7.3 GPS Observables and Data Processing

7.3.1 Observables

7.3.1.1 Classical View

Four basic observables can be identified:
- pseudoranges from code measurements,
- pseudorange differences from integrated Doppler counts,
- carrier phases or carrier phase differences, and
- differences in signal travel time from interferometric measurements.

A *pseudorange from code measurements* equals the time shift that is necessary to correlate the incoming code sequence with a code sequence generated in the GPS receiver, multiplied by the velocity of light [7.1.4], [7.3.1.2]. The fundamental observation equation for a single pseudorange is

$$\begin{aligned} PR_i &= |\boldsymbol{X}_i - \boldsymbol{X}_B| + cdt_u = c\tau_i \\ &= ((X_i - X_B)^2 + (Y_i - Y_B)^2 + (Z_i - Z_B)^2)^{\frac{1}{2}} + cdt_u, \end{aligned} \tag{7.35}$$

with the notations from Fig. 7.29:

R_i geometrical distance (slant range) between satellite antenna S_i and receiver antenna B,
$\boldsymbol{X}_i$ satellite position vector in the geocentric CTS [2.1.2] with the components X_i, Y_i, Z_i,
$\boldsymbol{X}_B$ position vector of the receiver antenna B in the CTS with the components X_B, Y_B, Z_B,
τ_i observed signal propagation time between satellite antenna S_i and observer antenna B,
dt_u clock synchronization error between GPS system time and receiver clock, and
c signal propagation velocity.

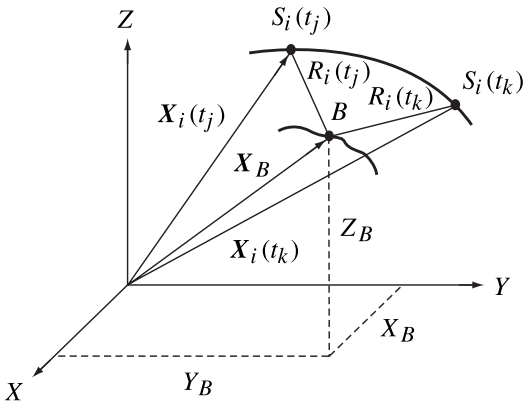

Figure 7.29. Geometric relations in satellite positioning

The coordinates of the observer antenna $\boldsymbol{X}_B$ can be derived from simultaneous range measurements to four satellites (Fig. 7.30(a)).

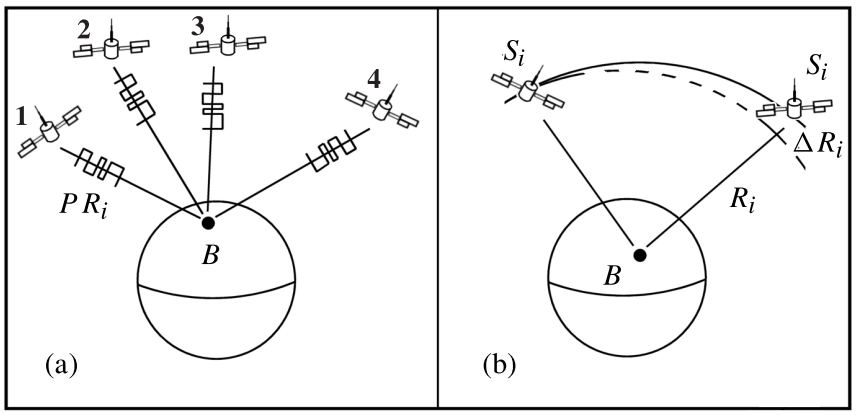

Figure 7.30. Observation concept with pseudoranges (a) and pseudorange differences (b)

Pseudorange differences can be derived from observations of the Doppler shift of the received carrier frequency, in a similar way to that done with the TRANSIT NNSS system (Fig. 7.30(b), cf. [4.2.3], [6.2]). The frequency shift of the received carrier, f_r, is measured with respect to a reference frequency, f_g, within the receiver, and yields the integrated Doppler count:

$$N_{jk} = \int_{t_j}^{t_k} (f_g - f_r)\, dt. \tag{7.36}$$

N_{jk} is a measure of the range difference between the receiver antenna, B, and two consecutive orbital positions of the same satellite, S_i, at two different epochs, t_j, and t_k, (Fig. 7.29). The related observation equation is, see (6.9):

$$\Delta R_i = |\boldsymbol{X}_i(t_k) - \boldsymbol{X}_B| - |\boldsymbol{X}_i(t_j) - \boldsymbol{X}_B| = \frac{c}{f_0}(N_{jk} - (f_g - f_s)(t_k - t_j)), \tag{7.37}$$

with

f_g a reference frequency generated in the receiver,
f_s the frequency of the signal emitted by the satellite antenna, and
f_r the frequency of the signal received at the observer antenna.

It is also possible to measure the Doppler shift of the codes. The resolution is, however, very poor because of the low code frequency when compared with the carrier frequency. The integrated Doppler measurement (7.37) should not be confused with the instantaneous Doppler measurement, used in velocity determination with navigational receivers [7.6.2.7].

The *carrier phase* is derived from a phase comparison between the received Doppler-shifted carrier signal, f_{CR}, and the (nominally constant) receiver-generated reference frequency, f_0. The proper observable is then the measured phase difference

$$\Phi_B = \Phi_{\mathrm{CR}} - \Phi_0. \tag{7.38}$$

With

λ the carrier wavelength,
N_{B_i} the integer number of complete carrier cycles within the range R_i, and
dt_u the clock synchronization error,

the fundamental observation equation of carrier phase measurements (cf. Fig. 7.31(a)) follows thus:

$$\Phi_{B_i} = \frac{2\pi}{\lambda}(|\boldsymbol{X}_i - \boldsymbol{X}_B| - N_{B_i}\lambda + cdt_u). \tag{7.39}$$

The main difficulty related to this method is the determination of the cycle ambiguity, N_{B_i}, because the observable only determines the phase within one wavelength. The *ambiguity term*, N_{B_i}, has to be determined with appropriate methods [7.3.2.3], or the range must be known a priori with an accuracy corresponding to the half cycle length ($\sim$10 cm).

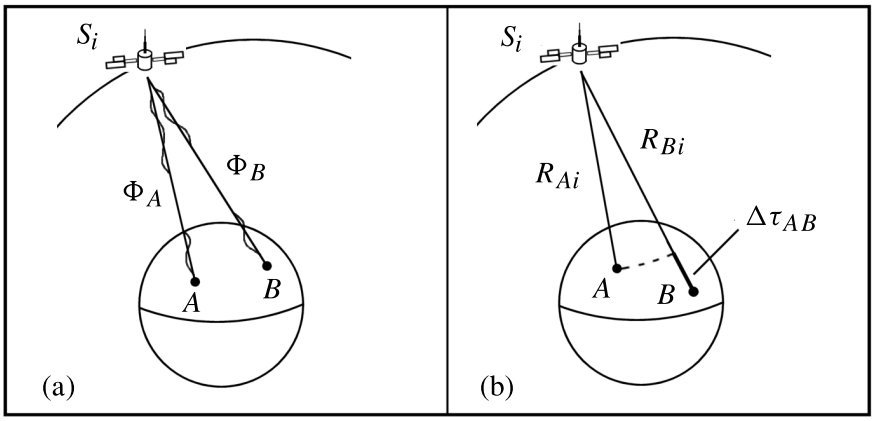

Figure 7.31. Observation concept with phase differences (a) and interferometric time differences (b)

Frequently, the phase difference of the same satellite signal, observed at two stations A and B, is considered as the basic observable. The observation equation of the *single phase difference* (or in short, *single difference*) is:

$$\begin{aligned}\Delta\Phi_{AB_i} = \Phi_{B_i} - \Phi_{A_i} = {} & \frac{2\pi}{\lambda}(|\boldsymbol{X}_i - \boldsymbol{X}_B| - |\boldsymbol{X}_i - \boldsymbol{X}_A|) \\ & - (N_{B_i} - N_{A_i})\lambda + c(dt_{u_B} - dt_{u_A}).\end{aligned} \tag{7.40}$$

The original observable in (7.39) is, for clarity, sometimes referred to as *undifferenced phase observable* or *zero phase measurement*.

In the case of pure *interferometric* observations (Fig. 7.31(b)) the GPS signals are used without knowledge or use of the signal structure (cf. [7.2.3]). The signals are recorded, together with precise time marks, at least at two stations, A and B, and are then correlated. The fundamental observable is the difference, $\Delta\tau_{A,B_i}$, between the signal arrival times at both stations with respect to a particular satellite S_i. The observable can be scaled into a range difference, $R_{B_i} - R_{A_i}$. The observation equation is:

$$\Delta\tau_{A,B_i} = \frac{(R_{B_i} - R_{A_i})}{c} = \frac{(|\boldsymbol{X}_i - \boldsymbol{X}_B| - |\boldsymbol{X}_i - \boldsymbol{X}_A|)}{c} + (dt_{u_B} - dt_{u_A}). \quad (7.41)$$

The technique is very similar to *Very Long Baseline Interferometry* (VLBI) which uses radio signals from Quasars [11.1].

Two of the four observables mentioned above are never or only rarely used in applied geodesy. The integrated Doppler count requires a rather long observation time (several hours) to allow the satellite configuration to change sufficiently, and it requires very stable oscillators in the user segment. The method is, however, implicitly used for the determination of ambiguities (cf. [7.3.2.3]). The genuine interferometric technique requires highly sophisticated instrumentation and high data processing expenditure.

In practice, therefore, only two fundamental observables are used that can be regarded as measurements of pseudoranges:

- code phases (pseudoranges from code observations), and
- carrier phases (pseudoranges from carrier observations).

7.3.1.2 Code and Carrier Phases

The main characteristics and differences between the two observables are summarized in Table 7.6. Signal propagation and the observation procedure are discussed below in more detail. Fig. 7.32 illustrates the signal propagation of the code and carrier phases, in which

T	is the satellite time (time system of the individual space vehicle),
t	is the receiver time (subscript RCV),
subscript t	refers to the transmitted signal,
subscript r	refers to the received signal,
f_{CD}	is the code frequency, and
f_{CR}	is the carrier frequency.

The signal states (phases) at the epoch of transmission are:

$$\begin{aligned} \text{Code} \quad \Phi_{\mathrm{CD}}(T_t) &= T_{t_{\mathrm{SV}}} f_{\mathrm{CD}}, \\ \text{Carrier} \quad \Phi_{\mathrm{CR}}(T_t) &= T_{t_{\mathrm{SV}}} f_{\mathrm{CR}}, \end{aligned} \quad (7.42)$$

according to the fundamental relation (2.80) between phase, frequency and time.

Table 7.6. Main characteristics of code- and carrier phases

	Code	Carrier
wavelength	P-code 29.3 m	L1 19.0 cm
	C/A-code 293 m	L2 24.4 cm
observation noise	P-code 0.3–1 m	
	C/A-code 3–10 m	1–3 mm
new development	P-code cm–dm	
	C/A-code dm–m	< 0.2 mm
propagation effects	ionospheric delay $+\Delta T_{\mathrm{ION}}$	ionospheric advance $-\Delta T_{\mathrm{ION}}$
ambiguity	non-ambiguous	ambiguous

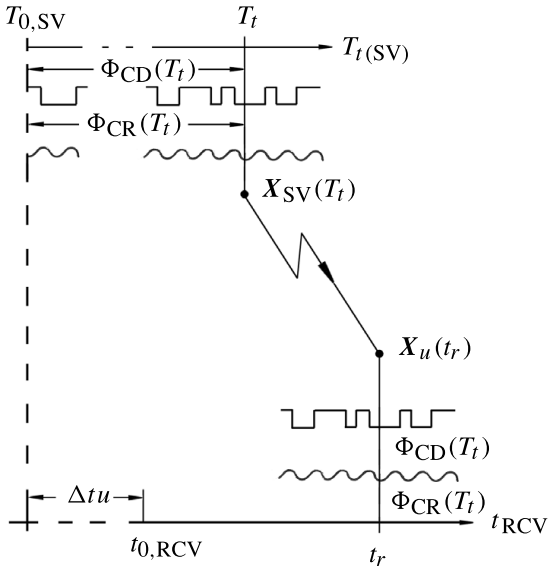

Figure 7.32. Signal propagation of code and carrier phases

The carrier phase, $\Phi_{\mathrm{CD}}(T_t)$, leaves the satellite antenna at the epoch, T_t, measured in the satellite time frame. The signal state propagates with approximately the velocity of light, and reaches the receiver antenna at the epoch, t_r, measured in the receiver time frame. The same is valid for the code signals. Note that the phase states of the received signals are identical to the phase states of the transmitted signals. In other words, measuring the phase state at the receiver means measuring the signal's emission epoch at the satellite.

The process of *code phase observation* is illustrated in Fig. 7.33. The code sequence, generated in the receiver, is shifted stepwise against the code sequence received

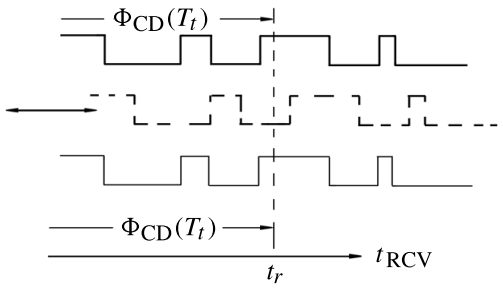

Figure 7.33. Code phase measurement

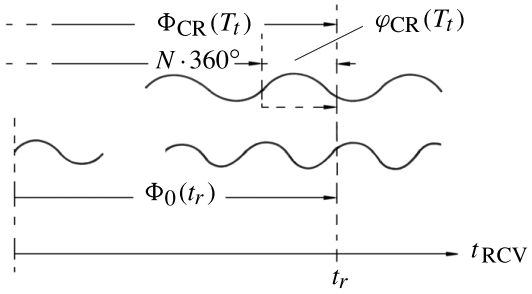

Figure 7.34. Carrier phase measurements

from the satellite, until maximum correlation is achieved [7.2.2]. At the moment of maximum correlation the code phase, $\Phi_{CD}(T_t)$, of the internal code sequence is measured in the receiver time frame, giving t_r. This code phase is identical to the code phase transmitted from the satellite, if we neglect propagation delays within the receiver. We hence obtain the epoch of transmission, T_t, of the code state in the satellite time frame. The difference between both clock readings yields the pseudorange, i.e.

$$PR = c(t_r - T_t). \tag{7.43}$$

Defining:

dt_s satellite clock error with respect to GPS system time,
dt_u clock synchronization error,
dt_a atmospheric propagation delay,
ε_R observation noise, and
R slant range

we obtain the more developed observation equation for *code measurements* (cf. (7.35)):

$$PR_{CD} = c(t_r - T_t) = R + cdt_u + cdt_a + cdt_s + \varepsilon_R. \tag{7.44}$$

In addition, propagation delays in the satellite or the receiver hardware may be included; they cannot be separated from clock errors, and hence are included in the respective clock model. The signs of the different terms come from convention and they differ in the literature.

Note the iterative character of equation (7.44) because the slant range, R, between receiver, B, and satellite, S, at the epochs of transmission and reception is given by

$$R^2 = (X_S(T_t) - X_B(t_r))^2 + (Y_S(T_t) - Y_B(t_r))^2 + (Z_S(T_t) - (Z_B(t_r))^2, \tag{7.45}$$

with

$$T_t = t_r - \frac{R}{c}.$$

The process of *carrier phase observation* is illustrated in Fig. 7.34. The observable is the difference between the transmitted and Doppler-shifted carrier phase, $\varphi_{CR}(T_t)$,

defined in the satellite time frame, and the phase of the reference signal, $\Phi_0(t_r)$, defined in the receiver time frame. The "observed" relative phase is:

$$\varphi_m(t_r) = \varphi_{\mathrm{CR}}(T_t) - \Phi_0(t_r). \tag{7.46}$$

The carrier phase can be written, using the relation (7.42), as

$$\varphi_{\mathrm{CR}}(T_t) = \varphi_m(t_r) + \Phi_0(t_r) = \varphi_m(t_r) + t_r f_0. \tag{7.47}$$

With the substitution (see Fig. 7.34),

$$\Phi_{\mathrm{CR}}(T_t) = N \cdot 360^\circ + \varphi_{\mathrm{CR}}(T_t), \tag{7.48}$$

and with N as the integer ambiguity, we obtain one expression for the epoch of transmission, T_t, of the carrier phase signal, referred to the satellite time frame:

$$T_t = \frac{\Phi_{\mathrm{CR}}(T_t)}{f_{\mathrm{CR}}} = \frac{\varphi_{\mathrm{CR}}(T_t) + N}{f_{\mathrm{CR}}} = \overline{T}_t + \frac{N}{f_{\mathrm{CR}}}. \tag{7.49}$$

The pseudorange from carrier phase measurements is then

$$PR_{\mathrm{CR}} = c(t_r - \overline{T}_t). \tag{7.50}$$

With the ambiguity term,

$$c \cdot \frac{N}{f_{\mathrm{CR}}} = N \cdot \lambda_{\mathrm{CR}},$$

the observation equation for *carrier phase measurements*, corresponding to (7.44), becomes

$$PR_{\mathrm{CR}} = R + cdt_u + cdt_a + cdt_s + c\left(\frac{N}{f_{\mathrm{CR}}}\right) + \varepsilon_R. \tag{7.51}$$

Note that the clock parameters, dt_u, dt_s, the ambiguity term, N, and the hardware signal delays are linear dependent. Hence, ambiguity fixing is not a trivial problem (Wübbena et al., 2001b). Either the parameters are eliminated by forming differences, or the singularity is carefully treated in the parameter estimation process [7.3.2.2].

7.3.2 Parameter Estimation

7.3.2.1 Linear Combinations and Derived Observables

We can see that both observables, carrier phases and code phases on both frequencies, lead to pseudoranges. It may therefore be of advantage to use all observables, or linear combinations thereof, in the parameter estimation process. In principle, an unlimited number of possibilities exists, to combine the different observables, and to form derived observables, but only some combinations are meaningful in the context of positioning. We distinguish combinations

– between observations at different stations,

– between observations of different satellites,
– between observations at different epochs,
– between observations of the same type, and
– between observations of different type.

One advantage of the use of derived observations is that errors that are present in the original observations are eliminated or reduced when differences are formed between observables. In some cases the ambiguities of derived observations are easier to solve than for those of the original observations. On the other hand, the noise level may be considerably increased on combination. The use of linear combinations in the parameter estimation process must hence be thoroughly studied. Common linear combinations are between stations and satellites. Following Fig. 7.35 we introduce

2 receivers i, j,
2 satellites p, q,
epoch t_1: position 1 of the satellites p, q,
epoch t_2: position 2 of the satellites p, q, and
8 pseudorange measurements:

$$PR^p_{1i},\ PR^p_{2i},\ PR^q_{1i},\ PR^q_{2i};\ PR^p_{1j},\ PR^p_{2j},\ PR^q_{1j},\ PR^q_{2j}\,.$$

These may be either code or carrier phase observations.

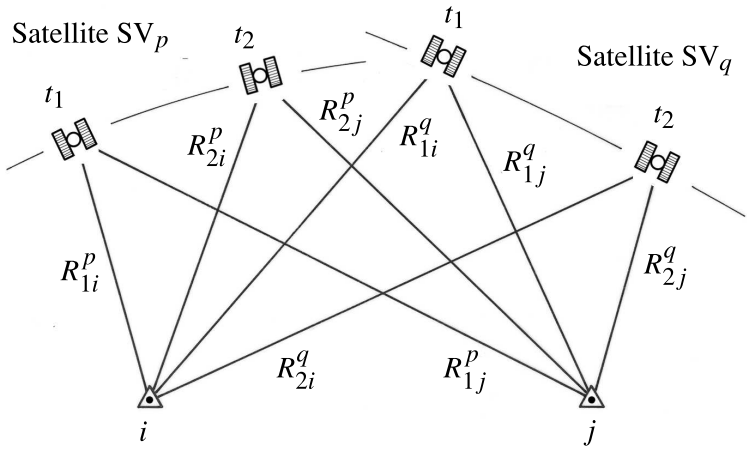

Figure 7.35. Differencing between receivers and satellites

Single differences can be formed between two receivers, between two satellites, or between two epochs. Following the notation of Wells (ed.) (1986) and Fig. 7.35 we introduce the differencing operators

for between-receiver single differences:

$$\Delta(\bullet) = (\bullet)_{\text{receiver } j} - (\bullet)_{\text{receiver } i}\,, \tag{7.52}$$

for between-satellite single differences:

$$\nabla(\bullet) = (\bullet)^{\text{satellite } q} - (\bullet)^{\text{satellite } p}\,, \tag{7.53}$$

for between-epoch single differences:

$$\delta(\bullet) = (\bullet)_{\text{epoch 2}} - (\bullet)_{\text{epoch 1}}. \tag{7.54}$$

Between-epoch single differences (7.54) correspond to the well-known *Doppler solution* [6.1].

