPRMC: Recommended Minimum Specific GPS/TRANSIT Data

Si tratta di una delle frasi più complete; comprende informazioni essenziali relativi a data/ora, posizione, velocità e validità dei dati.

Campo	Formato	Esempio	Descrizione
1	hhmmss.ss	125642.57	Ora UTC-GPS
2	A	A	Stato: A=Active, Attivo; V=Void, Nullo
3	llll.ll	4512.72	Latitudine della posizione attuale (numero variabile di decimali)
4	A	N	Emisfero della posizione attuale: N=Nord; S=Sud
5	yyyy.yy	00934.88	Longitudine della posizione attuale (numero variabile di decimali)
6	A	E	Verso della posizione attuale: E=Est; W=Ovest
7	x.x	023.45	Velocità al suolo, in nodi.
8	x.x	358.20	Track Made Good - Direzione di movimento, in gradi reali.
9	ddmmyy	291105	Data
10	x.x	1.2	Variazione / declinazione magnetica , in gradi
11	A	E	Verso della variazione / declinazione magnetica
12	A	D	Tipo di rilevazione: A=Autonomous, D=Differential, E=Estimated, N=Non valid data (dalla versione 2.3)

3.12.9 \$GPRMB: Recommended Minimum Navigation Info (Waypoint based navigation active)

Frase dedicata alla navigazione tra waypoint, con indicazioni relative all'avvicinamento alla destinazione.

Campo	Formato	Esempio	Descrizione
1	A	A	Stato: A=Active, Attivo; V=Void, Nullo
2	x.x	1.3	Cross Track Error (Scarrocciamento) in miglia nautiche
3	A	L	Verso dello scarrocciamento: L=Left; R=Right
4	A-A	BOA52	Identificatore del waypoint di destinazione
5	A-A	PORTO	Identificatore del waypoint di origine
6	llll.ll	4613.83	Latitudine del waypoint di destinazione
7	A	N	Emisfero del waypoint di destinazione
8	yyyy.yy	00945.99	Longitudine del waypoint di destinazione
9	A	E	Verso del waypoint di destinazione
10	x.x	123.45	Distanza (Range) dalla destinazione in miglia nautiche
11	x.x	356.2	Direzione verso la destinazione in gradi reali
12	x.x	12.2	Velocità di avvicinamento in nodi
13	A	A	Allarme arrivo: A=Arrivato; V=Non arrivato

3.12.10 \$GPGGA: Global Positioning System fix data

Frase che comprende dati relativi al fix tridimensionale (3D) ed alla correzione differenziale.

Campo	Formato	Esempio	Descrizione
1	hhmmss.ss	125642.57	Ora UTC - GPS
2	llll.ll	4512.72	Latitudine della posizione attuale (numero di decimali variabile)
3	A	N	Emisfero della posizione attuale: N=Nord; S=Sud
4	yyyy.yy	00934.88	Longitudine della posizione attuale (numero di decimali variabile)
5	A	E	Verso della posizione attuale: E=Est; W=Ovest
6	n	1	Qualità del rilevamento GPS: 0=non valido; 1=GPS; 2=DGPS
7	x	8	Numero di satelliti in vista (teorica, non necessariamente in uso o ricevuti)
8	x.x	1.9	HDOP - Horizontal Dilution Of Precision
9	x.x	540.4	Altitudine dell'antenna GPS relativa al livello medio del mare (geoide)
10	A	M	Indicatore dell'unità di misura dell'altitudine M=metri
11	x.x	45.8	Altezza (separazione) del geoide (livello medio del mare) rispetto all'elissoide WGS84
12	A	M	Indicatore dell'unità di misura della separazione del geoide M=metri
13	x.x	8.0	Tempo in secondi dall'ultimo aggiornamento DGPS
14	nnnn	1023	Identificatore della stazione DGPS (0000-1023)

3.12.11 \$GPGSA: GPS DOP and Satellites Active

Frase che comprende dati relativi alla qualità del fix e l'indicazione dei satelliti ricevuti (o di cui i vari canali stanno tentando la ricezione).

Esistono significative differenze di implementazione nell'ordine e organizzazione dei campi indicanti i satelliti attivi; alcuni ricevitori li raggruppano all'inizio della sequenza (lasciando campi vuoti alla fine), altri li ordinano per numero (Garmin), altri ancora li lasciano in ordine sparso con campi vuoti intermedi (Evermore).

Campo	Formato	Esempio	Descrizione
1	A	M	Tipo di selezione del tipo di rilevamento (2D o 3D): M=Manuale; A=Automatico
2	n	3	Tipo di rilevamento: 2=2D (bidimensionale); 3=3D (tridimensionale)
3	nn	03	Identificatore (SV = PRN) di satellite ricevuto
4	nn	12	Identificatore (SV = PRN) di satellite ricevuto
...			
...			
14	nn	01	Identificatore (SV = PRN) di satellite ricevuto
15	x.x	1.9	PDOP - Position Dilution Of Precision
16	x.x	1.2	HDOP - Horizontal Dilution Of Precision
17	x.x	2.2	VDOP - Vertical Dilution Of Precision

3.12.12 \$GPGLL: Geographic position - Latitude/Longitude

Frase minima: solo dati posizionali, ora e validità del fix.

Campo	Formato	Esempio	Descrizione
1	llll.ll	4512.72	Latitudine della posizione attuale (numero di decimali variabile)
2	A	N	Emisfero della posizione attuale: N=Nord; S=Sud
3	yyyyy.yy	00934.88	Longitudine della posizione attuale (numero di decimali variabile)
4	A	E	Verso della posizione attuale: E=Est; W=Ovest
5	hhmmss.ss	125642.57	Ora UTC GPS
6	A	A	Stato: A=Active, Attivo; V=Void, Nullo (freccia in su/giù)

3.12.13 \$GPGSV: Satellites in View

Indica dati relativi a tutti i satelliti che il ricevitore, in base ai propri dati di almanacco ed eventualmente alla maschera di elevazione, ritiene di poter ricevere.

Di ciascun satellite vengono indicati identificatore, azimuth, elevazione e livello di segnale ricevuto. La frase completa si struttura in un gruppo (ciclo) di più frasi di tipo GSV per poter contenere tutti i campi necessari; una frase GSV può contenere le informazioni relative al più a 4 satelliti.

Un ricevitore a 12 canali potrà indicare un massimo di 12 satelliti e quindi generare fino a 3 frasi GSV.

Campo	Formato	Esempio	Descrizione
1	n	3	Numero di frasi del gruppo
2	n	1	Numero progressivo della frase nel gruppo (1 di 3 nell'esempio)
3	nn	09	Numero totale di satelliti teoricamente ricevibili
4	nn	03	Identificatore (SV = PRN) di satellite ricevuto del primo dell'elenco
5	nn	85	Elevazione in gradi sull'orizzonte del primo satellite dell'elenco
6	nnn	224	Azimuth in gradi del primo satellite dell'elenco
7	nn	37	Livello del segnale ricevuto dal primo satellite dell'elenco
8	nn	09	Identificatore (SV = PRN) di satellite ricevuto del secondo dell'elenco
9	nn	125	Elevazione in gradi sull'orizzonte del secondo satellite dell'elenco
10	nnn	34	Azimuth in gradi del secondo satellite dell'elenco
11	nn	24	Livello del segnale ricevuto dal secondo satellite dell'elenco
...			
16	nn	31	Identificatore (SV = PRN) di satellite ricevuto del quarto dell'elenco
17	nn	17	Elevazione in gradi sull'orizzonte del quarto satellite dell'elenco
18	nnn	120	Azimuth in gradi del quarto satellite dell'elenco
19	nn	42	Livello del segnale ricevuto dal quarto satellite dell'elenco

3.12.14 Il messaggio NMEA

Come già anticipato, il protocollo NMEA utilizzato per l'interfacciamento digitale con ricevitori GPS, è un protocollo **only – out**; per questo motivo viene spesso indicato con l'espressione NMEA OUT. Una volta attivata l'interfaccia, il ricevitore inizia a trasmettere messaggi NMEA ad intervalli di tempo approssimativamente costanti (la frequenza è di circa 1 messaggio al secondo) finché l'interfaccia non viene disattivata; è importante osservare che il messaggio viene emesso comunque allo scadere del quanto di tempo, indipendentemente dal fatto che trasporti o meno informazioni valide; si è già visto che la validità dei dati forniti è infatti indicata dal protocollo stesso in alcuni sue frasi (es. \$GPRMC).