In GPS geodesy single differences between two receivers are usually formed, i.e. pseudorange measurements at one station are subtracted from simultaneous pseudorange measurements to the same satellite at a second station. The observables are

$$((PR^p_{1i} - PR^p_{1j}), (PR^p_{2i} - PR^p_{2j}), (PR^q_{1i} - PR^q_{1j}), (PR^q_{2i} - PR^q_{2j})). \tag{7.55}$$

For code phases, equation (7.52) reads with (7.44) (the unchanging indices are omitted for simplicity) as

$$\Delta PR_{\text{CD}_{ij}} = \Delta R_{ij} + c(dt_{u_j} - dt_{u_i}) + c(dt_{a_j} - dt_{a_i}) + c(dt_s - dt_s) + \varepsilon_{\Delta\text{CD}}, \tag{7.56}$$

or, in a simplified notation (e.g. Wells (ed.), 1986), as

$$\Delta PR = \Delta R + c\Delta dt_u + \Delta d_a + \varepsilon_{\text{CD}}. \tag{7.57}$$

The atmospheric delay, d_a, is often separated into the ionospheric and the tropospheric parts d_{ion}, d_{trop}, so that

$$\Delta PR = \Delta R + c\Delta dt_u + \Delta d_{\text{ion}} + \Delta d_{\text{trop}} + \varepsilon_{\text{CD}}. \tag{7.58}$$

For carrier phases we obtain, with (7.51):

$$\begin{aligned} \Delta PR_{\text{CR}_{ij}} = {} & \Delta R_{ij} + c(dt_{u_j} - dt_{u_i}) + c(dt_{a_j} - dt_{a_i}) \\ & + c(dt_s - dt_s) + \lambda_{\text{CR}}(N_i - N_j) + \varepsilon_{\Delta\text{CR}}, \end{aligned} \tag{7.59}$$

or, in its abbreviated form:

$$\Delta PR_{\text{CR}} = \Delta R_{ij} + c\Delta dt_{u_{ij}} + c\Delta dt_{a_{ij}} + \lambda_{\text{CR}}\Delta N_{ij} + \varepsilon_{\Delta\text{CR}}. \tag{7.60}$$

In a simplified notation (e.g. again Wells (ed.), 1986) the between-receiver single differences read:

$$\Delta\Phi = \Delta R + c\Delta dt_u - \Delta d_{\text{ion}} + \Delta d_{\text{trop}} + \lambda\Delta N + \varepsilon_\Phi. \tag{7.61}$$

It becomes evident that the satellite clock error, dt_s, has disappeared, and that the errors in the propagation delay, dt_a, only affect the range measurements with the remaining differential effect. For stations close together, dt_{a_j} and dt_{a_i} may be regarded as equal, and they consequently vanish from equations (7.56) and (7.59). The same is true for the effect of orbital errors (cf. [7.4.3]).

If single differences between satellites are formed, i.e. the observations of two satellites simultaneously recorded at a single station are differenced, the receiver clock term, dt_u, in (7.58) and (7.60) cancels.

With single differences between two epochs for the same satellite the ambiguity term, N, from equation (7.51) cancels because the initial phase ambiguity does not change with time (as long as no cycle slips occur (cf. [7.3.3.1]).

Double differences (DD) are usually formed between receivers and satellites. They are constructed by taking two between-receiver single differences (7.57), (7.61) and differencing these between two satellites, SV_p and SV_q. The derived observables are (Fig. 7.35):

$$((PR_{1i}^{p} - PR_{1j}^{p}) - (PR_{1i}^{q} - PR_{1j}^{q})), ((PR_{2i}^{p} - PR_{2j}^{p}) - (PR_{2i}^{q} - PR_{2j}^{q})). \quad (7.62)$$

The observation equation for carrier phases develops, from (7.60), as

$$\begin{aligned} \nabla\Delta PR_{CR} = {} & (\Delta R_{ij}^{p} - \Delta R_{ij}^{q}) + c(\Delta t_{u_{ij}} - \Delta t_{u_{ij}}) + c(\Delta t_{a_{ij}^{p}} - \Delta t_{a_{ij}^{q}}) \\ & + \lambda_{\mathrm{CR}}(\Delta N_{ij}^{p} - \Delta N_{ij}^{q}) + \varepsilon_{\nabla\Delta}, \end{aligned} \quad (7.63)$$

or in a simplified notation, as

$$\nabla\Delta\Phi = \nabla\Delta R - \nabla\Delta d_{\mathrm{ion}} + \nabla\Delta d_{\mathrm{trop}} + \lambda\nabla\Delta N + \varepsilon_{\Phi}. \quad (7.64)$$

The equivalent equation for code measurements is:

$$\nabla\Delta PR = \nabla\Delta R + \nabla\Delta d_{\mathrm{ion}} + \nabla\Delta d_{\mathrm{trop}} + \varepsilon_{\mathrm{CD}}. \quad (7.65)$$

Note that the receiver clock term, dt_u, vanishes. The double difference observables are free from satellite and receiver clock errors and include only reduced propagation and orbit errors. The double difference observable is the basic observable in many adjustment models for GPS observations (cf. [7.3.4]). Double differences are also the basic observables in many techniques used for the resolution of ambiguities [7.3.2.3].

Triple differences between receivers, satellites, and time, are constructed by taking two epochs, t_1 and t_2. The derived observables are (Fig. 7.35):

$$((PR_{1i}^{p} - PR_{1j}^{p}) - (PR_{1i}^{q} - PR_{1j}^{q})) - ((PR_{2i}^{p} - PR_{2j}^{p}) - (PR_{2i}^{q} - PR_{2j}^{q})). \quad (7.66)$$

As discussed earlier, the initial cycle ambiguity, N, cancels from the observation equation. What remains is a linear combination of all pseudoranges and the residual propagation biases as well as unmodeled orbital biases. The observation equation for triple differences is identical for code and carrier phases, except for the sign of the ionospheric delay. It is, however, usually not established for code observables. In the simplified notation we obtain directly that

$$\delta\nabla\Delta\Phi = \delta\nabla\Delta R - \delta\nabla\Delta d_{\mathrm{ion}} + \delta\nabla\Delta d_{\mathrm{trop}} + \varepsilon_{\mathrm{res}}. \quad (7.67)$$

The last three terms contain the residual propagation biases and unmodeled effects. Triple differences are often used to provide approximate (Doppler) solutions. They are also very useful to aid the removal of cycle slips [7.3.3.1] in an automatic editing process.

For a more complete presentation and for complete expressions see e.g. Wells (ed.) (1986); Wübbena (1991); Leick (1995); Hofmann-Wellenhof et al. (2001).

Linear combinations between observations of the same type can be formed between carrier phases and between code phases. The main uses of such linear combinations are for the elimination of the ionospheric delay [2.3.3.1], [7.4.4.1] and for the resolution of carrier phase ambiguities.

An arbitrary linear combination of the carrier phases on L1 and L2 is formed with integer coefficients, n, m (Wübbena, 1989):

$$\Phi_{n,m}(t) = n\Phi_1(t) + m\Phi_2(t). \tag{7.68}$$

With (2.80) the linear combination fulfils (for a given clock i) the equation

$$t^i(t) = \frac{\Phi^i_{n,m}(t)}{f_{n,m}}. \tag{7.69}$$

Here,

$$f_{n,m} = nf_1 + mf_2, \tag{7.70}$$

is the frequency of the derived signal. Scaling with the velocity of light yields the related wavelength:

$$\lambda_{n,m} = \frac{c}{f_{n,m}} = \frac{c}{nf_1 + mf_2}. \tag{7.71}$$

The ambiguity of the linear combination is then

$$N_{n,m} = nN_1 + mN_2, \tag{7.72}$$

i.e. $N_{n,m}$ is an integer if n and m are integers. Based on these formulas, the first order ionospheric effect and the resulting influence on the combined phase can be computed as (Wübbena, 1989):

$$\delta\Phi_{n,m,I} = -\frac{C_I}{f_1 f_2}(nf_2 + mf_1), \tag{7.73}$$

where C_i is a function of the total electron content (cf. [2.3.3.1], [7.4.4.1]).

The corresponding ionospheric influence on signal propagation time can be characterized with an amplification factor, V_I, as

$$\delta T_{n,m,I} = \frac{\delta\Phi_{n,m_I}}{f_{n,m}} = -\frac{C_I}{f_1 f_2}\frac{nf_2 + mf_1}{nf_1 + mf_2} = -\frac{C_I}{f_1 f_2}V_I. \tag{7.74}$$

The standard deviation of the original phase observation, σ_Φ, can be propagated into the standard deviation of the combined phase observation, with

$$\sigma_{\varphi_{n,m}} = \sqrt{n^2 + m^2}\sigma_\varphi, \tag{7.75}$$

or scaled into a range:

$$\sigma_{n,m} = \lambda_{n,m}\sigma_{\varphi_{n,m}}. \tag{7.76}$$

Among the unlimited number of linear combinations, only those that fulfill some important criteria for the combined signals are of interest:

– integer coefficients to produce *integer ambiguities*,
– reasonably *large wavelength* to help ambiguity fixing,
– *low ionospheric* influence, and
– limited observation noise.

Table 7.7. Selected linear combinations of carrier phases

Signal	n	m	λ cm	$\lambda_{1/2}$ cm	V_I	σ mm
L_1	1	0	19.0	19.0	0.779	3.0
L_2	0	1	24.4	12.2	1.283	3.9
L_Δ	1	-1	86.2	43.1	-1.000	19.4
L_Σ	1	1	10.7	5.4	1.000	2.1
L_{-12}	-1	2	34.1	34.1	2.168	12.1
L_{32}	3	-2	13.2	13.2	0.234	7.6
L_{43}	4	-3	11.4	5.7	0.070	9.1
L_{97}	9	-7	5.4	2.7	0.004	9.7
L_{54}	5	-4	10.1	10.1	-0.055	10.3
L_{65}	6	-5	9.0	4.5	-0.154	11.2
L_0	-	-	$\approx$5.4	$\approx$2.7	0.000	10.0
L_I	-	-	$\approx$10.7	$\approx$5.4	2.000	20.0

Table 7.7 (Wübbena, 1989; Wanninger, 1994) summarizes the leading properties of some selected linear combinations. The column, $\lambda_{1/2}$, contains the effective wavelength for a receiver with squaring technique on L2. The observation noise of the original observation is taken from the specification in the "Interface Control Document" as 0.1 rad, corresponding to 3 mm (ICD, 1993). Well known linear combinations are the *wide lane*:

$$L_\Delta = L_1 - L_2; \quad \lambda_\Delta = 86.2 \text{ cm}, \tag{7.77}$$

and the *narrow lane*:

$$L_\Sigma = L_1 + L_2; \quad \lambda_\Sigma = 10.7 \text{ cm}. \tag{7.78}$$

The advantage of the wide lane observable, compared with the original observation, is that the ambiguity has to be resolved for a signal with a wavelength four times larger; the related observation noise is, however, six times greater.

The narrow lane has the lowest noise level of all linear combinations and hence yields the best results. Its ambiguity is, however, difficult to resolve (cf. [7.3.2.3]). The narrow lane is mainly used over short interstation distances.

The magnitude of the ionospheric effect in the wide lane and in the narrow lane is equal, but it has an opposite sign. Hence the mean of the wide and the narrow lane

yields the *ionospheric free signal*:

$$L_0 = \frac{L_\Delta + L_\Sigma}{2}. \tag{7.79}$$

L_0 is not related to integer ambiguities (cf. [7.4.4.1]) and is therefore not a suitable signal for very precise solutions. The associated observation noise is rather high.

Finally, the *ionospheric signal*,

$$L_I = L_\Sigma - L_\Delta, \tag{7.80}$$

formed from the difference between wide and narrow lane, is of interest because it contains the complete ionospheric effect. The signal allows a detailed analysis of the ionospheric behavior and is helpful in ambiguity resolution strategies (cf. [7.3.2.3]). The ambiguities of the wide and narrow lanes are:

$$N_\Delta = N_1 - N_2; \quad N_\Sigma = N_1 + N_2 \tag{7.81}$$

i.e. they are not independent. When N_Δ is even, N_Σ has to be even, and when N_Δ is odd, N_Σ has to be odd:

$$N_\Delta \bmod 2 = N_\Sigma \bmod 2. \tag{7.82}$$

This *even-odd condition* implies that when the ambiguity is resolved for one of the two combinations the effective wavelength of the other combined signals is increased by a factor of two. For example, when the ambiguity of the wide-lane, N_Δ, is resolved then the effective wavelength of the narrow lane is 21.4 cm and N_Σ can be resolved much more easily. For a table of effective wavelength factors see Table 7.10. The condition (7.82) is used in the *extra wide laning* technique [7.3.2.3], and generates, under the assumption of vanishing ionospheric differences, an effective wavelength of

$$2\lambda_\Delta = 1.72 \text{ m}.$$

Some linear combinations of Table 7.7 are very close to the ionospheric free signal, but with integer ambiguities, e.g. L_{54}, L_{43} and L_{97}. The related wavelengths of L_{54} and L_{43} are larger than λ_0, hence the ambiguity may be estimated more easily. This is in particular true for L_{54}, because the effective wavelength remains unchanged for squaring receivers.

Linear combinations can also be formed for code phases (Wübbena, 1988). Table 7.8 gives an overview. Only P-code observations can be used because the C/A-code is not available on L2.

Because of the dispersive ionospheric effect [2.3.3] the propagation time of signals and signal combinations is different. If signals are received at exactly the same epoch the related transmission epochs are different. Fig. 7.36 illustrates the situation for selected code and carrier signals with respect to the satellite time frame. Taking the first order ionospheric effect into account, it can be shown (Wübbena, 1988, 1991) that the apparent transmission epochs are identical for L_Δ and C_Σ, as well as for L_Σ

Table 7.8. Code phase linear combinations

Signal	n	m	$-V_I$	$\overline{\sigma}$ [m]
C_1	1	0	-.779	0.47
C_2	0	1	-1.283	0.47
C_Δ	1	-1	1.000	2.68
C_Σ	1	1	-1.000	0.33

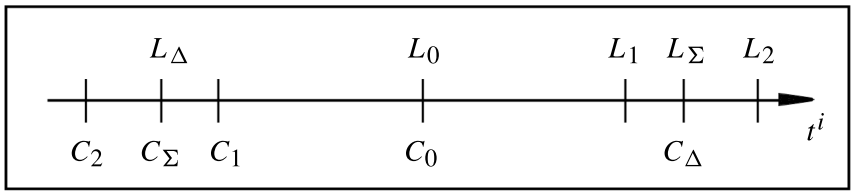

Figure 7.36. Apparent transmission epochs of linear signal combinations for an identical time of arrival

and C_Δ. This opens the possibility to combine code and carrier observations. The unambiguous code signals can be used in the estimation of ambiguities in the carrier signals (cf. [7.3.2.3]).

As part of the GPS modernization program [7.1.7] a third civil frequency, L5, located at 1176.45 MHz, will be available from about 2005. This third frequency can be used to form additional linear combinations, for instance:

$L_1 - L_5$, with $\lambda = 0.75$ m, and
$L_2 - L_5$, with $\lambda = 5.86$ m

These signals will certainly contribute to GPS data processing, although the noise level and ionospheric influences are rather high. Favorable applications are expected for processing short baselines (Hatch et al., 2000).

7.3.2.2 Concepts of Parametrization

As has been stated above, both primary observables, the code phases and the carrier phases, yield pseudoranges between the receiver antenna and a certain number of satellites. Instead of pseudorange the expression *biased range* (range with systematic error effects) is often used. The fundamental difference between both observables is the ambiguity term in equations (7.44) and (7.51). The handling of the ambiguity and bias problem leads to different evaluation concepts. Two main approaches may be distinguished (cf. [4.1]):

(a) parameter estimation, and
(b) parameter elimination.

(a) All systematic influences (biases) that have a stable and well-described structure are *estimated* together with the station parameters as so-called *nuisance* or *bias parameters*. Bias parameters may be, for instance, corrections to the satellite orbit, clock parameters, ambiguity terms, and tropospheric scale factors. These biases can be either directly measured by additional observations (e.g. the ionospheric delay) or they are included in an extended adjustment model (e.g. the tropospheric delay). The *undifferenced phase measurements* (7.39) are used as the basic observables.

(b) Most of the biases are *eliminated* by taking the difference between observables. It is assumed that the disturbing terms are linearly dependent with one another in the various data sets. Up to a certain degree this is correct, e.g. for clock biases, orbital biases and ambiguities. *Single, double* and *triple differences* of the carrier phase measurements are used as derived observables [7.3.2.1]. The concept is primarily applied for *baselines* between two stations.

Both procedures have advantages and disadvantages. The main advantage of the *parameter estimation* approach (a) is its flexibility and independence of the requirement for simultaneous observations at all participating stations. Coordinates of a point field (network) are determined, not derived quantities such as baselines. The behavior of certain parameters, like the comportment of clocks, can be controlled. Highly stable satellite oscillators can be used to improve the stability of the adjustment.

Additional biases like antenna phase center variations, multipath effects or hardware propagation delays can be modeled in a rigorous way and can be directly integrated into the observation equation. Modern approaches like *Precise Point Positioning* (PPP) [7.3.4] are only possible with undifferenced data. In total, it can be stated that carrier phase observations are physically much better represented by undifferenced observables than by double differences. On the other hand, the undifferenced approach requires that all bias and nuisance parameters have to be included explicitly in the observation equation.

The *baseline concept* (b) was introduced from experiences with Very Long Baseline Interferometry (VLBI) (Counselman, Steinbrecher, 1982). Its main advantage is that many common error effects are eliminated from the observations by differencing, thus simplifying the parameter estimation. This is in particular true for the satellite and receiver clock errors, but also to some extent for orbit and signal propagation errors. For larger station separations, however, the elimination process no longer works out in a rigorous manner; it becomes more difficult or even impossible to fix the ambiguities to whole numbers (integers). Also, the number of independent observations is considerably reduced by the differencing, and information is lost.

Through the differencing process all absolute biases are eliminated and only the differences between the biases remain in the data. Differenced observables are no longer single station related but vector related. The error modeling of these differences, however, is much more difficult than is the error modeling for the undifferenced original bias effects (see e.g. Wübbena et al., 2001b). The double difference approach introduces mathematical correlations into the resultant observables. These correlations have to be modeled in the variance–covariance matrix.

The principle of positioning with GPS in the *parameter estimation* approach follows equations (7.44) and (7.51). It is illustrated in Fig. 7.37. Here T_{t_i} is the epoch of signal transmission in the time system of the individual satellite, S_i. The relation to the GPS system time (satellite time frame) is written as

$$T_t = T_{t_i} - dt_{s_i}. \tag{7.83}$$

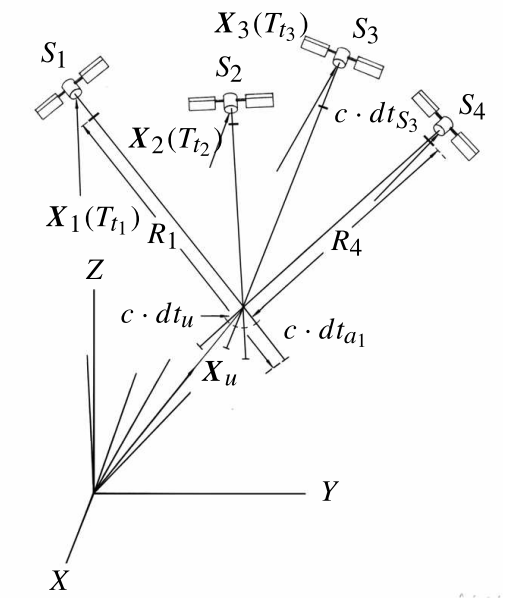

Figure 7.37. Main solution parameters in positioning with pseudoranges

Observations are:

- t_r the epoch of signal reception in the receiver time frame, and
- T_{t_i} the epoch of signal transmission in the individual time frame of satellite, S_i.

Known quantities are the coordinates, $\boldsymbol{X}_i(T_t)$ of the satellite, S_i, at the epoch, T_t, in the CTS coordinate system. In their fundamental simplified formulation the observation equations (7.44) and (7.51) are solved for four unknowns. These are:

- $\boldsymbol{X}_u$ three coordinates (X_u, Y_u, Z_u) of the user antenna in the CTS coordinate system, and
- dt_u clock synchronization error between the user's clock and GPS system time.

The parameter vector of the linearized observation equation is then

$$\boldsymbol{X} = (X_u, Y_u, Z_u, dt_u). \tag{7.84}$$

For a more developed model of parameter estimation, additional parameters can be introduced, for example:

- 3 clock biases per station,
- 3 clock biases per satellite,
- 6 orbital biases per satellite,
- 1 parameter for solar radiation pressure,
- 1 tropospheric parameter per station and satellite, and
- 1 ambiguity parameter per station and satellite.

The parameter for tropospheric propagation delay, dt_{a_i}, is already contained in Fig. 7.37 and in the observation equations (7.44) and (7.51). The clock behavior can be described by a polynomial model with terms for clock bias, drift and ageing (cf. (7.4), [2.2.5]):

$$\begin{aligned} dt_u &= a_{0u} + a_{1u}(t - t_0) + a_{2u}(t - t_0)^2, \\ dt_s &= a_{0s} + a_{1s}(t - t_0) + a_{2s}(t - t_0)^2. \end{aligned} \tag{7.85}$$

The observation equation must be extended accordingly. More parameters can be included if necessary, for example antenna phase center variations or receiver hardware delays.

The singularity in equation (7.51), namely the linear dependency between clock parameters, signal propagation delay, and ambiguity term, has to be treated carefully by fixing certain ambiguity parameters. It is important to choose the correct number of parameters so that the singularity can just be removed, and the ambiguities remain integers (Wübbena et al., 2001b).

The unknown parameters are combined into the vector of unknowns, $\boldsymbol{X}$, and the observations into the vector of observations, $\boldsymbol{PR}$. The linearized observation equation is then written as

$$\boldsymbol{l} = \boldsymbol{PR} - \boldsymbol{PR}(\boldsymbol{X}_0). \tag{7.86}$$

$\boldsymbol{X}_0$ is the approximation vector of the unknown parameters and $\boldsymbol{l}$ the vector of reduced observations. With

$$\boldsymbol{x} = \boldsymbol{X} - \boldsymbol{X}_0, \tag{7.87}$$

it follows that

$$\boldsymbol{l} = \boldsymbol{A}\boldsymbol{x}. \tag{7.88}$$

The design matrix $\boldsymbol{A}$ contains the partial derivatives of the observations with respect to the unknowns:

$$\boldsymbol{A} = (\partial \boldsymbol{PR}/\partial \boldsymbol{X}). \tag{7.89}$$

The solution vector of the system is

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{l}. \tag{7.90}$$

With ε_R as the measurement noise of the pseudoranges (cf. (7.44)), the error ε_x of the adjusted parameters is found to be

$$\varepsilon_x = \boldsymbol{A}^{-1}\varepsilon_R. \tag{7.91}$$

For more detailed information on adjustment techniques see e.g. Vaníček, Krakiwsky (1986, chap. 12), Leick (1995, chap. 4), Strang, Borre (1997, chap. 1) or Niemeier (2002).