Un messaggio NMEA è sempre costituito da un insieme di frasi; la struttura del messaggio, ossia il tipo, il numero e la sequenzialità delle stesse, dipende dalla particolare implementazione; per qualunque ricevitore, però, un messaggio NMEA è sempre costituito dall'insieme di frasi compreso tra due successive ripetizioni di una stessa frase la cui lunghezza massima è sempre inferiore a 82 caratteri; tutte le frasi analizzate in precedenza, ad eccezione di \$GPGSV, soddisfano questo requisito e possono essere assunte come header del messaggio NMEA.

```

$PGRME,28.2,M,50.0,M,57.4,M*15
$PGRMZ,-2,f,2*35
$GPRTE,1,1,c,*37
$GPRMC,122400,A,4153.0879,N,01230.6083,E,0.0,289.9,190803,1.7,E,A*19
$GPRMB,A,,,,,,,,,A,A*0B
$GPGGA,122400,4153.0879,N,01230.6083,E,1,03,4.0,-0.7,M,45.1,M,,*62
$GPGSA,A,2,08,10,,,,,27,,,,,4.2,4.0,1.0*3C
$GPGSV,3,1,09,08,88,261,39,10,53,305,32,13,24,089,00,17,17,313,00*71
$GPGSV,3,2,09,24,23,233,00,26,05,285,00,27,59,050,43,28,29,160,00*77
$GPGSV,3,3,09,29,27,293,00*46
$GPGLL,4153.0879,N,01230.6083,E,122400,A,A*4A
$GPBOD,,T,,M,,*47

$PGRME,15.0,M,22.5,M,15.0,M*1B
$PGRMZ,628,f,3*17
$GPRTE,1,1,c,*37
$GPRMC,143628,V,4149.8208,N,01242.1278,E,10.8,0.0,041003,1.7,E,S*2E
$GPRMB,V,,,,,,,,,A,S*0E
$GPGGA,143628,4149.8208,N,01242.1278,E,8,10,2.0,191.4,M,44.8,M,,*44
$GPGSA,A,3,04,05,07,09,14,18,23,26,29,30,,,,,2.0,3.0*13
$GPGSV,3,1,10,04,11,094,38,05,55,270,48,07,25,047,43,09,80,056,51*76
$GPGSV,3,2,10,14,13,322,39,18,08,252,37,23,11,294,37,26,26,174,43*78
$GPGSV,3,3,10,29,09,166,37,30,24,256,42*7D
$GPGLL,4149.8208,N,01242.1278,E,143628,V,S*4B
$GPBOD,,T,,M,,*47

$PGRME,4.7,M,5.4,M,7.2,M*29
$PGRMZ,117,f,3*1C
$GPRTE,1,1,c,*37
$GPRMC,150632,A,4154.1889,N,01231.1482,E,15.7,315.3,190803,1.7,E,A*28
$GPRMB,A,,,,,,,,,A,A*0B
$GPGGA,150632,4154.1889,N,01231.1482,E,1,08,1.7,36.1,M,45.1,M,,*79
$GPGSA,A,3,07,08,09,10,18,26,28,29,,,,,2.7,1.7,2.0*38
$GPGSV,2,1,08,07,29,122,00,08,22,073,39,09,26,276,44,10,25,201,41*77
$GPGSV,2,2,08,18,16,320,43,26,72,312,47,28,52,055,48,29,85,146,36*7A
$GPGLL,4154.1889,N,01231.1482,E,150632,A,A*46
$GPBOD,,T,,M,,*47

```

Figura 3.42 Tre messaggi NMEA: il primo ed il terzo trasportano dati validi; il terzo ha una frase \$GPGSV in meno in quanto l'almanacco suggerisce al ricevitore un numero di satelliti che dovrebbero risultare in vista pari a 8 [vedi frasi \$GPGSV]; nel primo e nel secondo messaggio tale numero risulta, rispettivamente 9 e 10.