In the *parameter elimination* process (b) various combinations of differences are formed, as described in [7.3.2.1]. The formulation of differences implies some consequences that have to be considered for data processing. The most important aspects are briefly discussed.

Algebraic correlation, introduced into the differences, has to be taken into account in a rigorous adjustment. The covariance matrix of observations has to be updated with the increasing number of receivers and receiver types. The problem is discussed e.g. by Beutler et al. (1987); Goad, Müller (1988); Goad (1998) and in most GPS textbooks like Leick (1995); Hofmann-Wellenhof et al. (2001).

A *pre-selection of differences* is necessary if several stations and satellites are involved in the same session, because only one part of the possible differences is

independent. One possible concept is the definition of a *reference satellite* and a *reference station* (e.g. Goad, 1985, 1998; Hofmann-Wellenhof et al., 2001) that have to be introduced into all differences. Difficulties arise with data gaps for the reference satellite or the reference station.

Differencing of observations eliminates the *absolute information* that is contained in the undifferenced observations. A reliable absolute datum has to be introduced into the solution in order to avoid mismodeling [7.6.1.3].

Identical observation epochs are required for the use of differenced observations. For receivers observing at different epochs within the same session, the observations have to be reduced to identical epochs with appropriate interpolation models, e.g. stochastic clock modeling (e.g. Wübbena, 1988).

Because of the simple basic model and the good results in processing short baselines, most commercial software packages use the baseline approach (b) (parameter elimination) with double differences as primary observables. The parameter estimation technique (a) with undifferenced observables is preferred in scientific software packages such as GEONAP and GIPSY-OASIS II. The well-known scientific package BERNESE, however, is also based on the double difference concept [7.3.4].

7.3.2.3 Resolution of Ambiguities

Carrier phase measurements are affected by the ambiguity term N (7.51), that is, by an unknown number of complete wavelengths between the satellite and the receiver antenna. This *initial ambiguity* has to be determined with appropriate techniques to exploit the full accuracy potential of the GPS carrier phase measurements. Ambiguity determination is one of the most demanding problems in the geodetic technique of evaluating GPS observations. On the other hand, it is the integer nature of the phase ambiguities that guarantees the high accuracy of relative positioning with GPS, in particular when the observation time is short.

The best and simplest possibility for determining the ambiguity would be the use of additional frequencies or signals, as is the case for terrestrial electronic distance measurements (e.g. Kahmen, Feig, 1988). Unfortunately, for the time being, GPS does not provide more than two frequencies, hence other strategies were developed to solve the ambiguity problem. The main approaches are (see e.g. Han, Rizos, 1997):

(a) the geometric method (coordinate domain search),
(b) code and carrier phase combinations (observation domain search),
(c) ambiguity search methods (ambiguity domain search), and
(d) combined methods.

Currently the approach (c) is considered to be the most effective and powerful search method, in particular for fast solutions, and is discussed widely in the literature.

(a) *Coordinate domain search*
The *geometric method* makes use of the time-dependent variation in the geometric relation between receiver and satellites. In general, continuous phase measurements are used, and the ambiguities are estimated as real-number parameters. A simple

explanation of the procedure is as follows. Once the satellite signals are identified by the receiver, the whole number of incoming cycles is measured and counted. The unknown initial ambiguity N is maintained throughout the observation process and can be represented by a single parameter (bias). Continuous tracking of the carrier phases results in the determination of ambiguity-free range differences. These are used in a Doppler solution [6.1], to determine the coordinates of the user antenna. Ambiguity-free pseudoranges R_i between the user antenna and the satellites can be derived from the Doppler solution, and compared with the ambiguous range observations. The ambiguities are directly derived from this comparison.

The technique works if the change of geometry is sufficiently large, i.e. with a rather long observation time. If the receiver loses lock on one or more satellites and the remaining number of simultaneously-observed satellites is less than four, a new initial ambiguity has to be introduced. This is the *cycle slip problem* [7.3.3.1]. With more than four satellites, the ambiguities are not independent. A geometrical interpretation of the technique is given with Fig. 7.38.

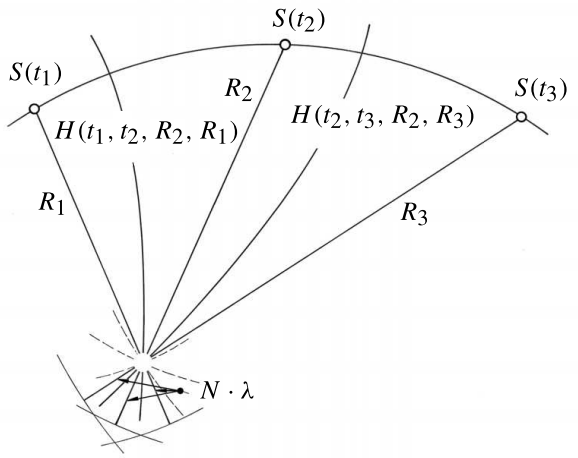

Figure 7.38. Geometric method of ambiguity resolution

The estimated ambiguities are real numbers (*ambiguity float* or *ambiguity free solution*). They can be fixed to integer numbers (*ambiguity fixed solution*) if the estimated values are very close to an integer number, or in other words if the relative position error in the direction of the satellite is smaller than half a cycle length. This requirement leads to the rather long observation time in the pure geometric method. A reduction of the observation time is possible with additional satellites, better geometry, or when signals with a larger wavelength are used (e.g. the *wide lane* L_Δ (7.77)).

One advantage of the geometric method is the clear and simple modeling. For short interstation distances the method usually works out properly and provides reliable results. Problems arise when unmodeled systematic effects remain in the observations, e.g. due to troposphere, ionosphere, and the satellite orbit. The influences of these effects increase with station separations, and make reliable ambiguity fixing difficult or impossible. A wrong determination of the integer ambiguity introduces systematic errors into the coordinates.

The geometric method of ambiguity fixing is one of the earliest methods and has been widely discussed in literature (e.g. Gurtner et al. (1985); Frei, Beutler (1989); Wübbena (1991)). One specific disadvantage of the method is the long observation time necessary to obtain real-value results good enough to resolve the ambiguities. To overcome this problem, at least for short baselines, special observational techniques

have been developed such as "antenna swapping" or "re-occupation" (see [7.3.5.3]).

One particular development, the *ambiguity function* method, was described for the first time by Remondi (1984). The basic idea is as follows: Observables are the *single differences* [7.3.2.1] between two stations where the coordinates of one station are taken as known. Unknown parameters are only the coordinates of the second station and the difference in the receiver clock errors. A search algorithm is defined that varies the baseline vector until the related computed single differences correspond best with the observed single differences. Further developments of this method are described by Mader (1990); Remondi (1990); Han, Rizos (1997) and Hofmann-Wellenhof et al. (2001, p. 229ff). The technique is, from today's understanding, of rather poor computational efficiency and consequently it is of only historical interest (Kim, Langley, 2000). Advantages and disadvantages of the geometric methods are summarized as follows:

Advantages

- basically simple and clear modeling,
- works with few satellites,
- usable for short, long and very long distances, and
- the ambiguity float solution rapidly provides approximate results.

Disadvantages

- long observation time necessary for sufficient geometric rigor,
- influenced by unmodeled effects like ionosphere, orbits, etc.,
- no a priori use of the integer nature of ambiguities, and
- sensitive to unrecovered cycle slips.

(b) *Observation domain search*

In the second approach to ambiguity solution the *combination of code* and *carrier phase* observations is applied. The non-ambiguous code phase measurements are used as an additional wavelength to resolve the carrier phase ambiguity:

$$PR_{CR} - PR_{CD} = \lambda N + dT_A + d\varepsilon. \tag{7.92}$$

This method is independent of the geometry and it is sometimes referred to as the *geometry-free* technique (Hatch et al., 2000). The difference between both observables contains, however, the residual errors dT_A (see below). The basic idea is to make code measurements until the noise level of the code solution is less than half the wavelength of the carrier wave. Because of the much larger (mainly multipath induced) noise in the code measurements, determination of the cycle ambiguity requires that the code observations be smoothed over multiple epochs.

The idea behind this approach was discussed early by Bossler et al. (1980) and Hatch (1982b), but it was not considered to be operational. First results were presented independently by Melbourne et al. (1985) and Wübbena (1985). Today the technique is widely applied, in particular for kinematic applications over short distances (Hatch et al., 2000; Hofmann-Wellenhof et al., 2001). It also promises good results with the forthcoming availability of three frequencies (Vollath et al., 1999).

It is obvious that the method requires a receiver with a low noise level for code-measurements. Even then it is difficult to resolve the ambiguities of the original signals L_1 and L_2, because of their short wavelengths. Instead, the *wide lane* L_Δ (7.77) with 86.2 cm is used. The situation will improve for receivers with a code noise as low as a few centimeters [7.2.4.2].

For proper modeling the different propagation properties of codes and carriers in the ionosphere have to be considered. The signals L_Δ and C_Σ are combined because they have identical propagation times (cf. [7.3.2.1], Fig. 7.36, Wübbena (1988)).

The term dT_A in (7.92) contains the different propagation delays of both signal types due to the satellite and receiver hardware, and in particular due to *multipath effects* [7.4.4.3]. Note that the method is completely independent of the observation geometry, of the satellite and receiver clocks, and of the atmospheric delays. The method hence also works for longer baselines and in kinematic mode.

Once the wide lane ambiguity is correctly resolved, the ambiguities of further signals can be determined with the geometrical method if the ionosphere can be properly modeled; these are for example the ambiguity of the ionospheric free signal or the ambiguity of the narrow lane. The narrow lane is of particular interest to precise applications because of its very low noise level (cf. Table 7.7). It is, however, very difficult to solve for larger distances, because of its very small wavelength, and it is mainly applied in short range applications.

For short interstation distances, where the ionospheric effect can be considered as equal on both stations, the *extra wide laning* technique can be applied (Wübbena, 1988, 1989). Because of the even-odd condition (7.82), the effective wavelength of the wide lane increases to 1.72 m.

To give an example, for short distances the ionospheric signal in the single or double difference must be zero, that is

$$L_I = L_\Sigma - L_\Delta \stackrel{!}{=} 0. \tag{7.93}$$

This condition is only fulfilled if the wide lane ambiguity has been resolved correctly. Let the wide lane ambiguity be estimated as 0.6 cycles, and fixed to 1 cycle. Following the condition (7.93) the narrow lane ambiguity will then be estimated to 8 cycles because

$$\lambda_\Delta \approx 8\lambda_\Sigma. \tag{7.94}$$

This disagrees with the even-odd condition, consequently the wide lane ambiguity was erroneous and will be fixed to 0.

The application of the extra-wide laning technique leads to dramatic savings in observation time, and it is also applicable for ambiguity resolutions "*on the way*" (or "*on the fly*") in kinematic surveying ([7.3.5.4], [7.5.2]). Advantages and disadvantages of the code/carrier methods are summarized as follows:

Advantages:

- independence of geometry,
- kinematic application, and

– long and very long baselines possible.

Disadvantages:

– dual frequency P-code receiver necessary,
– sensitive to multipath, and
– only wide lane ambiguities are resolved.

(c) *Ambiguity domain search*

Ambiguity search methods have been developed with the objective of cutting down the necessary observation time for an individual observation station. The more satellites that are available, the better this method works. The basic idea is to search for the optimum ambiguity combination of L_1, L_2 or derived signals. The search algorithm usually starts with an initial ambiguity float solution and then restricts the solution vector to discrete integer values applying some optimization techniques. The possible combinations within a pre-defined "ambiguity-space" are examined. The procedure is illustrated in Fig. 7.39 for the case of two and three satellites. Each additional satellite restricts the number of possible solutions.

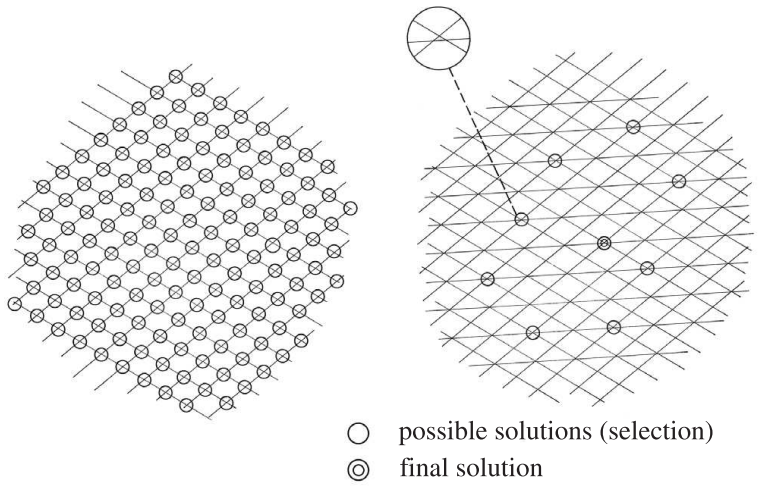

Figure 7.39. Possible solutions for the ambiguities are selected; situation for two satellites (left) and three satellites (right)

The basic problem with this is that the number of necessary mathematical operations increases rapidly beyond all limits. If n is the number of cycles within the search interval, and m the number of ambiguities to be determined, then the number of necessary operations is as given in Table 7.9. It becomes evident that the problem is not solvable by the examination of all possible combinations, but that appropriate selection strategies have to be applied.

Various proposals for selection strategies exist in the literature, and have been introduced in software packages (see e.g. Hatch, 1991; Leick, 1995; Teunissen, 1998; Hofmann-Wellenhof et al., 2001). Early examples are the treatment as a *neural network*

Table 7.9. Number of necessary operations for fixing ambiguities in all possible combinations (Wübbena, 1991)

$n \backslash m$	1	3	10	20
4	1 620	48 020	$3.9 \cdot 10^6$	$5.6 \cdot 10^7$
8	$4.7 \cdot 10^5$	$4.1 \cdot 10^8$	$2.7 \cdot 10^{12}$	$5.7 \cdot 10^{14}$
20	$1.5 \cdot 10^{12}$	$3.3 \cdot 10^{19}$	$1.1 \cdot 10^{29}$	$7.5 \cdot 10^{34}$

(Landau, 1990), or the *Fast Ambiguity Resolution Approach* (FARA) (Frei, Beutler, 1990). A very powerful modern technique is the *LAMBDA method*, developed at the Delft University of Technology (Teunissen et al., 1995; Teunissen, 1998; Joosten, Tiberius, 2000). LAMBDA stands for "Least Squares Ambiguity Decorrelation Adjustment". The basic idea is to transform the original real-valued double difference ambiguities, which are highly correlated, into decorrelated real-valued ambiguities. As such the number of solution candidates is considerably reduced. By this procedure the original highly elongated search space is transformed into a sphere-like search space with the same volume, which allows a much more efficient identification of the integer ambiguities (see Fig. 7.40).

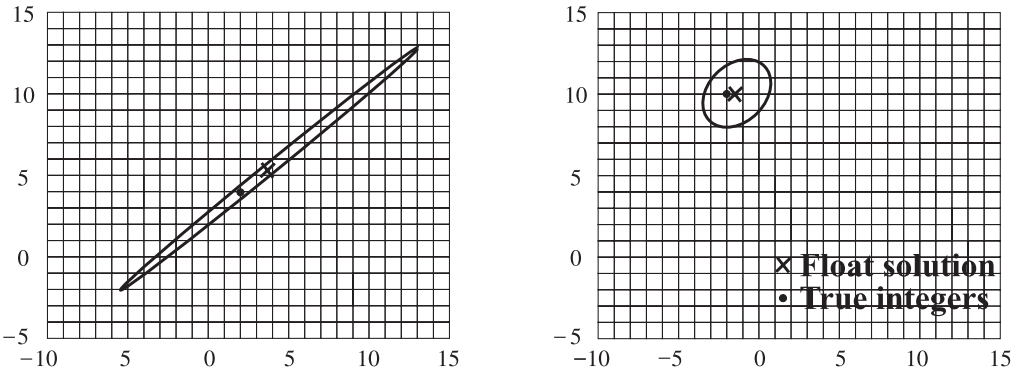

Figure 7.40. Lambda method: ambiguity search space before (left) and after (right) decorrelation (adapted from Joosten, Tiberius 2000)

Other representative techniques are, for example, FAST (Fast Ambiguity Search Filter, Chen, Lachapelle, 1995) and OMEGA (Optimal Method for Estimating Ambiguities, Kim, Langley, 1999). Further algorithms are still under development and will improve the methods of fast ambiguity resolution, in particular for surveying applications over short distances (cf. [7.3.5.2]) but also for long-range real-time kinematic positioning. For a short review of status and trends see e.g. Kim, Langley (2000). Advantages and disadvantages of the ambiguity search methods are summarized as follows:

Advantages:
- allows fast ambiguity resolution (e.g. in rapid static applications),
- true kinematic applications possible, and
- uses the integer nature of ambiguities.

Disadvantages:
- sensitive to systematic errors, and
- requires the observation of as many satellites as possible.

(d) *Combined methods*
These methods include a combination of all the possibilities mentioned above, and should yield the best results. Other techniques may be added. The basic idea is that each fixed ambiguity improves and stabilizes the solution in the next iteration step. This is not only true for the resolution of individual ambiguities but also for fixing of arbitrary integer linear combinations of different ambiguities.

We have seen that the effective wavelength of the wide or narrow lane increases by a factor of two if the ambiguity of one linear combination has been already resolved (7.93), (7.94). Table 7.10 shows the effective wavelength factors for some frequently used linear combinations. If the ambiguity of the "fixed" signal is known, the effective wavelength of the "free" (ambiguity not yet fixed) signal increases by the given factor.

Table 7.10. Effective wavelength factors

Fixed Signal⇒ / ⇓ Free Signal	L_1	L_2	L_Δ	L_Σ	L_{54}	L_{43}	L_I	L_0
L_1	–	1	1	1	4	3	9	7
L_2	1	–	1	1	5	4	7	9
L_Δ	1	1	–	2	1	1	2	2
L_Σ	1	1	2	–	9	7	16	16
L_{54}	4	5	1	9	–	1	17	1
L_{43}	3	4	1	7	1	–	15	1
L_I	9	7	2	16	17	15	–	32
L_0	7	9	2	16	1	1	32	–

Options to use these signals are included in some scientific software packages. The adjustment of a complete network can be improved if the ambiguities for some of the baselines are known a priori. Ambiguities of shorter baselines are in most cases easier to resolve than ambiguities in long baselines. Techniques have been developed for including very short baselines within large scale networks, and then starting with the ambiguity resolution for the short baselines as a first step (Blewitt et al., 1988).

Considerable benefits will be derived from the addition of a third frequency, L5. The new signals $L_1 - L_5$ and $L_2 - L_5$ can be used to resolve ambiguities in a step-wise approach, starting with the longest wavelength. This procedure is also referred to as

cascade ambiguity resolution. For the promises and limitations see e.g. Hatch et al. (2000).

Particular methods have been developed to rapidly determine the initial phase ambiguities for *kinematic surveys* [7.3.5]. One powerful early procedure is the *antenna swapping* technique, i.e. the exchange of two antennas over a very short baseline (several meters) before the start of the kinematic survey (Remondi, 1985). This procedure is described in more detail in section [7.3.5.2]. Today, with the availability of a large number of visible satellites (sometimes eight and more) the ambiguity search methods are in particular a powerful means for the rapid resolution of ambiguities, often based on a single epoch observation.

The resolution of ambiguities is a key factor for precise GPS surveying. In many cases, in particular if the interstation distances are small, and if the data quality is good, the ambiguity resolution works out satisfactorily with the routine options in the software supplied by the manufacturer. In all those cases where

- the interstation distances are large ($> 10\,\text{km}$) and highest accuracy ($\leq 1\,\text{cm}$) is required,
- the data quality is poor (e.g. multipath, cycle slips),
- only a few satellites are visible,
- the ionosphere is disturbed, and/or
- the observation time is short,

problems may arise in solving for ambiguities. In such cases, a careful and interactive data processing operation with multipurpose GPS adjustment software may be necessary. The proper use of the different possibilities, as they have been discussed in this chapter, usually requires trained and experienced personnel.

For precise differential GPS (PDGPS) in reference station networks [7.5.3] a modeling of the error state in real-time helps to reduce the *Time To Fix Ambiguities* (TTFA) and to improve the ambiguity success rate.

One key question is whether the ambiguities have been fixed correctly (*ambiguity validation*). This question can be formalized in a probabilistic measure, the *ambiguity success rate* (Joosten, Tiberius, 2000). Since ambiguities are determined from noisy data, the estimated integer ambiguities can be treated as stochastic variates, similar to standard adjustment practice. If the success rate is sufficiently high, for example 99%, it is likely that the correct integer has been found. For details see Teunissen (1998); Joosten, Tiberius (2000). Other procedures are to compute *contrast* or *ratio* values between the best and the second best solution, namely the two smallest values of the square root sum of residuals. Only when the ratio of these two values exceeds a selected threshold, the solution with the smallest value is chosen as the correct solution (Hatch et al., 2000).

With today's satellite coverage it is possible to extend the observation time at a site up to several hours or even days. In such cases the *ambiguity float solution* already provides excellent results, also for large interstation distances, without the necessity to resolve the integer ambiguities. For applications like the establishment of fundamental geodetic control or the monitoring of crustal deformation the problem of ambiguity fixing therefore is of minor relevance.

7.3.3 Data Handling

7.3.3.1 Cycle Slips

Cycle slips occur if the receiver loses phase lock of the satellite signal. The reasons for cycle slips may be
observation dependent, e.g.:

- obstructions, in particular for kinematic observations,
- signal noise, in particular caused by multipath and ionospheric scintillation,
- low satellite elevation, causing low signal strength,

or *receiver dependent*, e.g.:

- weak signals, partly caused by signal interference,
- antenna inclination in kinematic application (airplane, ship),
- caused by signal processing.

In a cycle slip, the carrier phase shows a sudden jump by an integer number of cycles; the fractional part of the phase observable remains unchanged (Fig. 7.41). The cycle slip may be as small as one or a few cycles, or contain millions of cycles.

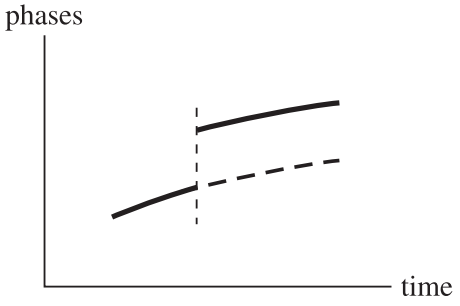

Figure 7.41. Representation of a carrier phase cycle slip

Cycle slips have either to be removed from the data at the preprocessing level, or a new ambiguity has to be determined for the particular pseudorange.

A cycle slip can easily be detected if double and triple differences are formed. This is demonstrated in Table 7.11. The notation corresponds to Fig. 7.35 and to chapter [7.3.2.1]. A cycle slip SL is introduced into the phase observation between station j and satellite p at epoch t. All single and double differences, starting with epoch t, are corrupted by the cycle slip whereas only one of the triple differences is affected. It is evident that the triple difference technique of fixing cycle slips belongs to the very early and classical methods (e.g. Remondi, 1985).

Two aspects have to be distinguished: *cycle slip detection* and the elimination of cycle slips from the data, the *cycle slip repair*, also denoted as *cycle slip fixing*. Most modern receivers have built-in algorithms that identify all or most of the cycle slips, and indicate (*flag*) the slips in the data set. These indications are very helpful for data preprocessing. Advanced receivers also have sophisticated implementations of the phase lock loops with fewer occurrences of cycle slips (Misra, Enge, 2001). Cycle slip repair belongs to the process of data editing, either automatic or interactive.

Several methods are in use and have been widely discussed in the literature. For a review see Gao, Zuofa (1999); Hofmann-Wellenhof et al. (2001) and Bisnath et al. (2001). The main differences come from

- the available data sets (single or dual frequency, codes),

Table 7.11. Effect of a cycle slip SL on single, double and triple differences

Carrier Phases			
$\Phi_i^p(t-2)$	$\Phi_j^p(t-2)$	$\Phi_i^q(t-2)$	$\Phi_j^q(t-2)$
$\Phi_i^p(t-1)$	$\Phi_j^p(t-1)$	$\Phi_i^q(t-1)$	$\Phi_j^q(t-1)$
$\Phi_i^p(t)$	$\Phi_j^p(t)+SL$	$\Phi_i^q(t)$	$\Phi_j^q(t)$
$\Phi_i^p(t+1)$	$\Phi_j^p(t+1)+SL$	$\Phi_i^q(t+1)$	$\Phi_j^q(t+1)$
$\Phi_i^p(t+2)$	$\Phi_j^p(t+2)+SL$	$\Phi_i^q(t+2)$	$\Phi_j^q(t+2)$
Single Differences			
$\Delta\Phi_{ij}^p(t-2)$		$\Delta\Phi_{ij}^q(t-2)$	
$\Delta\Phi_{ij}^p(t-1)$		$\Delta\Phi_{ij}^q(t-1)$	
$\Delta\Phi_{ij}^p(t)+SL$		$\Delta\Phi_{ij}^q(t)$	
$\Delta\Phi_{ij}^p(t+1)+SL$		$\Delta\Phi_{ij}^q(t+1)$	
$\Delta\Phi_{ij}^p(t+2)+SL$		$\Delta\Phi_{ij}^q(t+2)$	
Double Differences		Triple Differences	
$\nabla\Delta\Phi_{ij}^{pq}(t-2)$			
$\nabla\Delta\Phi_{ij}^{pq}(t-1)$		$\delta\nabla\Phi_{ij}^{pq}(t-1,t-2)$	
$\nabla\Delta\Phi_{ij}^{pq}(t)-SL$		$\delta\nabla\Phi_{ij}^{pq}(t,t-1)-SL$	
$\nabla\Delta\Phi_{ij}^{pq}(t+1)-SL$		$\delta\nabla\Phi_{ij}^{pq}(t+1,t)$	
$\nabla\Delta\Phi_{ij}^{pq}(t+2)-SL$		$\delta\nabla\Phi_{ij}^{pq}(t+2,t+1)$	

- the available a priori knowledge (station coordinates, satellite coordinates), and
- the kind of observations (static or kinematic).

Of particular interest are methods that can be applied for a single receiver and hence can be used for single station preprocessing.

Cycle slips cannot be identified and corrected from a single series of phase measurements alone. Basically, the phase observations are compared and combined with other quantities, and the behavior of the differences is analyzed. The differences can be regarded as "test quantities" that must show a smooth behavior. Any discontinuity in the time series of the test quantities indicates a cycle slip. Different approaches are in use to identify the discontinuities:

- a low degree *polynomial* is fitted to the time series; this method is widely used,
- a dynamic model is set up to predict subsequent observations by *Kalman filtering* (in new approaches also by *wavelets*); a comparison between predicted and observed data indicates possible cycle slips, and
- a scheme of first, second, third and fourth differences can be set up; discontinuities show rather strong signals in the higher order differences.

The methods applied have not changed much since the 1980s, hence it is worth studying the cited publications. In the following some of the principal procedures are described:

(a) analysis of double differences and computed ranges,
(b) analysis of ionospheric residuals,
(c) analysis of code/carrier combinations, and
(d) methods usable in kinematic applications.

(a) *Analysis of double differences and computed ranges*
This method requires knowledge of approximate satellite and station coordinates. Slant ranges R_{ij}^{pq} computed from the coordinates are compared with the observed double differences. The test quantity is the residual, $r_{\nabla\Delta}(t)$:

$$r_{\nabla\Delta}(t) = \nabla\Delta\Phi_{ij}^{pq}(t) - \frac{1}{\lambda}\left[R_j^q(t) - R_j^p(t) - R_i^q(t) + R_i^p(t)\right] + \nabla\Delta A. \qquad (7.95)$$

The time dependent behavior of the test quantity, $r_{\nabla\Delta}(t)$, is analyzed by fitting a curve to the series, e.g. a low degree polynomial. A discontinuity in the curve fit identifies the cycle slip.

The term $\nabla\Delta A$ contains the residual atmospheric effect, in particular due to the ionosphere. The ionosphere either has to be modeled, or neglected for short interstation distances. The method hence suffers from ionospheric variations. Other disadvantages are that two stations are involved, and that approximate coordinates are required for stations and satellites. The method is furthermore rather sensitive to observation noise. The main advantage is that the technique works with single frequency receivers. For more details see e.g. Bock et al. (1986); Beutler et al. (1988); Lichtenegger, Hofmann-Wellenhof (1990).

(b) *Analysis of ionospheric residuals*
If dual frequency receivers are available, the differences between the L1 and the L2 signal can be analyzed. Based on a proposal of Goad (1985), see also Cross, Ahmad (1988); Landau (1990), the following test quantity may be used:

$$r_\Phi(t) = \Phi_{L1}(t) - \frac{f_{L1}}{f_{L2}}\Phi_{L2}(t). \qquad (7.96)$$

The main advantage of this test variable is that only frequency dependent parameters are present in equation (7.96). The variable, $r_\Phi(t)$, only depends on the ionosphere (and multipath), and varies very slowly with time. No a priori knowledge of station or satellite coordinates is required. This "geometry-free" method works for a single receiver and can hence be applied in undifferenced phase processing [7.3.4].

If cycle slips ΔN_{L1}, ΔN_{L2} occur on both frequencies they cannot uniquely be identified because they are connected by the Diophantine equation,

$$\Delta N = \Delta N_{L1} - \frac{f_{L1}}{f_{L2}} \cdot \Delta N_{L2}. \qquad (7.97)$$

ΔN is the total cycle slip effect resulting from an analysis of equation (7.96). Equation (7.97) can be solved if the cycle slips in the L1 and L2 signals are known from other methods with an accuracy of about 6 to 8 cycles (e.g. Bastos, Landau, 1988).

(c) *Analysis of code-carrier combinations*
For low-noise code receivers the unambiguous code result can be compared with the phase measurements. The difference between both data sets depends only on the number of integer ambiguities, N:

$$r_{CD,\Phi}(t) = \lambda\Phi + \lambda N - R_{CD}(t). \tag{7.98}$$

The equation holds for both carriers. With a code noise level of the order of ± 0.6 m, the method allows the determination of cycle slips to about ± 3 cycles. This is sufficient for input values to method (b).

If the noise level of the code measurements is sufficiently low, the cycle slips can be determined and removed immediately. The same is true for methods that smooth the P-code measurements over a few epochs and allow quasi-real-time ambiguity resolution through code-carrier combination [7.3.2.3]. In such cases the cycle slip problem no longer exists, because a new ambiguity can be introduced for each cycle slip, and be immediately resolved. The main advantages of code-carrier combination are the simplicity of the underlying model and the possibility of using it in kinematic applications (e.g. Landau, 1988).

A new proposal uses the widelane carrier phase minus the narrowlane pseudorange as test quantity. This observable is geometry-free as well as ionosphere-free. It is insensitive to receiver motion with good results for automatic cycle slip correction both in static and kinematic mode (Gao, Zuofa, 1999; Bisnath et al., 2001).

(d) *Methods usable in kinematic mode*
A particular situation evolves for kinematic surveying, i.e. if at least one of the participating receivers is moving [7.3.5.3], [7.3.5.4]. In early applications, the integer phase ambiguities were resolved before the start of the moving receiver. As long as the phase lock is maintained, kinematic positions of the moving antenna can be derived. If more than four satellites are tracked, and at least four satellites continue to stay on lock, the cycle slips can be corrected with respect to the remaining satellites, and the survey can continue.

Loss of carrier lock is detected by most GPS receivers within 20 milliseconds. Once a momentary loss of lock is detected a multichannel state-of-the-art GPS receiver should be able to determine the number of carrier cycles which were lost when track has again been established. The redundancy available in the additional satellites permits the reconstruction in the missing carrier cycles. In addition, if separate tracking of both carriers is performed within the receiver, this information can be used to determine the number of slipped cycles on one carrier, as long as both carriers do not lose lock over the same time period. If less than four satellites are left, the following procedures can be applied to recover the ambiguities.

1. Return to the last coordinated point and start with a new determination of the initial phase ambiguity. This early method is not practicable for marine and airborne applications, and it is today also not used for land applications because of the availability of efficient and rapid ambiguity fixing techniques.

2. With dual frequency receivers the ionospheric residuals (method (b)) can be used, together with a Kalman filter, to predict the dynamic behavior of the observation platform and to close the gaps in the observations. The technique is very sensitive to high dynamics and low signal/noise ratio.
3. A new ambiguity "on the fly" (OTF) is determined after loss of lock. With low-noise code receivers the extra-wide laning technique can be applied. If a sufficient number of satellites can be tracked after the data gap (6 or more) then advanced ambiguity search techniques [7.3.2.3] are a powerful tool. For each cycle slip, a new ambiguity parameter is simply introduced into the adjustment.
4. The integration of an additional sensor helps to bridge the gaps caused by cycle slips. An external atomic clock (rubidium oscillator) replaces one satellite. An inertial sensor package can be used to interpolate the GPS positions if signals to particular satellites are shaded by obstructions (e.g. Colombo et al., 1999; Böder, 2002).

7.3.3.2 The Receiver Independent Data Format RINEX

Each receiver type has its own binary data format, and the observables are defined following the manufacturers' individual concepts. Time tags may be defined in transmission time, or in receiver time; phase measurement may be expressed in whole cycles, or in fractional parts of cycles; code and phase may have different or identical time tags, and satellites may be observed simultaneously or at different epochs.

As a consequence, data of different receiver types cannot easily be processed simultaneously with one particular GPS data processing software package.

To solve this problem, either all manufacturers have to use the same data output format, or a common data format has to be defined that can be used as a data interface between all geodetic receiver types, and the different processing software systems. The first has not been realized to date. However, a successful solution has been found to define and accept a common data format for international data exchange.

Based on developments at the University of Berne, Switzerland, the *Receiver Independent Exchange Format* RINEX was proposed by Gurtner et al. (1989) at the Fifth International Geodetic Symposium on Satellite Positioning in Las Cruzes. The proposal was discussed and modified during a workshop at this symposium, and recommended for international use. More discussions followed in 1989 and 1990, and brought some modifications and extensions to the data format. A review of the historical development is given by Gurtner (1994).

RINEX has indeed been accepted by the international user community and by the community of receiver manufacturers. For most geodetic receivers *translator software* is provided by the manufacturers that converts the receiver dependent data into the RINEX format. In addition, all major data processing software requires RINEX data as an input. RINEX hence serves as a general interface between receivers and multipurpose data processing software.

With RINEX, one of the most serious obstacles to the routine mixing of data from different receiver types is removed. It is an important precondition for large

international cooperative projects like the IGS [7.8.1], and it found its first important application in the EUREF campaign in 1989 for the establishment of the European Reference Frame [7.6.2.1].

Since the first publication, in 1989, several revisions and modifications have been introduced. The current revision status is version 2.10. A detailed document is available, for instance via the IGS server (Gurtner, 2001). The following definitions are taken from this document.

RINEX defines three fundamental quantities in the GPS observables: *Time, Range*, and *Phase*. The *time* of measurement is the receiver time of the *received* signals. It is identical for the phase and range measurements and is identical for all satellites observed at that epoch. It is expressed in GPS time (not in UT).

The *pseudorange* is the distance from the receiver antenna to the satellite antenna, including receiver and satellite clock offsets and other biases:

$$\begin{aligned}\text{Pseudorange} &= \text{Geometrical distance} \\ &\quad + c \cdot (\text{RCVR clock offset} - \text{Satellite clock offset} + \text{Biases}),\end{aligned} \tag{7.99}$$

so that the pseudorange reflects the actual behavior of the receiver and satellite clocks. The pseudorange is written in units of meters and is unambiguous; i.e., C/A code ranges add the correct number of milliseconds to obtain the definition of pseudorange given above.

The *phase* is the carrier phase measured in whole cycles at both L1 and L2. The half-cycles measured by squaring-type receivers must be converted to whole cycles and this fact is noted by the wavelength factor in the header records. The phase changes in the same sense as the range (negative Doppler); i.e. range increases equal phase increases. The phase observations between epochs must be connected by including the integer number of cycles. The phase will not contain any systematic drifts from intentional offsets of the reference oscillator.

The observables are not corrected for external effects like atmospheric refraction, satellite clock offsets, etc. The sign of the *Doppler shift* as additional observable is defined as usual, namely positive for approaching satellites.

The basic RINEX format consists of three ASCII file types:

1. *Observation Data File*
2. *Meteorological Data File*
3. *Navigation Message File*.

Each file type consists of a header section and a data section. The observation file usually contains the data collected by one receiver at one station during one session. Since RINEX version 2 it is also possible to include observation data collected in sequence by a roving receiver during rapid static or kinematic surveys. From the long list of revision details only some major items are indicated:

- inclusion of GLONASS data (since 1997),
- continuous numbering of the GPS week; no rollover (1998),
- inclusion of navigation data from GEO satellites (2000), and
- inclusion of navigation data from LEO satellites (2001).

For detailed information see the cited documents, in particular Gurtner (2001).

RINEX is the international exchange format for the postprocessing of GPS data. For the transmission of data corrections, in real-time, in relative (Differential) GPS applications, a particular data format is available: the RTCM format. Details of this data format are given in the section on Differential GPS (DGPS) [7.5.1.2].

7.3.4 Adjustment Strategies and Software Concepts

All observations made simultaneously and continuously in the course of a particular GPS project are called a *session*. A session may be as short as a few minutes, if fast ambiguity resolution techniques are applicable in small networks, or it may last several hours, or even days, if the highest accuracy is wanted in larger networks. During the development phase, with a limited number of satellites available, a typical observation session lasted between one and three hours. Since continuous worldwide coverage was established in 1993/1994, sessions can last several days. For practical reasons and for analysis purposes it may be advisable to break down the complete data set into individual sessions of several hours, for instance one session per day or three sessions of eight hours each per day. The following observation and evaluation strategies are in use:

(a) *single-station* adjustment,
(b) processing of single *baselines* and subsequent combination of baselines into networks,
(c) processing of all simultaneously-observed data of a single session in a joined adjustment (*multi-station adjustment*), and
(d) combination of several session solutions into a rigorous overall network solution (*multi-session adjustment*).

The *single-station adjustment* (a) provides absolute station coordinates referred to WGS84. Because of the low accuracy [7.4.1], the results are of little interest to geodetic applications, but they often meet the requirements for some tasks in geophysical prospecting, GIS data acquisition, or in remote sensing. The typical application field is navigation (cf. [7.6.2.7]).

In a rigorous geodetic adjustment relative *and* absolute information [7.4.3] is required. This is why a single-station solution is incorporated into many software packages for multi-station post-processing. The single-station adjustment is also used for preprocessing and editing the data (e.g. because of cycle slips, Earth rotation, relativity, ionosphere, troposphere and formation of normal points), before they enter the level of multi-station adjustment. More accurate absolute positions, at the level of a few meters or better, can be achieved if data from several days of observation are used. Along with the accurate modeling of orbits and clocks [7.4.3] the concept of *Precise Point Positioning* (PPP) has been developed (see below).

The *single baseline concept* (b) was widely used in early software development for the processing of GPS data. The observations from two simultaneously-operating receivers are processed in a joint adjustment, mostly in double difference form [7.3.2.1].

Results are the components ΔX, ΔY, ΔZ of the baseline vector and the associated variance-covariance matrix.

The individual baselines can be used as input data for a network adjustment program and combined into larger networks. The procedure is rigorous, if only two GPS receivers observe simultaneously and if all the stochastic information of the complete variance-covariance matrix is exploited. However, if the station pairs are selected from a larger number of simultaneously operating receivers, the possible baseline combinations are not all independent of each other. Fig. 7.42 gives an example for the case of three receivers. If baselines a and b are considered as independent, baseline c is not. It is called a *trivial baseline* because it can already be derived from the results of baselines a and b. A general rule is given in terms of the number of simultaneously operating receivers, r:

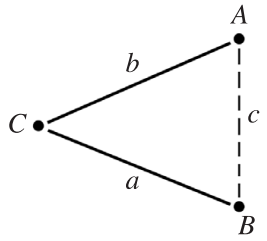

Figure 7.42. Independent and trivial baselines in the case of three simultaneously operating receivers

$$r(r-1)/2 \quad \text{number of possible baselines, and} \tag{7.100}$$

$$(r-1) \quad \text{number of independent baselines} \tag{7.101}$$

If only baseline processing software is available, the independent "non-trivial" baselines have to be identified using suitable selection criteria like the baseline length or number of observations. Nevertheless, the procedure is not rigorous for network solutions because the stochastic information between the simultaneously observed baselines is neglected. Careful weighting and decorrelation is necessary to improve the solution (e.g. Goad, Müller, 1988).

In the simplest case of the baseline concept only the length of the baseline vector is utilized, and GPS is used as a method of distance measurement. This technique was frequently applied in the early years of geodetic exploitation of GPS. The procedure is correct; however one part of the information contained in GPS, namely the spatial orientation of the baseline, is neglected.

Most manufacturers offer software along with the receiver equipment that utilizes the baseline concept. This software is suitable for small projects, for in-the-field data verification and for real-time kinematic (RTK) surveying [7.3.5.4], [7.5.2].

In a *multi-station adjustment* (c) all data that have been observed simultaneously with three or more participating receivers are processed jointly. No baselines are determined, but rather coordinates in a network with the associated complete variance-covariance matrix. Hence, it is a rigorous adjustment of the observations using all mutual stochastic relationships. For geodetic purposes, the multi-station adjustment has conceptual advantages over the baseline approach, because the accuracy potential of GPS is completely exploited. If the observations stem from one session, one speaks of a *session-solution*.

Several session-solutions can be combined into a *multi-session adjustment* (d) or, more precisely, into a *multistation-multisession solution*. This is the usual procedure, if larger networks have to be broken into parts because of a limited number of available GPS receivers. The basic condition is that each session is connected to at least one other session of the network through one or more identical stations where observations have been carried out in both sessions. An increasing number of identical stations increases the stability and the reliability of the total network (cf. [7.6.1.3]).

The multi-session solution is completely rigorous and equivalent to an all-in-one joint adjustment, if the variance-covariance matrices of the individual session solutions are properly used. The stepwise procedure, starting with session solutions, has the advantage of requiring less computer capacity. In addition, comparison of the individual session results provides an excellent insight into the network's accuracy if sufficient redundant observations at identical stations have been included. Software packages for GPS data processing of large networks are usually based on the multistation-multisession concept.

The development in the field of GPS software is fast, hence only a few basic considerations are made here.

A first classification is possible into *commercial software*, provided from receiver manufacturers, and multipurpose *scientific software*, that originates from developments at scientific institutes. Software of the first group is primarily designed for processing of data from a particular receiver type. Advanced packages, however, also accept data from other receivers via the RINEX interface. As a rule, only the executable object code is available to the user, and the basic mathematical models are mostly not documented in detail. Commercial software is adequate for everyday surveying work. It usually offers a large variety of possible applications and can be operated easily enough by personnel with an average background in engineering and GPS technology. In some cases the basic software includes baseline adjustment, and additional software is necessary for network computation. Usually this kind of software allows for static and kinematic [7.3.5] applications and includes extensive mission planning capabilities [7.6.1.1]. The "Real-Time Kinematic" (RTK) capability with OTF ambiguity techniques (see [7.3.5.4]) is today considered to be a standard option. Current examples of this first group are

SKI-Pro	from Leica Geosystems
TGO	(Trimble Geomatics Office) from Trimble Navigation, and
Pinnacle	from Javad/Topcon Positioning Systems.

The development of a general-purpose GPS post-processing system (second group) is a major operation (Beutler et al., 1990). It requires several man-years of development and consists of a large number of individual programs, adding up to tens of thousands of lines of code. Usually, these software packages are not restricted to just one receiver type but accept data from a large variety of geodetic receivers. The packages serve in most cases for

- professional standard use in smaller networks for rapid processing,
- professional use in high accuracy surveys, also over large distances,
- scientific use in research and education, and

– data analysis and scientific investigations including geodynamic research and analysis of permanent arrays.

Besides the standard options for rapid processing, these kinds of software packages offer many particular alternatives for scientific processing. Interactive operation is essential. Some packages include options for orbit determination, or the estimation of atmospheric models.

Scientific processing requires a lot of experience and a deep understanding of GPS signals and error behavior. Data processing is particularly difficult if the data are contaminated by ionospheric disturbances (cf. [7.4.4.1]) and when the highest accuracy over large distances is required from noisy data. The mathematical models, and the rationale behind the scientific general-purpose software packages, are in most cases well documented and discussed in published literature. In some cases, the user has access to the source code and can include modifications or new parts. Current examples of this second group of GPS software are

BERNESE	developed at the University of Berne, Switzerland (e.g. Beutler et al., 1988; Hugentobler et al., 2001)
GEONAP	originally developed at the University of Hannover, Germany (e.g. Wübbena, 1989, 1991), and
GIPSY-OASIS II	developed at the Jet Propulsion Laboratory, USA (e.g. Blewitt et al., 1988; Webb, Zumberge, 1993).

A multi-purpose software package consists of several parts. Three main groups can be identified:

– the *pre-processor level* prepares the data for the main processing,
– the *main-processor level* deals with the estimation of unknown parameters, and
– the *post-processor level* summarizes various information in tables or in graphical form, and combines sessions to networks, if required.

Fig. 7.43 shows a simplified functional flow diagram of a generic software package for static multistation - multisession GPS processing. The structure of the GEONAP software is very similar to this diagram.

Starting from the *raw data* of all receivers involved in a single session, these data must be acquired, translated into a readable ASCII format, and tested for rough errors (blunders). In most cases the RINEX format is used as a data interface between receiver and software. RINEX requires additional information that is not always provided by the receivers, e.g. antenna height, approximate station coordinates, meteorological data etc. These data can be introduced into a *database*.

The broadcast message can be separated from the observation data, checked and organized in a session dependent *message file*. Smoothing algorithms [3.3.3.2] for different portions of the message can be applied. At this level, external orbital information, e.g. IGS orbits [7.4.3.2], [7.8.1], can be introduced if required.

Single station solutions from code measurements or carrier-smoothed pseudoranges are usually generated at the level of the *main program*. Necessary data corrections can be applied at this stage; these are, for example, corrections for ionosphere,

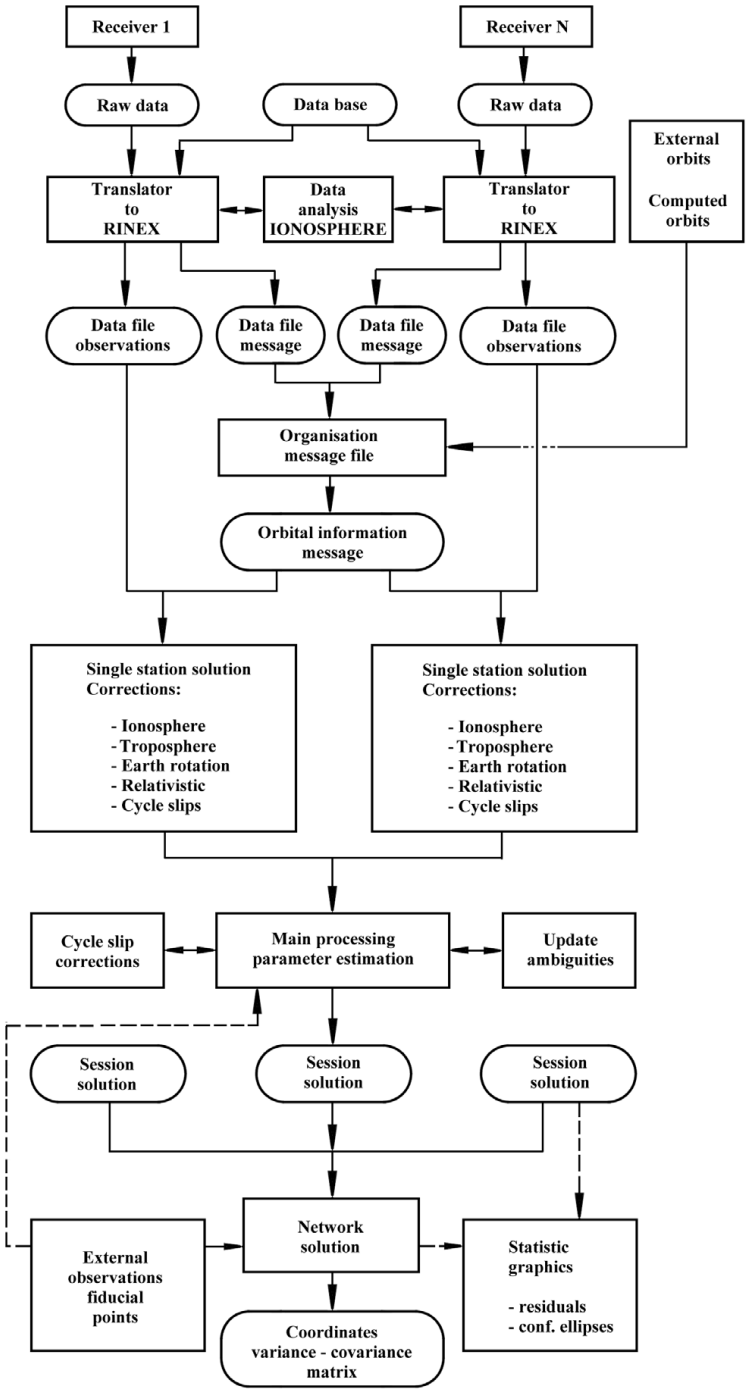

Figure 7.43. Simplified functional flow diagram of a generic GPS software package

troposphere [7.4.4], antenna phase center [7.4.5], Earth rotation and relativistic effects [7.4.1]. The data can be controlled for cycle slips, and be condensed to normal points [3.3.3.2].

Parameter estimation in the *main program* follows either the concept of parameter elimination, or parameter determination, or both [7.3.2]. Basic observables may be the undifferenced phase data (e.g. GEONAP, GIPSY-OASIS II), or double differences (e.g. BERNESE).

If double differences are used, sophisticated decorrelation and weighting techniques have to be applied (Goad, Müller, 1988; Beutler et al., 1990) to account for the mathematical correlations between the double-differences in a GPS network. It may also be advisable to generate an optimum set of independent double difference observables with respect to the shortest interstation distances and minimum influences of data gaps at individual receivers (Rothacher, Mervart, 1996).

Parameters to be estimated can be

- station coordinates,
- biases of satellite clock and receiver clock (second degree polynomial plus stochastic clock model [2.2.5]),
- hardware delays in the satellite and receiver electronics (second degree polynomial plus 1 stochastic parameter),
- orbit improvement with short arc model (e.g. up to 6 Keplerian parameters, solar pressure model),
- tropospheric scale parameter for each station (constant or stochastic process),
- local ionospheric corrections (e.g. improved parameters in the Klobuchar model, time varying parameters of a single layer ionospheric model), and
- parameters for each non-resolved ambiguity and each unrecovered cycle slip.

More parameters can be included. The Bernese GPS software version 4.2 allows, among others, to estimate precise orbits, Earth rotation parameters, precise ionospheric models and the precipitable water content of the atmosphere (Hugentobler et al., 2001). The adjustment process in the main program can be repeated in order to fix as many ambiguities as possible. The principal results of the session solution are the coordinates of all participating stations and the variance-covariance matrix.

At the *post-processing level*, all session solutions can be rigorously combined in a network adjustment program, if this is not already provided in the main program. The results of the network solution can be re-substituted into the session solution with the objective of fixing remaining ambiguities. External information, like coordinates of fiducial stations, can be introduced at this level. Also, improved orbits can be introduced again into the ephemeris file, and a second processing run can be initiated.

The final results from the network solution can be transformed into a local or global datum [2.1.5], [12.1], and compared or combined with existing terrestrial data sets [7.6.2.1]. A final evaluation of the results is supported by statistical tests and a series of graphical representations. These may include statistics on usable simultaneous observations, and the behavior of particular signals and linear combinations, such as the ionospheric signal (7.80).

A new development is *Precise Point Positioning* (PPP). This is a powerful strategy for estimating the coordinates of a single station using precise satellite orbits and satellite clocks. The necessary information can be taken either from the IGS [7.8.1] or from other sources like the JPL [7.4.3.2].

The idea behind PPP is as follows: Precise orbits and satellite clocks are estimated, based on observations from a high quality global fiducial network. This information is taken to solve for station parameters of any site in the world (position, clock and wet troposphere). Only one single station is processed at a time. A disadvantage is that the method is unable to take account for correlations between stations, and that orbits are assumed to be perfect, which is not true. The final formal errors hence have to be scaled to more realistic values.

If only code measurements are used, the observation noise level is well above the precision of orbits and clocks. The positioning accuracy hence depends mainly on the code observation and reaches several decimeters. With the use of phase measurements (e.g. with GIPSY-OASIS II software) the achievable accuracy is comparable to the accuracy in global networks, namely 1 cm or better (e.g. Völksen, 2000).

One particular advantage of the PPP strategy, when compared with network adjustment, is that the processing time increases only linearly with the number of stations. Note that original undifferenced observables are essential for this approach.

For international cooperation in the processing of global and regional networks it is necessary to exchange the results from processing centers, that use different software packages, with the objective to combine global and regional solutions. To this end a *Solution (Software/Technique) Independent Exchange Format*, SINEX, has been defined (IGS, 1996). The SINEX format contains coordinate estimates and the corresponding covariance information, as well as additional information like receiver types, antenna types, phase center values, eccentricities, and a priori weights. SINEX is mainly used by the IGS community [7.8.1].

Information about the latest developments in the software sector can be taken from the proceedings of the series of GPS symposia, such as ION GPS and IAG Symposia, or from related journals (e.g. *GPS World*, *GPS Solutions*, also *Journal of Geodesy*).

7.3.5 Concepts of Rapid Methods with GPS

7.3.5.1 Basic Considerations

Various techniques have been developed in recent years that exploit the capability of GPS to provide precise coordinates after a very short observation time, or even while the receiver (including the antenna) is moving along a trajectory. Sometimes misleading, the related rapid methods were named *kinematic GPS*. In addition, different terms describing particular types of rapid GPS surveying procedures have been created, such as *semi-kinematic, pseudo-kinematic, true kinematic, rapid kinematic, pseudo-static, stop-and-go kinematic* etc. In some cases, different terms were used to describe the same procedure or, more confusing, the same term was used to describe different procedures. The related literature must therefore be read with care. A clarifying overview is given by Kleusberg (1990).

Rapid methods require the resolution of ambiguities in order to exploit the high accuracy potential of GPS phase measurements. Otherwise the noise level of real-valued solutions for the short observation times would be too high. One prerequisite

for the rapid solution of ambiguities is that the distant-dependent errors (see [7.4.3], [7.4.4]) be small. Hence, the rapid methods only work well for short distances (up to several kilometers) between the participating stations. For longer ranges, it is necessary to model the distance dependent errors, e.g. in active reference networks (see [7.5.3]).

Different possibilities exist for subdividing the rapid methods of GPS. The scheme used here is into

- rapid static methods,
- semi kinematic (stop and go) methods, and
- pure kinematic methods.

The rationale behind this subdivision is whether the receiver is taking measurements while it is in motion, and the coordinates of the trajectory can be determined (*kinematic mode*), or whether the receiver is switched off during transportation, and coordinates can only be determined when the antenna is stationary (*static mode*). A third mode is in between these possibilities, in that the receiver has to maintain lock during the times of transportation, but coordinates are not usually derived for the trajectory (*semi-kinematic mode*).

A further distinction between static and kinematic surveying can be seen with respect to accuracy issues (Kleusberg, 1990). In static GPS surveying, most random measurement errors are absorbed in the residuals after adjustment, while in kinematic surveying, most random measurement errors are absorbed in the coordinates. This is why the accuracy potential of static GPS cannot completely be reached with pure kinematic methods.

Only precise methods are considered here, i.e. with an accuracy level of a few centimeters, for kinematic surveys. This implies the use of carrier phase data as the basic observables. Less accurate methods for determining coordinates of a trajectory, i.e. when code measurements are used as primary observables, are discussed in [7.3.6] and [7.5.1]. The dividing line between kinematic and navigational use of GPS is debatable.

As in nearly all geodetic applications at least two receivers are needed to determine relative coordinates. In the concept of rapid methods, one receiver usually remains fixed during the operation, while a second receiver, the *roving* receiver, moves between stations or along a trajectory.

The first two methods were frequently applied after 1990, when GPS had developed into a powerful technique for detailed surveying (cadaster, GIS). With improving satellite coverage after 1994, and the availability of rapid OTF algorithms, currently mainly the third (pure kinematic) method is applied for local surveys. The stop-and-go technique has nearly disappeared from use. It is, however, explained in this book because of its conceptual importance and because it is still offered by a number of manufacturers.

7.3.5.2 Rapid Static Methods

Two different modes can be distinguished (Fig. 7.44):

(a) rapid static mode with single station occupation

(b) rapid static mode with station *re-occupation* after about one hour.

In the first mode (a) fast ambiguity resolution techniques are required (cf. [7.3.2.3]). These can be for example

- code/carrier combination with dual frequency, low code-noise receivers, and
- ambiguity search methods with 6 and more satellites.

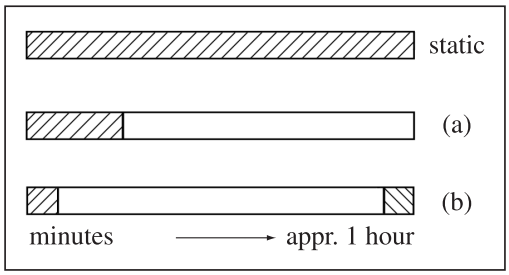

Figure 7.44. Modes of rapid static surveying

Basically, the same techniques are used as for classical static positioning. Depending on the receiver type, satellite coverage, and interstation distance, observation times of several minutes up to 15 minutes are sufficient. The method is particularly powerful over short distances, with dual-band low noise receivers, a high number of visible satellites, and fast ambiguity resolution algorithms. The key factor is the necessary time to fix ambiguities (TTFA) and the *ambiguity success rate*. The procedure is very flexible and effective, and is widely used in surveying applications [7.6.2.4], mostly together with near real-time data processing.

Rapid static applications are of particular interest with respect to reference services like SAPOS [7.5.1.3]. In order to augment the possible distance to the nearest reference station, concepts like *active networks* or *virtual reference stations* play an increasing role.

The demand for surveying equipment with rapid static capability will grow further. Equipment will be assessed mainly on the basis of its capability to resolve ambiguities and to provide precise position results after as short a time as possible.

In the second mode (b) each station has to be re-occupied after an interval of about 50 to 120 minutes. The observation time required at each station is relatively short, about four to eight minutes. Tracking during the transitions is not necessary; the receiver might be turned off while traveling. The rationale behind this procedure is that data from a different geometric configuration are required to resolve the ambiguities (geometrical method, [7.3.2.3]), but not because there is a need for extra observations. Ten minutes of data are completely sufficient to absorb most random measurement errors in the adjustment residuals. Both data sets are considered as one set with one cycle slip in between. The same processing software is used as in static GPS surveying; the "cycle slip" can be fixed with triple and double difference techniques.

Cycle slip fixing over more than 30 minutes, however, only works properly if the data quality is high (low noise, low multipath, low ionospheric effects) and if the repeated station occupations are identical (forced centering). A further requirement is that the same satellites have to be observed for both station occupations.

The re-occupation method was widely used in the early years of rapid GPS methods. Today it has nearly disappeared from the surveying market because of its rather complicated procedure and the high efficiency of method (a). Note that it is a static

method, although the procedure is frequently called *pseudo-kinematic*. Other names are *broken-static*, *intermittent-static*, or *snapshot-static*.

7.3.5.3 Semi Kinematic Methods

This procedure can be traced back to the pioneering work of Remondi (1984, 1986), and it is often also referred to as *stop and go surveying* or simply *kinematic surveying*. The fundamental idea is that there is basically no difference between static and kinematic methods once the ambiguities are resolved and maintained. Kinematic surveying can hence be understood as the "transfer of ambiguities from one station to the other". In the concept of "semi-kinematic" or "stop and go" surveying, the antenna is mounted for a short time (seconds to minutes) on the survey marker that is to be coordinated. The required stop-time comes rather from the need to identify the station and to mount the antenna pod vertically than from the need to gather sufficient GPS data at the station.

The trajectory between stations is usually of no interest (other than in the concept of "pure kinematic surveying" [7.3.5.4]). However, it is essential that phase lock to at least four satellites with a geometrically strong configuration is maintained during antenna movement. The fundamental problem is to determine the initial phase ambiguity before the survey starts. This can be achieved with static initialization procedures. Three main techniques have been developed:

(1) determination of a start baseline with a static survey prior to kinematic operations,
(2) short observation on a known baseline, and
(3) antenna swapping.

The *first method* is rather time consuming, but rapid static methods [7.3.5.2] can be applied as well. The *second method* requires a precisely known baseline. Note that the three dimensional Cartesian coordinate differences of the baseline must be predetermined with an accuracy level of a few cm. The method is very fast because only data from about one minute of observations are required. After initialization, one receiver remains at the station and the second (roving receiver) starts with the survey.

The mathematical background is simple. With the notation of (7.64) the double difference carrier phase data, collected for a short time, are given by

$$\nabla\Delta\Phi = \nabla\Delta R + \lambda\nabla\Delta N. \tag{7.102}$$

The range double difference, $\nabla\Delta R$, can be computed because the station coordinates and the satellite coordinates are known. The computed double difference range is subtracted from the observed carrier phase double difference and yields the ambiguity

$$\nabla\Delta N = (\nabla\Delta\Phi - \nabla\Delta R)/\lambda. \tag{7.103}$$

The *third method* (Fig. 7.45) has been widely used because it is fast, precise, reliable, and it does not require a priori knowledge of a baseline. The procedure was first described by Remondi (1985).

Two receivers are set up close-by at two stations. One station should have known coordinates because it serves as a reference station for the survey. The second station can be arbitrarily selected. About one minute of common data are observed. Then, both antennas are exchanged, maintaining phase lock on the satellites, and again about one minute of data is collected. One receiver remains at the reference station, and the second (roving) receiver is taken to the survey markers.

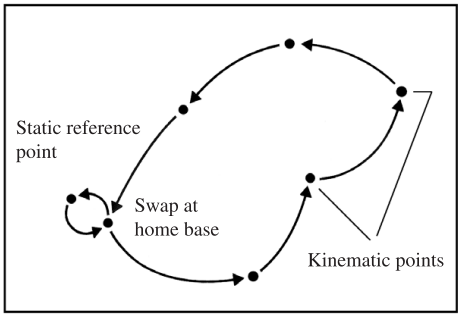

Figure 7.45. Antenna swapping technique

Again, the mathematical background is simple. The first antenna set-up yields the ambiguous double difference phase observation given by

$$\nabla\Delta\Phi(t_1) = \nabla\Delta R(t_1) + \lambda\nabla\Delta N. \tag{7.104}$$

After the antenna exchange the observations provide a second double difference equation:

$$\nabla\Delta\Phi(t_2) = -\nabla\Delta R(t_2) + \lambda\nabla\Delta N. \tag{7.105}$$

Subtraction of the observation equations eliminates the ambiguity term so that

$$\nabla\Delta\Phi(t_1) - \nabla\Delta\Phi(t_2) = \nabla\Delta R(t_1) + \nabla\Delta R(t_2). \tag{7.106}$$

From equation (7.106) the baseline can be determined and substituted into equation (7.103) (known baseline method), to derive the ambiguity term $\nabla\Delta N$.

In essence, equation (7.106) describes the triple difference. Without the antenna exchange, the right-hand side of (7.106) would nearly vanish, because the sum would change to a difference of nearly two identical quantities, and hence could not be solved. The geometrical message of equation (7.106) is that, due to the antenna swap, the geometry has changed enough to solve the triple difference equation after a very short time. A detailed derivation of the antenna exchange technique, with complete expressions, can be taken from Remondi (1985) or Hofmann-Wellenhof, Remondi (1988).

The semi-kinematic method is highly efficient in open areas where no loss of lock, due to signal obstructions, is to be expected. In cases where cycle slips occur, and cannot be recovered, the roving receiver has to go back to the last coordinated survey mark. The stop and go method is still meaningful for local surveys, when two GPS receivers are available but no data link for real-time application. Otherwise, today, RTK techniques [7.3.5.4], [7.5.2] or the use of active reference network services [7.5.3.2] is advisable.

7.3.5.4 Pure Kinematic Method

For many purposes, precise coordinates of the trajectory of a moving GPS receiver have to be determined, in particular in marine and airborne applications. In these cases, a loss of lock without the possibility of recovering cycle slips or ambiguities while the antenna platform is moving cannot be accepted. Hence, methods are required that are independent of static initialization techniques, and that include the capacity to recover cycle slips and/or to resolve ambiguities during motion. These techniques are referred to as ambiguity solution *on the way* (Seeber, Wübbena, 1989), or *on the fly* (e.g. Abidin, Wells, 1990; Cocard, 1995). Only with such methods at hand can kinematic surveying be *purely* or *truly kinematic*. Suitable methods for ambiguity resolution while the receiver is moving are [7.3.2.3]

- code/carrier combination using the *extra wide laning technique*, and
- *ambiguity search functions* for six or more satellites.

The efficiency of these techniques will be improved when low noise code receivers [7.2.5], or combined GPS/GLONASS receivers [7.7.1], are available, as well as new satellite signals [7.1.7].

For the ambiguity resolution on the fly a real-time data link with sufficient capacity is required [7.5.1.2]. Methods for cycle slip recovery in true kinematic mode are

- use of redundant satellites ($\geq$ 4 four satellites),
- use of dual frequency data, and
- use of code/carrier combination.

The inclusion of external sensors can support the recovery of cycle slips and the resolution of ambiguities as well, for example the use of (e.g. Lachapelle, 1990; Colombo et al., 1999; El-Sheimy, 2000; Böder, 2002)

- high quality clocks (e.g. rubidium),
- inertial navigation systems (INS), and
- barometric altimeters.

The accuracy level of pure kinematic surveying is well below 10 cm and can reach a few centimeters under favorable conditions (satellite coverage, low noise receivers, no multipath, low platform dynamics).

The fields of application are broad and continuously broadening [7.6.2]. They include land, air, and ocean surveying, traffic and machine control, engineering-surveying and GIS. In many cases it is sufficient to process data afterwards. Most major software packages offer options for kinematic data. When the results are required in real-time, it is essential to establish a data link between reference and user station. Either the reference data are transmitted from an active network of control points like SAPOS [7.5.1.3], or a particular data link between a local reference station and the roving receiver is established. The latter solution is known as the RTK option (*Real Time Kinematic with OTF (On the Fly) ambiguity resolution*) and it is mainly used for limited ranges (several km). RTK systems are routinely applied for surveying tasks, and they form part of every modern GPS receiver system [7.2.4.2]. For more details about RTK technology see [7.5.2], [7.6.2.4]. Real-time kinematic methods over long distances are still under development.

When accuracy requirements are less demanding, it is not necessary to resolve the ambiguities, and the code measurements can be used as the primary observable. These techniques are discussed in the following section [7.3.6].

7.3.6 Navigation with GPS

GPS was primarily designed as a navigation system with a worldwide real-time capability. The following modes are in use (see also [7.5.1]):

(1) absolute observations with code phases,
(2) absolute observations with code and carrier phases,
(3) relative observations with code phases,
(4) relative observations with carrier-smoothed code phases, and
(5) relative observations with code and carrier phases (the carrier phases are the primary observables).

Absolute observations with code phases (1) are important for general navigation purposes if the accuracy requirements are not too high (cf. [7.4.1]). The observation equations are given with (7.35). After deactivation of Selective Availability (SA) the accuracy level is about 10 m. This is sufficient for most purposes of general navigation, but not for particular tasks in land navigation, in marine geodesy and certainly not in hydrography (cf. [7.6.2.7]). It is therefore not very meaningful to use carrier-smoothed code observations (2) for a single receiver, because a series of observations (relative in time) is affected by time variable effects and also by change in the satellite constellation. Carrier phases can, however, be used for the determination of the instantaneous velocity. Using (6.4)

$$f_r(t) = f_s \left(1 - \frac{1}{c}\frac{dr}{dt}\right),$$

with f_s the frequency emitted from the satellite and $f_r(t)$ the Doppler-shifted frequency, measured in the receiver, it follows that

$$\frac{ds}{dt} = c \cdot \left(1 + \frac{f_r(t)}{f_s}\right). \tag{7.107}$$

In all navigational applications with accuracy requirements better than, say, 10 m, relative observations are essential, either in mode (3) or in mode (4). In both cases, simultaneous observations at a fixed reference station with known coordinates are required (cf. Fig. 7.60, p. 326), [12.3.2].

From the reference observations differences are computed, either between the actual position and the known position or between the actually observed pseudoranges and the ranges derived from the satellite coordinates and the known station coordinates. The differences are transmitted to the moving platform and are used as corrections to the navigation solution. The procedure is known as *Differential GPS* (DGPS) and is treated at length in [7.5.1]. Fig. 7.46 gives a schematic view of a generic navigational software package for differential navigation.

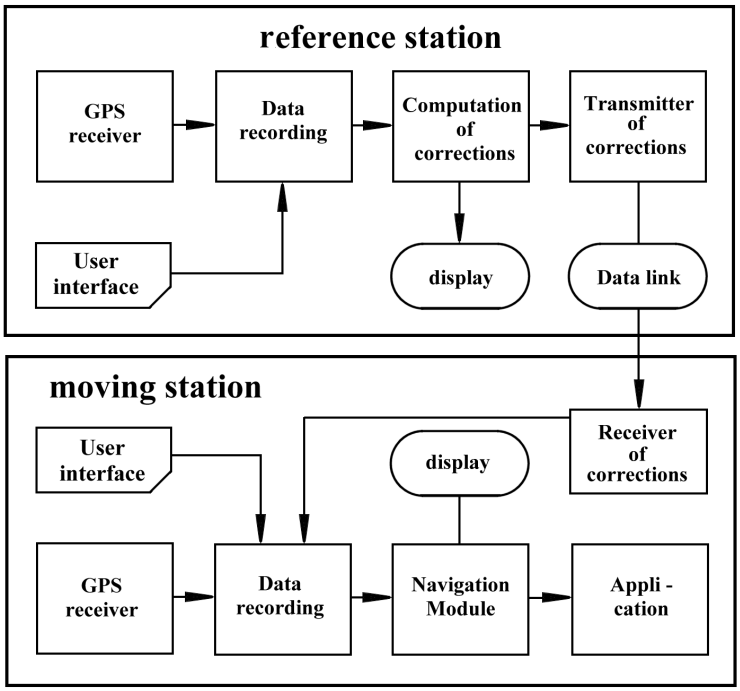

Figure 7.46. Schematic view of a software package for differential navigation

Modern high performance algorithms for shipborne positioning are nearly always based on a combination of code and carrier phase measurements (mode (4)). The carrier phase observations are considered as time-differenced pseudoranges with a much higher accuracy level than the pseudoranges from code measurements. A combination of both observables with proper weighting yields a smoothed series of pseudoranges (Fig. 7.47).

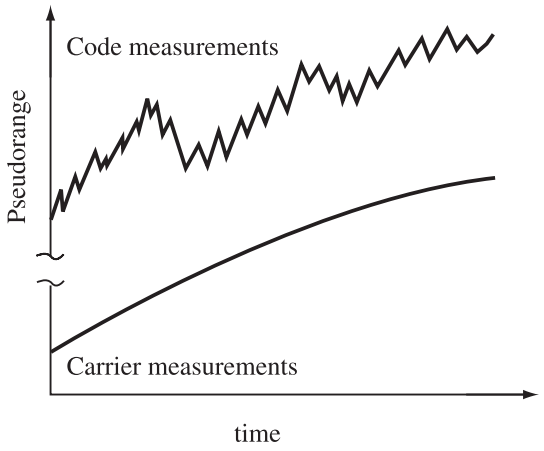

Figure 7.47. Carrier smoothed pseudoranges

The recursive algorithm was proposed by *Hatch* as early as 1982 and is as follows (Lachapelle et al., 1986; Lachapelle, 1991), see also Westrop et al. (1990):

$$PR_\Phi(t) = w_{PR}(t)PR_r(t) + w_\Phi(t)(PR_\Phi(t-1) + (\Phi(t) - \Phi(t-1))), \qquad (7.108)$$

where

$PR_r(t)$ raw observed pseudorange at t,
$PR_\Phi(t)$ phase-smoothed pseudorange at t,
$w_{PR}(t)$ weight of raw pseudorange at t,
$w_\Phi(t)$ weight of $(PR_\Phi(t-1) + (\Phi(t) - \Phi(t-1)))$,

$$w_\Phi(t) = 1 - w_{PR}(t). \tag{7.109}$$

The use of this filter progressively increases the weight of the phase measurements, Φ. In practice, the ambiguity is determined after some time with a resolution of three to five cycles for C/A-code receivers, corresponding to an accuracy level of about 1 meter. Cycle slips may cause discontinuities if less than three satellites remain on lock. In such cases the algorithm has to be reset. Positions and velocities are usually estimated in a Kalman filter approach (e.g. Gelb (ed.), 1974; Schwarz, 1991).

In essence, the code pseudoranges are used for a rough estimation of the position, and the relative carrier phase measurements provide precise position change estimates. These position changes are used to refer all position estimates to one epoch for averaging. The technique works particularly well now after the deactivation of SA. A combination of precise IGS orbits and carrier smoothed pseudoranges provides high precision navigation results with a single receiver (Bisnath et al., 2002).

The mode (5) yields the highest accuracy and is identical to the pure kinematic method [7.3.5.4]. Navigation with GPS is treated in more detail in the sections [7.5.1], [7.6.2.7] and [12.3]. For an overview on the achievable accuracy see Fig. (7.61, p. 327).

7.4 Error Budget and Corrections

7.4.1 Basic Considerations

From a general point of view, errors are introduced into the process of parameter estimation if the modeling is too simple and does not conform to physical reality. The simple concept of pseudorange measurement and navigation in Fig. 7.2 and Fig. 7.29 fails to consider some physical circumstances, or does so insufficiently, namely:

- the Earth-fixed geocentric reference system (CTS) is not an inertial system,
- Newtonian mechanics are not strictly applicable, and
- the signals are not propagating in a vacuum.

It is therefore necessary to correct the satellite coordinates, the satellite clocks, and the observations for

- Earth rotation,
- relativistic effects, and
- ionospheric and tropospheric propagation effects.

In addition, corrections may be necessary for imperfect orbit modeling, signal propagation delays inside the satellite and the receiver hardware, and multipath effects. Finally, the error propagation is affected by the geometric distribution of the satellites. Since most influences that affect the accuracy of GPS measurements have already been treated in full earlier in this book, a summary presentation will be sufficient for this section.

Corrections to the observations are based on measurements or model assumptions. If corrections are applied, we are left with a certain residual error budget (cf. Table 7.12). The error budget can be reduced by refined modeling and by additional observations.

Usually, the contribution of a particular error source is analyzed in terms of its effect on the range determination. The combined effect of ephemeris uncertainties, propagation errors, clock and timing errors, and receiver noise, projected onto the line connecting observer and satellite, is called *User Equivalent Range Error* (UERE) or *User Range Error* (URE). Sometimes the total error is divided into *Signal-in-Space* (SIS) URE, also abbreviated as SISRE, and the *User Equipment Error* (UEE). The rationale behind this division is that the Operational Control Segment (OCS) is only responsible for the SIS performance whereas the UEE depends on the particular user's equipment and correction models.

The SIS URE includes satellite clock and ephemeris prediction errors, OCS state estimate process noise, and some minor residual noise. SIS does not include instantaneous single-frequency ionospheric model errors, tropospheric model errors, receiver noise, receiver antenna phase center variations, or multipath effects. These influences contribute to the UEE.

Official statements can be taken from the document "GPS SPS Performance Standard" (DOD, 2001). This document contains the specific capabilities provided by the *Standard Positioning Service* SPS to all users on a continuous, worldwide basis without any direct user charge. Following the "Federal Radio Navigation Plan" (DOD/DOT, 2001a), access to the *Precise Positioning Service*, PPS, is restricted to U.S. Armed Forces, U.S. Federal agencies and selected allied armed forces and governments (see also [7.1.6].

With disabled *Selective Availability* (SA) the SIS performance for PPS and SPS is nearly identical. According to the cited document the accuracy standard for the SPS Signal-in-Space URE is

$$\sigma_{\mathrm{URE}} \approx 6 \text{ m}.$$

The related accuracy standards for position and height are for a global average positioning domain (95%, SIS only):

$$\begin{array}{rll} \leq & 13 \text{ m} & \text{horizontal, and} \\ \leq & 22 \text{ m} & \text{vertical.} \end{array}$$

For a worst site positioning domain the numbers are (95%, SIS only):

$$\begin{array}{rll} \leq & 36 \text{ m} & \text{horizontal, and} \\ \leq & 77 \text{ m} & \text{vertical.} \end{array}$$

Experiences show that in practice the achievable accuracy is much higher (cf. [7.1.6]). The particular error sources are assigned to three main groups, namely

- satellite position and clock errors,
- signal propagation errors, and
- receiver errors.

Table 7.12 includes average numerical values of the individual error sources as they are generally accepted for operational GPS.

Table 7.12. Main GPS error contributions to the single range observation

Error Source	RMS Range Error
Satellite	
– orbit	1 – 2 m
– clock	1 – 2 m
Signal propagation	
– ionosphere (2 frequencies)	cm – dm
– ionosphere (model, best)	1 – 2 m
– ionosphere (model, average)	5 – 10 m
– ionosphere (model, worst)	10 – 50 m
– troposphere (model)	dm
– multipath	1 – 2 m
Receiver	
– observation noise	0.2 – 1 m
– hardware delays	dm – m
– antenna phase center	mm – cm

Another separation is into

- distance dependent errors (orbit, ionosphere, troposphere), and
- station dependent errors (antenna phase center variation, multipath).

This latter grouping is used together with the error modeling in multiple reference station networks [7.5.3].

The *Earth rotation correction* is necessary if satellite coordinates are computed in an Earth-fixed reference frame at the epoch of signal transmission. During signal propagation from the satellite antenna to the receiver antenna the CTS coordinate system rotates with respect to the satellite; consequently the position of the transmission antenna changes in the rotated CTS system. The original satellite coordinates must be rotated about the Z-axis by an angle, α, which is defined as the product of the propagation time, τ, and Earth's rotational velocity, ω_e (cf. [7.1.5.3]):

$$\alpha = \omega_e \tau. \tag{7.110}$$

Let X', Y', Z' be the original, and X, Y, Z the corrected satellite coordinates, then

$$X = X' \cos\alpha + Y' \sin\alpha, \quad Y = Y' \cos\alpha - X' \sin\alpha, \quad Z = Z'. \tag{7.111}$$

The rotation angle, α, is smaller than $1.''5$. Hence, the trigonometric functions in (7.111) can be replaced by the first elements of a series expansion.

A correction for *relativistic effects* is required because the satellite clock and the main clock by which GPS system time is defined operate at places with different gravitational potential and are moving with different velocities. The relativistic effect causes an apparent frequency shift in the satellite oscillator. The main part of this effect is compensated because the satellite oscillator (10.23 MHz) is operated at a slightly reduced nominal frequency (0.0045 Hz less), see Van Dierendonck et al. (1980). What remains is a small constant component, due to different orbital heights, and a periodic component. The constant effect is absorbed by the satellite clock's drift parameter, a_1, cf. (7.4). Due to the periodic effect the satellite clock reading must be corrected (cf. [7.1.5.3]):

$$\Delta t_r[s] = -4.443 \cdot 10^{-10} e\sqrt{A}[m] \sin E, \tag{7.112}$$

and the drift of the satellite clock:

$$\dot{\Delta} t_r[s] = -4.443 \cdot 10^{-10} \sqrt{A}[m] \cos E \frac{dE}{dt}. \tag{7.113}$$

The effect of the correction (7.112) on the satellite time, for one revolution, is demonstrated in Fig. 7.48. The maximum value can reach 70 nanoseconds in time, and 0.01 nanoseconds/sec for the clock drift. Remaining relativistic effects are compensated in relative observation. For a deeper treatment of the subject see e.g. Ashby, Spilker (1996).

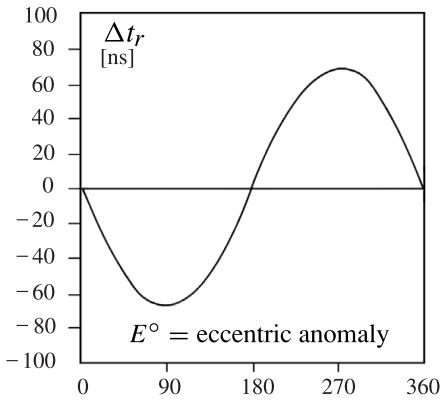

Figure 7.48. Relativistic correction for one satellite revolution

7.4.2 Satellite Geometry and Accuracy Measures

The accuracy of GPS positioning depends on two factors:

- the accuracy of a single pseudorange measurement, expressed by the User Equivalent Range Error (UERE) [7.4.1] or by the associated standard deviation, σ_r, and
- the geometric configuration of the satellites used.

The relation between σ_r and the associated standard deviation of positioning, σ^*, is described by a scalar quantity which is frequently used in navigation and called DOP (*Dilution of Precision*):

$$\sigma^* = \mathrm{DOP} \cdot \sigma_r. \tag{7.114}$$

Different DOP designations are in use:

$$\begin{aligned}
\sigma_H &= \text{HDOP} \cdot \sigma_\text{r} \qquad \text{for horizontal positioning,} \\
\sigma_V &= \text{VDOP} \cdot \sigma_\text{r} \qquad \text{for vertical positioning,} \\
\sigma_P &= \text{PDOP} \cdot \sigma_\text{r} \qquad \text{for 3D positioning, and} \\
\sigma_T &= \text{TDOP} \cdot \sigma_\text{r} \qquad \text{for time determination.}
\end{aligned} \tag{7.115}$$

The combined effect for position and time is called GDOP:

$$\text{GDOP} = \sqrt{(\text{PDOP})^2 + (\text{TDOP})^2}. \tag{7.116}$$

PDOP can be interpreted as the reciprocal value of the volume, V, of a tetrahedron that is formed from the satellite and user positions (Milliken, Zoller, 1980):

$$\text{PDOP} = \frac{1}{\text{V}}. \tag{7.117}$$

Fig. 7.49 gives a geometrical explanation. The best geometric situation exists when the volume is maximized, and hence PDOP in (7.117) is minimized.

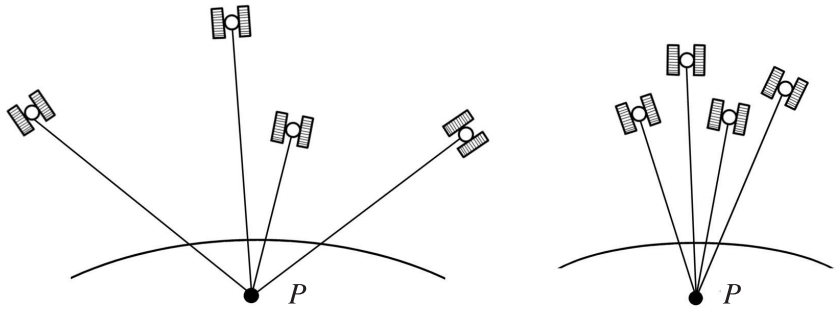

Figure 7.49. Satellite geometry and PDOP

The same result can be derived from the rules of adjustment and error propagation, e.g. Leick (1995); Misra, Enge (2001). Following the equations (7.86) to (7.91) we obtain the covariance matrix for GPS positioning:

$$\boldsymbol{C}_{xx} = \sigma_r^2 (\boldsymbol{A}^\text{T} \boldsymbol{A})^{-1}, \tag{7.118}$$

with the following elements

$$\boldsymbol{C}_{xx} = \sigma_r^2 \begin{pmatrix} q_{xx} & q_{xy} & q_{xz} & q_{xt} \\ q_{yx} & q_{yy} & q_{yz} & q_{yt} \\ q_{zx} & q_{zy} & q_{zz} & q_{zt} \\ q_{tx} & q_{ty} & q_{tz} & q_{tt} \end{pmatrix}. \tag{7.119}$$

From the covariance matrix, the variance of a position determination is found to be

$$\sigma_p^2 = \sigma_r^2 \, (q_{xx} + q_{yy} + q_{zz}), \tag{7.120}$$

or

$$\sigma_P = \sigma_r \cdot \text{PDOP}. \tag{7.121}$$

If we introduce a local Cartesian coordinate system (cf. Fig. 2.8) in which the x-axis points north, the y-axis east, and the z-axis to the zenith, it then follows for the horizontal position error, that

$$\sigma_H^2 = \sigma_r^2 \, (q_{xx} + q_{yy}), \tag{7.122}$$

or

$$\sigma_H = \sigma_r \cdot \text{HDOP};$$

and for the vertical position error, that

$$\sigma_V^2 = \sigma_r^2 \, q_{zz}, \tag{7.123}$$

or

$$\sigma_V = \sigma_r \cdot \text{VDOP}.$$

The fourth parameter t contains primarily the timing error. Hence

$$\sigma_t^2 = \sigma_r^2 \, q_{tt}, \tag{7.124}$$

or

$$\sigma_t = \sigma_r \cdot \text{TDOP}.$$

In general

$$\text{GDOP} = \sqrt{\text{trace}(\boldsymbol{A}^{\text{T}}\boldsymbol{A})^{-1}}. \tag{7.125}$$

Until the full satellite configuration was installed the DOP values for a given location changed considerably during the day, and the best "*observation window*" had to be selected. Since the operational 24-satellite configuration is complete, the variations of DOP values are less critical, and PDOP remains at most times below 3, or even 2.

Today, with the full constellation, the DOP numbers are in general of less interest to surveying applications, because all visible satellites can be tracked with modern receivers, and can be introduced into a rigorous adjustment process. Accuracy estimates result from the adjustment algorithm rather than from PDOP calculations. DOP indicators remain, however, an important tool for survey planning and control, in particular in the rapid static, kinematic or navigational mode [7.3.5], where short time signal blocking caused by obstructions may occur. For a detailed treatment of the DOP-issue see e.g. Spilker (1996c)

The accuracy measure usually applied in surveying and geodesy is the *standard deviation*, σ, mostly considered to be identical with the *root mean square error* RMS.

The probability of a location being within a certain region is described by a *confidence ellipse* (for two dimensions), or a *confidence ellipsoid* (for three dimensions) with the estimated position at its center. The axes of the confidence ellipse are a function of the standard deviation of the particular coordinates (e.g. σ_Φ for latitude and σ_Λ for longitude), and the level of probability. Usually applied levels of probability are 68.3 % (corresponding to 1 σ), 95.5 % (corresponding to 2 σ), or 99.7 % (corresponding to 3 σ). Accuracy numbers throughout this book are usually at the 1 σ level. For more information, see textbooks on adjustment and statistics e.g. Leick (1995); Strang, Borre (1997) or Niemeier (2002).

The accuracy measures usually applied in navigation are quite different. In most cases they are based on the accuracy of so-called *lines of position* LOP in a plane, or "surfaces of position" in three-dimensional navigation. Every navigation system defines its proper system of LOPs. The user position is located at the intersection of two or more such lines or surfaces of position. Some general information is given as follows, without going into details. For further reading see e.g. Forsell (1991) .

The relation between geodetic and navigational definitions is illustrated in Fig. 7.50. A common way to express two-dimensional accuracy is the *Distance Root Mean Square* (DRMS):

$$\mathrm{DRMS} = \sqrt{\sigma_\varphi^2 + \sigma_\lambda^2}. \qquad (7.126)$$

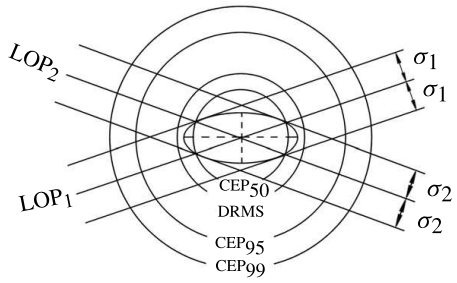

Figure 7.50. Geodetic and navigational accuracy measures

The probability of being within a circle with radius DRMS varies between 63.2 % and 68.3 %. Alternative names for DRMS are *Circular Radial Error* or *Mean Squared Position Error* (MSPE). One parameter which is frequently used (for example in the *U.S. Federal Radio Navigation Plan* and related documents (DOD, 2001; DOD/DOT, 2001a,b)) is the 2 DRMS:

$$2\,\mathrm{DRMS} = 2 \times \mathrm{DRMS} = 2\sqrt{\sigma_\varphi^2 + \sigma_\lambda^2}. \qquad (7.127)$$

The probability level is between 95.4 % and 98.2 %. Note that 2 DRMS must not be confused with 2-D RMS, the *two-dimensional root mean squared error*, that is essentially identical with DRMS (7.126).

The *Circular Error Probable* (CEP) is also widely used for different levels of probability. The most used measure is:

$$\mathrm{CEP} = 0.59\,(\sigma_\varphi + \sigma_\lambda), \qquad (7.128)$$

for 50 % probability. It defines the radius of a circle, centered at the true position, containing 50 % of the estimated positions. Other measures are:

$$\mathrm{CEP}_{95} = \mathrm{CEP} \cdot 2.08, \qquad (7.129)$$

for 95 % probability, and

$$\mathrm{CEP}_{99} = \mathrm{CEP} \cdot 2.58, \tag{7.130}$$

for 99 % probability.

Accuracy measures in use for three dimensions are the *Mean Radial Spherical Error* (MRSE):

$$\mathrm{MRSE} = \sqrt{\sigma_\varphi^2 + \sigma_\lambda^2 + \sigma_h^2}, \tag{7.131}$$

with a probability of 61 %, and the *Spherical Error Probable* (SEP):

$$\mathrm{SEP} = 0.51(\sigma_\varphi + \sigma_\lambda + \sigma_h), \tag{7.132}$$

with a probability of 50 %. The relationship between 2 DRMS, CEP, and SEP is

$$2\,\mathrm{DRMS} = 2.4 \cdot \mathrm{CEP} = 1.18 \cdot \mathrm{SEP}. \tag{7.133}$$

It becomes evident that numbers, indicating accuracies, are only meaningful if the corresponding accuracy measure is identified. As an example, the achievable accuracies under SPS are

$\approx$ 13 m	(φ and λ)	2 DRMS	(95 %),
$\approx$ 5 m	(φ and λ)	CEP	(50 %), and
$\approx$ 11 m	(φ, λ and h)	SEP	(50 %).

Note that all three accuracy numbers describe an identical situation. A short overview of accuracy measures is given by Van Diggelen (1998).

7.4.3 Orbits and Clocks

7.4.3.1 Broadcast Ephemerides and Clocks

Discrepancies between the predicted ephemeris available to the user and the actual orbit propagate into the determined positions of the user antenna. It is evident (Fig. 7.51) that the radial component of the orbital error corrupts the range determination, and hence the location of the user position, to a much higher degree in single station positioning than in relative positioning. For nearby stations, most of the orbit errors are cancelled out in the differencing and in the determination of relative coordinates. As a *rule of thumb* we have for the effect, db, of the orbit error, dr, on the determination of the baseline, b:

$$\frac{db}{b} = \frac{dr}{\rho}. \tag{7.134}$$

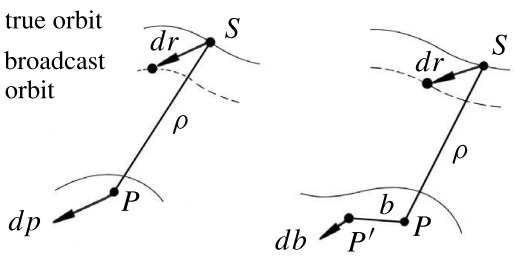

Figure 7.51. Effect of an orbit error on the single-point position (left) and the baseline determination (right)

The baseline error, db, thus depends mainly on the ratio of the baseline length, b, to the satellite range, ρ.

The maximum range between a GPS satellite and the observer is about 25 000 km. If a baseline error of 1 cm is accepted, an admissible orbit error for a specified baseline length is summarized in Table 7.13. The table clearly shows that, for relative coordinate determination over short distances, the required orbit accuracy is not a critical factor. On the other hand, the requirement for 1 cm accuracy over very great distances, for example, in geodynamic applications over 1000 km and more, implies an orbit accuracy of better than 1 m which is not yet provided by the broadcast ephemerides.

Table 7.13. Relation between orbit errors and corresponding 1 cm baseline errors

Baseline length	Admissible orbit error
0.1 km	2500 m
1.0 km	250 m
10 km	25 m
100 km	2.5 m
1000 km	0.25 m

In many cases, the accuracy of the baseline determination is set as a ratio of the base length, estimated in parts per million (ppm). To give an example, an orbit error of 25 m propagates into a relative accuracy of 1 ppm, and an orbit error of 5 m corresponds to 0.2 ppm, i.e. 1 cm over 50 km. The last figure is the critical limit for reference station networks (see [7.5.3]).

The formula (7.134) implies a considerable approximation, and it is widely regarded as too pessimistic (e.g. Zielinski, 1989; Beutler et al., 1998, p. 104). Indeed, the equation is derived from geometrical considerations and only reflects an instantaneous situation. For a whole session, the changing geometric configuration for all satellites has to be included. The resulting error evolves as the sum of all individual satellite orbit errors, integrated over the whole observation period. Zielinski (1989) estimates the resulting baseline errors as too large by a factor of 4 to 10. The "rule of thumb" (7.134) should therefore only be used with care, and for a rough estimation. More detailed equations are given in the cited literature.

Because of the great altitude of the GPS satellites, their orbits are only slightly affected by surface forces and higher-order potential coefficients of Earth's gravity field [3.2.4]. For the computation of short orbit arcs a gravity field expansion up to degree and order (6,6) or (8,8) is sufficient. It is necessary to include the gravitation effects of the Sun and Moon, as well as the Sun's radiation pressure, in orbit computation models. In particular, the non-gravitational forces on GPS satellite orbits have to be modeled carefully (e.g. Fliegel, Gallini, 1989; Fliegel et al., 1992; Beutler et al., 1998; Ziebart et al., 2002), cf. [3.2.3.4]. The ROCK42 model is used for Block II/IIa satellites, and a particular new model for Block IIR satellites (Marquis, Krier, 2000). For precise

computations a model for thermal re-radiation and a particular y-bias may be included (e.g. Rothacher, Mervart, 1996).

Experiences with the accuracy of broadcast orbits (status 1998/2000) indicate a level of 5 to 10 meters (Jefferson, Bar-Sever, 2000). However, there are times where significantly larger orbit errors may occur. Broadcast orbits in 2002 are accurate to about 3 m, following the estimation of the IGS (cf. Tab. 7.31, [7.8.1], p. 402). Along with the Accuracy Improvement Initiative (AII) [7.1.7] an orbit accuracy of about 1 m can be expected.

According to Tab. 7.13, this accuracy is in general sufficient for work with point distances up to several tens of kilometers. When greater point distances are associated with high accuracy requirements, the accuracy of the operational broadcast ephemerides is usually not sufficient. Hence, orbit improvement techniques have to be applied along with the data adjustment in large scale networks [7.3.4], or a posteriori precise ephemerides, based on observations from globally distributed *tracking stations* have to be used [7.4.3.2].

GPS time is operationally defined as the time scale used by the GPS system (cf. [2.2.3]). Each GPS satellite carries clocks which act as the time and frequency base for the realization of the *GPS system time* in the particular satellite. Navigation signals and carrier signals are time-tagged to the particular *satellite time frame*. GPS time is monitored by the Main Control Station (MCS) [7.1.3]. Its relationship to the other atomic time scales is demonstrated in Fig. 2.14, cf. [2.2.3, p. 38].

The satellite clocks run fast by 38.5 μs/day; this correction absorbs more than 99.6% of the relativistic clock effect. If necessary, the MCS applies other corrections to synchronize the individual satellite clocks with the system time. A synchronization error of 1 μs in a satellite clock will produce an error of 300 m in the pseudorange. When meter-level position accuracy is required, the clock synchronization between the satellites must be controlled within a few nanoseconds. This is why rubidium and cesium oscillators are used in the GPS satellites [2.2.5]. These clocks have a short-term stability of 10^{-9} to 10^{-10}, and a long-term stability of 10^{-12} to 10^{-13}. Block II/IIa satellites carry two rubidium and two cesium atomic frequency standards, and each Block IIR carries three rubidium standards.

The performance of each clock is observed by the Control Segment [7.1.3], and one of the clocks is selected to generate the signals. The deviation of a particular clock from GPS system time is modeled as a quadratic function of time. The parameters of this model are estimated, uploaded to the satellite, and are broadcast within the navigation message [7.1.5.3]. The coefficients a_0, $a_1$anda_2 in equation (7.4) are also called the bias, drift, and aging parameters of the clock. Typically, parameter a_0 is below 1 μs; $a_1 \approx 10^{-11}$ s/s, and $a_2 \approx 0$ s/s^2. With these correction terms of the broadcast message the satellite clocks can be kept synchronized within 5 to 10 ns (Misra, Enge, 2001, p. 92).

The actual behavior of each clock slightly differs from this model because of unpredictable, correlated frequency errors. For the highest accuracy requirements, the satellite clock parameters can be estimated in the adjustment process. A stochastic

correlation model (cf. [2.2.5], Wübbena (1989)) can be included for the growth of the random frequency error with time. Alternatively, a posteriori clock models based on observations can be used [7.4.3.2].

The requirements of the *receiver clock* are not very high. The user clock in the receiver need only be stable enough to do the pseudorange measurements with code-phases. A quartz oscillator of medium quality usually suffices. In most geodetic adjustment models, receiver clock errors are eliminated by means of double differences of the carrier phase observations [7.3.2.1].

However, the use of a more precise external clock (e.g. a rubidium oscillator) is of importance in cases where only few satellites are available. The clock substitutes one satellite.

7.4.3.2 Precise Ephemerides and Clocks, IGS

A posteriori precise ephemerides (PE) and clock parameters are based on observations from globally distributed *tracking stations*. At such stations, dual-frequency receivers are installed that can measure both the code phases and carrier phases of all satellites in view. Orbit errors can be separated from the station clocks' time errors through the use of high-precision oscillators (rubidium, cesium atomic standard). The tropospheric propagation delay can be determined with water vapor radiometers.

The data files usually conform to the *SP3 data format* (Standard Product 3), finalized by the U.S. National Geodetic Survey (NGS), (Remondi, 1991; Hilla, 2002). This format is precise to 1 mm and 1 picosecond. Several agencies provide precise ephemerides and adjusted clock parameters, among them the NIMA, JPL, and IGS.

The *National Imagery and Mapping Agency* NIMA (formerly *Defense Mapping Agency* (DMA)) generates precise ephemeris (PE) data files and improved clock parameters based on observations from 20 monitor stations. These are twelve NIMA and five Air Force stations (see Fig. 7.8, p. 217), and three IGS stations (Maspalomas, Kerguelen, Yakutsk).

The ephemerides files give position and velocity vectors for each satellite every 15 minutes. Two PE data types are available, one referred to the satellite's center of mass, and the other with respect to the satellite's antenna phase center. A comparison in 2001 between the IGS final orbits (see below) and the NIMA precise ephemerides showed differences of less than 20 cm.

Precise orbits and adjusted clock parameters can be used either for the post-processing of data in multi-station GPS networks; or they can be used for the processing of single receiver data in the *Precise Point Positioning* (PPP) mode [7.3.4]. The NIMA precise ephemerides are freely available via Internet (anonymous ftp).

Another resource for precise orbits are the NASA JPL (Jet Propulsion Laboratory) precise ephemerides and adjusted clock parameters. Positions and velocities are given for every 15 minutes, and clock parameters for every 5 minutes. Orbits and clocks result from the same estimation process and are hence completely consistent with each other. Final orbits are available after about 2 weeks. Rapid orbits are given within 20 hours; they agree with the final orbits at the level of about 20 cm. The JPL orbits

are also given as *non-fiducial* (NF) orbits, i.e the orbits are estimated in a free datum independent from the ITRF coordinates of the tracking stations. The datum instead is derived from the orbits and clocks. The adjusted network can then be transformed to any other datum without problems. Precise JPL orbits and clocks are primarily required for the processing of single receiver GPS data in the *Precise Point Positioning* (PPP) mode with the software package GIPSY OASIS II [7.3.4]. The necessary accuracy of the adjusted clocks is in the order of 100 picoseconds.

The most important source for precise ephemerides and other GPS products today is the IGS. The IGS, a service established by the IAG, officially started its activities on January 1, 1994, after a successful pilot phase of more than one year. In 1999 the name was changed from *International GPS Service for Geodynamics* to *International GPS Service*. Following the *Terms of Reference*, the primary objective of the IGS is to "provide a service to support through GPS data and GPS products geodetical and geophysical research activities". For more details of the structure, organization, and the various and growing services of the IGS see [7.8.1]. In the following, only the main information on data, orbits and clocks is given.

IGS collects, archives and distributes GPS observation data sets from more than 300 globally distributed stations. The stations have to meet certain quality criteria. About 120 stations are classified as *Global Stations* because they are regularly analyzed by at least three Analysis Centers. The IGS core products consist of weekly final products, namely

- GPS ephemeris and clock values, tabulated at 15-minutes intervals for each day (in SP3 format),
- Earth Orientation Parameters (EOP), and
- Geocentric station coordinates and velocities.

With respect to orbits and clocks, three different products are available (see Table 7.14, status August 2002). The *ultra-rapid orbits (predicted orbits)* are updated twice daily (at 03.00 and 15.00 UT) and are valid for a period of 48 hours. The first 27 hours are based on actual observations and the second 21 hours are a predicted orbit.

Table 7.14. Precise IGS GPS orbits and clocks

Orbits	Accuracy	Latency	Updates	Sample Interval
Broadcast	$\sim$ 260 cm/ $\sim$ 7 ns	real time	–	daily
Ultra-Rapid	$\sim$ 25 cm/$\sim$5 ns	real time	twice daily	15 min/15 min
Rapid	5 cm/0.2 ns	17 hours	daily	15 min/5 min
Final	$<$ 5 cm/0.1 ns	$\sim$ 13 days	weekly	15 min/5 min

With the IGS products at hand, all requirements for precise orbits are completely fulfilled. Together with the precise station coordinates and the original observation data from IGS stations, it is possible to connect every new station worldwide directly to the geocentric reference frame (see [7.6.2.1]). Note that also the individual analysis

centers of the IGS provide precise orbits, for example the NGS in the U.S. and CODE in Europe (see [7.8.1]).

7.4.4 Signal Propagation

The GPS signals, when propagating from the satellite antenna to the user antenna are subject to the following propagation effects:

- propagation delay in the ionosphere,
- propagation delay in the troposphere, and
- multipath propagation at the satellite and in the vicinity of the receiver antenna.

The atmospheric propagation delays are basically treated in section [2.3]. In this chapter, some of the more important properties and relationships with respect to GPS are pointed out.

7.4.4.1 Ionospheric Effects on GPS Signals

The propagation delay in the ionosphere (between about 50 km and 1000 km above the Earth's surface) depends on the electron content along the signal path and on the frequency used. The influencing parameters are mainly solar activity and the geomagnetic field. Hence, ionospheric refraction varies with frequency, geographic location, and time. The resulting range error, for GPS frequencies, can vary from less than 1 m to more than 100 m (Wells (ed.), 1986; Klobuchar, 1991, 1996). Dual frequency receivers make use of the fact that the L1 and L2 signals experience different propagation delays in the ionosphere. In addition, we have to note that the ionosphere is a dispersive medium [2.3.1.2], and that therefore the phase velocity (propagation of the carrier) is not the same as the group velocity (propagation of the codes).

To be exact, we observe the combined effect from the ionosphere and the *plasmasphere*, because the GPS orbits are located far above the ionospheric layers. The electron content below about 2000 km is also called the *Faraday content*. For a detailed study of the time-variable ionospheric behavior, e.g. in atmospheric physics for *ionospheric tomography*, it is hence advisable to combine measurements from Low Earth Orbiters (LEO) and GPS satellites, or to install GPS receivers in satellites at low orbital height. An observer at the surface of the Earth, who uses GPS as a tool for positioning or navigation, has no need to separate the ionospheric and the plasmaspheric propagation delay. In this book, as in most literature, the term *ionospheric delay* is therefore understood as the combined effect. For more information on the physics of the ionosphere with particular reference to the propagation of radio waves see e.g. Davies (1990) or Klobuchar (1996).

For carrier phase measurements we have the refraction coefficient from equation (2.95):

$$n_p = 1 - 40.3 \frac{n_e}{f^2},$$

with

n_e electron content along the signal propagation path, and

f carrier frequency.

The ionospheric effect on code propagation (group delay) is, at first order, of the same size as the carrier phase propagation but has the opposite sign (2.99):

$$n_g = 1 + 40.3 \frac{n_e}{f^2}.$$

Integration over the entire propagation path, s, then yields the total effect of ionospheric refraction on the pseudorange measurement, R, with code phases:

$$\delta R_{\mathrm{ION}_g} = \int_s (n_g - 1)\, ds \tag{7.135}$$

$$\delta R_{\mathrm{ION}_g} \approx \frac{40.3}{f^2} \int_s n_e\, ds. \tag{7.136}$$

The corresponding expression for carrier phase measurements is

$$\delta R_{\mathrm{ION}_p} \approx -\frac{40.3}{f^2} \int_s n_e\, ds. \tag{7.137}$$

Hence, the range from a code phase observation is measured as too long, and a range from a carrier phase observation is measured as too short.

The unknown integral can be determined by measurements of the ranges $R_1 = R\,(\mathrm{L1})$ and $R_2 = R\,(\mathrm{L2})$ on both frequencies:

$$R = R_1 - \delta R_{1,\mathrm{ION}}, \quad R = R_2 - \delta R_{2,\mathrm{ION}}. \tag{7.138}$$

By substitution of (7.136) into (7.138), omitting subscript g or p for simplicity, it follows that the expression of range correction for code phase measurements on L1, derived from dual frequency observations is

$$\delta R_{1,\mathrm{ION}} = \frac{R_1 - R_2}{1 - \left(\frac{f_1^2}{f_2^2}\right)} = \frac{f_2^2}{f_1^2 - f_2^2}(R_2 - R_1). \tag{7.139}$$

Because of the approximation in (2.99), this equation (7.139) is called the *first-order ionospheric refraction correction*. The remaining model errors reach only a few centimeters (see Table 7.15). Therefore the ionospheric effect in GPS can be very largely modeled by dual frequency observations. Table 7.15 shows the maximum range errors that can be expected for both GPS frequencies, and for the dual frequency corrected signal, both in vertical direction. For inclined directions the influence increases with the appropriate mapping function (cf. [2.3.2]). For more detailed information see e.g. Wanninger (1994); Langley (2000a); Misra, Enge (2001).

Table 7.15. Maximum vertical ionospheric range error [m] (Wübbena, 1991)

Frequency	1st order effect $(1/f^2)$	2nd order effect $(1/f^3)$	3rd order effect $(1/f^4)$
L1	32.5	0.036	0.002
L2	53.5	0.076	0.007
L1/L2	0.0	0.026	0.006

Corresponding equations can be derived for carrier phase observations (e.g. Wells (ed.), 1986; Misra, Enge, 2001):

$$\delta\Phi_{\mathrm{ION}}(\mathrm{L1}) = \frac{f_2^2}{f_2^2 - f_1^2}\Big(\Phi(\mathrm{L1}) - N(\mathrm{L1}) - \frac{f_1}{f_2}(\Phi(\mathrm{L2}) - N(\mathrm{L2}))\Big), \tag{7.140}$$

where $N(\mathrm{L1})$ and $N(\mathrm{L2})$ are the respective ambiguity terms. Equation (7.140) describes the *ionospheric phase advance* for L1 observations. Combining observations on L1 and L2 yields the *ionospheric free linear combination* (e.g. Leick, 1995; Hofmann-Wellenhof et al., 2001):

$$\Phi(\mathrm{L_0}) = \frac{f_1^2}{f_1^2 - f_2^2}\Phi(\mathrm{L1}) - \frac{f_1 f_2}{f_1^2 - f_2^2}\Phi(\mathrm{L2}). \tag{7.141}$$

Some authors call the ionospheric free combination the L_3 observable. The corresponding equation for ionospheric-free code range observations can be reached from (7.139) as (e.g. Misra, Enge, 2001):

$$R_0 = \frac{f_1^2}{(f_1^2 - f_2^2)}\,R_1 - \frac{f_2^2}{(f_1^2 - f_2^2)}\,R_2 = 2.546\,R_1 - 1.546\,R_2. \tag{7.142}$$

If only single-frequency receivers are available, a correction according to equations (7.139) or (7.140) is impossible. An attempt can then be made to use an ionospheric correction model, the coefficients of which are transmitted as part of the GPS satellite message (Klobuchar, 1987, 1996), [7.1.5.4]. This *Klobuchar model* is described by eight coefficients α_n, β_n, and removes about 50% of the ionospheric delay at mid-latitudes. The correction formula is

$$\begin{aligned} \Delta T_{\mathrm{ION}} &= DC + A\cos(2\pi(t - \Phi)/P) \quad [\mathrm{day}] \\ \Delta T_{\mathrm{ION}} &= DC \quad [\mathrm{night}], \end{aligned} \tag{7.143}$$

in which

ΔT_{ION}	vertical delay (ns),
DC	constant night-day offset (5 ns),
A	amplitude of the cosine function for daytime values,
Φ	constant phase offset corresponding to the peak of the cosine function, fixed at 54 000 s or 14 h local time,
t	local time,
P	period of the cosine function, and furthermore

$$A = \sum_{n=0}^{3} \alpha_n \Phi^n \,(\text{seconds}); \qquad P = \sum_{n=0}^{3} \beta_n \Phi^n \,(\text{seconds}).$$

The vertical ionospheric delay has to be scaled into the slant delay with the slant factor, F:

$$F = 1.0 + 16.0 \times (0.53 - E)^3. \tag{7.144}$$

E is the satellite elevation angle. Details of the calculations are given in Klobuchar (1996); Misra, Enge (2001), and in the GPS Interface Control Document (ICD, 1993).

The Klobuchar model is based on empirical data, but has severe limitations because the number of parameters is restricted to eight, and because it can only be updated once daily. Due to the fast changing ionospheric environment the remaining error in zenith delay is estimated to be about 10 m during the day at mid-latitudes, and much worse when the solar activity is high (Misra, Enge, 2001).

Alternative approaches are to model the ionosphere with the help of LEO observations, as with TRANSIT [6.2] or with PRARE on ERS-2 [4.3.3.3] (Flechtner, 2000), or with reference observations from one or more dual-frequency GPS receivers located in the working area. Both procedures were initiated early on, e.g. Lohmar (1985) and Georgiadou, Kleusberg (1988), but have severe limitations. With LEO satellites, only the Faraday content below 2000 km can be determined but not the electron content of the plasmasphere above 2000 km altitude (Davies, 1990). The contribution of the plasmasphere reaches, however, 10% to 50% of the total electron content. TEC measurements with GPS, hence, cannot directly be compared with TEC results from other radio systems. On the other hand, GPS occultation measurements with a GPS receiver on a satellite in low orbit (LEO) help to map plasma irregularities in the lower ionosphere (Hocke, Tsuda, 2001), see [7.6.2.9].

In the second technique, the vertical propagation delay is observed with dual-frequency equipment, and is introduced into a polynomial model describing the local ionosphere for the correction of single frequency observations. With one reference receiver located at the center of the working area, the method works well under homogeneous and moderate ionospheric conditions. However, it is not very effective in regions and times with strong *ionospheric disturbances* and/or *very high electron content*. The technique has been further developed using three and more dual frequency receivers (e.g. Webster, Kleusberg, 1992), and it is now well established in active multiple-reference-station networks. The actual behavior of the ionosphere can be measured and modeled in real-time for the complete working area. This concept

of "wide area augmentation" or "area correction parameters" or "virtual reference stations" is treated at length in [7.5.3].

At a global scale, ionospheric TEC models are derived from data of the International GPS Service, IGS (see [7.8.1]). Five so-called "Ionospheric Analysis Centers" (IAACs) deliver every 24 hours an "Ionospheric Map Exchange" (IONEX) file (Schaer et al., 1998) with 12 maps containing global TEC information with 2-hour time resolution. For the northern hemisphere, under normal conditions, the different TEC maps agree with the IGS mean by about five TEC units or less. At the equator and for southern latitudes, the situation is still more problematic because of poor station coverage. However, the use of regional networks for monitoring regional TEC behavior is being investigated, e.g. for South America (Fedrizzi et al., 2001).

The IGS is preparing for the establishment of an independent IGS ionospheric model and a near-real-time service (IGS, 2002a).

Residual errors in the ionospheric modeling are cancelled out, for the most part, through relative observations at two stations over short distances, since the satellites are observed through nearly the same atmosphere. The remaining error for single frequency equipment is estimated to be 1 to 2 ppm of the interstation distance, corresponding to 1 to 2 cm over 10 km (Campbell et al., 1984). These numbers are valid for a quiet ionosphere, and for observations in mid-latitudes only. The last periods of high solar activity have demonstrated that the residual error can be significantly larger. It is hence advisable to use only dual frequency equipment for high precision application.

Irregularities in Earth's ionosphere can produce short-term signal variations in amplitude and phase (e.g. Wanninger, 1992, 1994; Langley, 2000a). These *scintillation effects* mainly occur in a belt of $\pm$ 30 degrees either side of Earth's geomagnetic equator, and in the polar auroral zones (see Fig. 7.52). A very high electron content only occurs in equatorial regions.

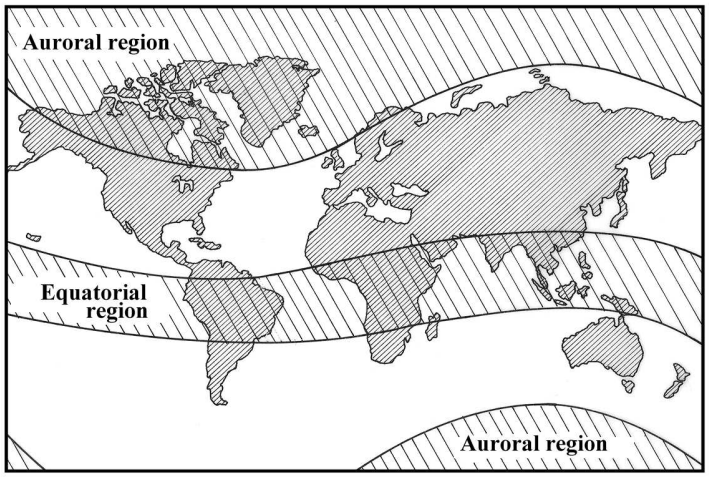

Figure 7.52. Regions of the world with high ionospheric activity

Equatorial scintillation effects have their maximum typically from approximately one hour after local sunset until approximately midnight (Klobuchar, 1991). Scintillation effects are less significant from April through August in the American, African, and Indian longitude regions, but maximize in the Pacific region. From September through March the situation is reversed.

Scintillation effects may cause a large number of cycle slips because the receiver cannot follow the short-term signal variations and fading periods. A very high electron content produces strong horizontal gradients and corrupts the ambiguity solution with the geometrical method, even over short distances, because the *ionospheric signal* (cf. [7.3.2.1]) overlaps even the wide lane wavelength within a few minutes of observation time (Wanninger, Jahn, 1991; Wanninger, 1994). In such situations, the only reliable possibility of ambiguity resolution so far found is the code-carrier combination [7.3.2.3] using data from low-code-noise receivers. Ionospheric effects are visible in double difference phase observations even over short distances. Relative errors up to 30 ppm have been observed in single frequency baseline determination over 10 km in Brazil (Campos et al., 1989).

At times, ionospheric perturbations also occur in mid-latitude regions (Wanninger, 1992, 1994). In particular, so-called *Medium Scale Travelling Ionospheric Disturbances* (MSTIDs) may generate serious problems for precise relative positioning in mean latitudes when the observation time is short (< 20 minutes).

In the short term, the situation will improve because the current sunspot cycle is now in the declining phase, with a minimum expected in 2006. As additional GPS frequencies become available, as part of the "GPS Modernization Program", multi-frequency receivers will enter the market, so that the local ionospheric delay can be directly measured and eliminated.

7.4.4.2 Tropospheric Propagation Effects

The tropospheric propagation delay is critical for precise position and baseline determination, in particular in the height component, because the tropospheric parameters are only poorly correlated over larger distances. Furthermore, it is difficult to separate error components stemming from the radial orbital errors, signal propagation errors, clock errors, antenna phase center variation, and errors in the station height. This is one of the reasons why the height component is much worse than the horizontal components in precise GPS positioning.

For frequencies in the radio spectrum [2.3] the tropospheric delay is independent of the frequency; hence it cannot be determined from dual-band measurements. The near-surface atmospheric structure has to be adequately modeled. Either mean atmospheric parameters, or measured data on temperature, atmospheric pressure, and water vapor content along the signal propagation path must be included in the model. Some of the currently accepted models are dealt with at length in [2.3.3.2]. For further information see also (Mendes, Langley, 1994, 1999; Spilker, 1996d; Langley, 1998b).

Usually, the influence of the neutral atmosphere on range measurements to satellites in the radio frequency domain is expressed by two integral terms: the dry component

and the wet component. The wet component depends on the distribution of water vapor in the atmosphere and is therefore harder to model. The wet portion, however, comprises only 10% of total tropospheric refraction. The total delay in the zenith direction comes to about 2.3 m, and increases near the horizon (10° elevation) to about 20 m. The dry component is precisely described (with an accuracy of $\pm 1\%$) by the available models. The wet delay can be modeled, depending on the atmospheric conditions, with an accuracy no better than 1 to 2 cm (Langley, 1998b).

Most studies conclude that none of the available models has a clear priority over the others. For low elevation angles, the *Niell model* (Niell, 1996) is usually preferred (e.g. Hay, 2000). This is of particular relevance, because the observation of low satellites (down to 5° elevation) is essential in precise GPS height determinations (Dach, 1999), [7.6.2.3]. The Niell mapping function is of the Marini type (2.116) and uses coefficients depending on latitude and season. For details see also Schüler (2001, p. 157ff).

If the stations are close together, the tropospheric residual error almost completely disappears by differencing in the relative observation mode. It is hence not advisable to introduce the observed meteorological data separately for each station into the adjustment of a small network in non-mountainous regions. The local measurements usually do not represent the regional atmospheric situation with sufficient rigor, and hence introduce biases into the solution. Instead, appropriate identical standard atmospheric parameters should be used for all stations. In this respect, it is of interest that the Niell dry and wet functions are completely independent of surface meteorological measurements.

When station distances are greater (say > 50 km), or when the height differences are larger (in mountainous regions), atmospheric conditions are no longer sufficiently correlated with one another. Adequate modeling, hence, is of growing importance, in particular for precise DGPS or WADGPS applications (e.g. Collins, Langley, 1999) [7.5].

One way of determining the water vapor content of the atmosphere along the propagation path is direct measurement with *water vapor radiometers* (e.g. Nothnagel, 2000). The instruments are, however, very elaborate and expensive and can only be used for major tasks (cf. [2.3.3.2]).

Another approach is to introduce a station dependent *zenith scale factor* for each satellite pass. This parameter can only be estimated reliably after an observation time of 1.5 to 2 hours. To allow for the time variable behavior of the tropospheric zenith delay, stochastic modeling has been successfully applied (e.g. Völksen, 2000). Another option is elevation-dependent weighting (Rothacher et al., 1998). In global networks the introduction of a scale factor is self-calibrating because mis-modeling in the atmospheric delay would produce a scale-factor and hence result in mis-modeling of the orbits and in a violation of orbital mechanics (Nothnagel, 2000).

A very successful approach is real-time monitoring of the tropospheric effects in active multiple reference station networks and the immediate correction of user positions [7.5.3]. Along with the Accuracy Improvement Initiative [7.1.7] enhanced orbital data can be expected when new tropospheric mapping functions will be applied

in the Master Control Station (Hay, Wong, 2000).

Tropospheric modeling remains one of the most demanding tasks in the precise use of GPS. A wealth of information already exists, but research in this respect certainly will continue. The availability of near real-time global and regional high-resolution tropospheric models coming from ground-based (e.g. Schüler, 2001) or space-based (e.g. Reigber et al., 2002) GPS observations considerably contributes to improved data correction. The IGS provides a tropospheric product in the form of combined zenith path delay estimates for more than 210 sites at the level of 3 to 5 millimeters, which corresponds to ∼1 mm in water vapor (Gendt, 2000), see also [7.6.2.9], [7.8.1].

7.4.4.3 Multipath

Multipath propagation means that one or more reflected signals reach the antenna, in addition to the direct signal. Under particular circumstances only the reflected signal may reach the antenna.

There can be reflections off horizontal, vertical, and inclined surfaces (Fig. 7.53), possible examples being streets, buildings, waterways, and vehicles. This should be considered when selecting observation sites, in particular for permanent reference stations.

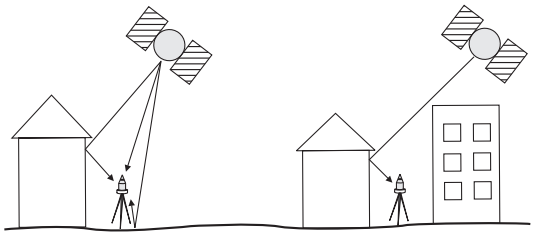

Figure 7.53. Multipath propagation

Multipath propagation affects both code and carrier measurements. The effect on *P-code observations* is two orders of magnitude larger than on carrier phase observations, and can reach decimeters to meters. The effect on *C/A-code observations* is at the order of several meters, and can even reach, in extreme situations, 100 m or more (Braasch, 1996). Under the worst conditions the code signal multipath may cause the receiver to lose phase lock. Many cycle slips are produced by multipath effects.

The code signal multipath becomes particularly critical when the code/carrier combination technique is used for ambiguity resolution, e.g. the extra wide laning technique for "ambiguity solution on the fly" (cf. [7.3.2.3]) in a surveying airplane. Reduction of code multipath is essential for the precise code-only Differential GPS [7.5.1], for example when sub-meter accuracy is required in GIS applications.

Multipath influences on *carrier phase observations* produce a phase shift that introduces a significant periodic bias of several centimeters into the range observation. The direct and the reflected signals are, in a simplified presentation:

$$A_D = A\cos\Phi_D, \quad A_R = \alpha A\cos(\Phi_D + \Phi), \tag{7.145}$$

where

A_D amplitude of the direct signal,

A_R amplitude of the reflected signal,

α damping factor ($0 \leq \alpha \leq 1$),
0: no reflection,
1: reflected signal as strong as direct signal,
Φ_D phase position of the direct signal, and
Φ phase shift of the reflected signal with respect to the phase of the direct signal.

The superposition of both signals gives the expression

$$A_\Sigma = A_D + A_R = A\cos\Phi_D + \alpha A\cos(\Phi_D + \Phi) = \beta A\cos(\Phi_D + \Theta). \quad (7.146)$$

With $A_{D,\max} = A$, and $A_{R,\max} = \alpha A$, it follows that the equation for the resultant multipath error, Θ, in the observed carrier phase is

$$\Theta = \arctan\left(\frac{\sin\Phi}{\alpha^{-1} + \cos\Phi}\right). \quad (7.147)$$

The signal amplitude is expressed as

$$B = \beta A = A\sqrt{1 + \alpha^2 + 2\alpha\cos\Phi}. \quad (7.148)$$

Inspection of the above equations demonstrates that, for $\alpha = 1$, the maximum value of Θ is

$$\Theta_{\max} = 90^\circ. \quad (7.149)$$

Hence, the maximum error in the L1 signal ($\lambda = 19.05$ cm) is about 5 cm. For linear combinations of L1 and L2 the values can be correspondingly larger or smaller. Their propagation into height errors may reach ± 15 cm (Georgiadou, Kleusberg, 1990). Due to the changing satellite geometry, the multipath effect in carrier phases shows a cyclic behavior. Typical periods are between 15 and 30 minutes, depending also on local reflectors. Fig. 7.54 gives an example for the double difference phase observable.

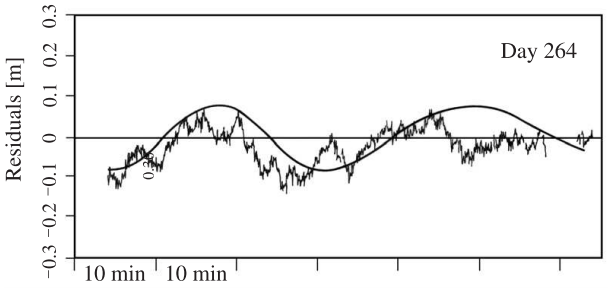

Figure 7.54. Observed and calculated (solid line) multipath effect

The multipath effect on position results can be minimized with observations over a larger time period, at least over one of the effective cycles. This is not possible in kinematic or rapid static surveying. It is hence important to avoid, or at least to mitigate, multipath propagation, in particular for reference stations, because multipath biases propagate into all rover positions. Possible measures to minimize the effects are

(a) observation design,
(b) receiver and software design, and
(c) station calibration.

(a) *Observation design*
Several measures and actions are possible:

- select sites carefully, avoid nearby reflectors,
- use antenna ground plane to avoid reflections from the ground,
- deploy absorbing material on the ground,
- select carefully designed antennas, e.g. choke ring antennas, and
- use multiple antenna arrays or controlled antenna motion to average out the multipath variation near the antenna.

One particular procedure is to observe sidereal differences. Since the satellite geometry repeats after 24 hours of sidereal time ($23^{h}\ 56^{min}$ UT), the multipath effect for a given site also repeats. By forming *sidereal differences* between the observables, for example double differences DD, at two consecutive days it is hence possible to generate *multipath free observables* DD_{Sid}. These can, for example, be used for the absolute calibration of antennas [7.4.5.1] and also for highly precise control measurements (Seeber et al., 1997a).

(b) *Receiver and software design*
A number of methods for reducing multipath effects use real-time signal processing in the receiver. For an overview see e.g. Weill (1997). The basic idea is to use particular signal properties for improving the correlation process. GPS signals are left-hand polarized. Reflected signals change their polarization, and hence can be detected in the receiver. Reflected signals also arrive later at the antenna and hence can be discriminated. The various techniques have been given names like "narrow-correlator", "strobe correlator", "correlation function", "Everest technology", and so on. Usually, the details are not revealed by the manufacturers; some basic concepts, however, are published, mostly in the ION-GPS Proceedings.

(c) *Station calibration*
A rather new idea is to calibrate stations, in particular reference stations, for multipath effects. A first step is the detection of multipath. Several techniques are possible. In *double differences* over short baselines, most errors are eliminated. The remaining residuals mainly contain the multipath differences. However, it is not possible to separate between the participating stations. Multipath effects are also visible in the signal-to-noise ratio (SNR). The signal strength varies in a sinusoidal form depending on the multipath. Finally, the inspection of sidereal differences helps to analyze the variation of multipath effects.

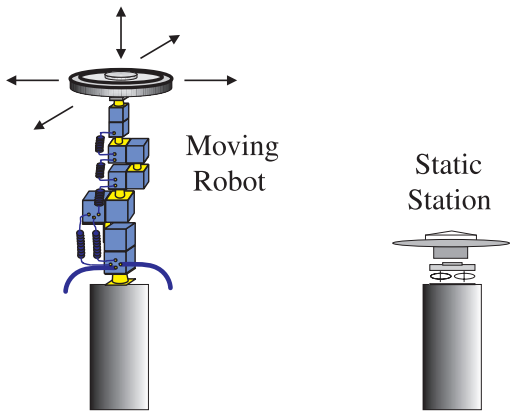

Figure 7.55. Calibration of a reference station for multipath

A method for the absolute field calibration of multipath has recently been

reported (Böder et al., 2001). One prerequisite is the availability of absolute calibrated antennas, because otherwise the multipath cannot be separated from the antenna phase center variations [7.4.5.1]. The basic idea of this method is to decorrelate the multipath through controlled motion of a robot (Fig. 7.55). The robot operates near the station to be calibrated. The fixed station senses the complete multipath. The moving station eliminates the multipath through the controlled motion. In the double differences only the multipath effects for the fixed station are present and can be described in a functional model, e.g. with spherical harmonics.

Multipath effects at satellites have been reported but seem to be less critical (Young et al., 1985).

7.4.4.4 Further Propagation Effects, Diffraction and Signal Interference

Two main influences have to be considered, signal diffraction and signal interference. *GPS signal diffraction* comes about when the direct GPS signal is obstructed but a diffracted signal is received. Fig. 7.56 explains the geometric situation. Following Walker, Kubik (1996) we distinguish the regions A, B, C, and D. In A, B, C we have direct reception of the GPS signals. In addition we may expect in

- A: reflected signals from the ground in front of the obstacle and from the obstacle,
- B: only little or no reflection, and
- C: reflection from the ground behind the obstacle.

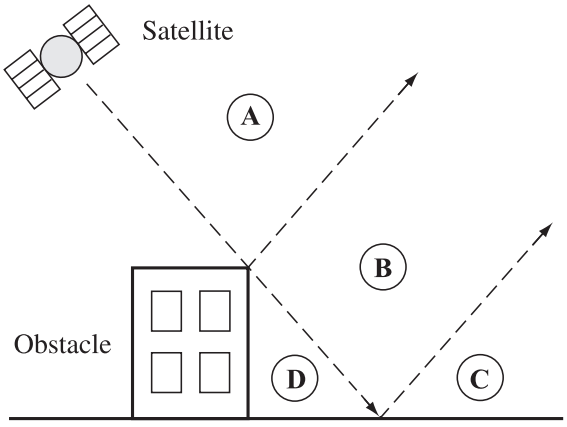

Figure 7.56. Signal propagation near an obstacle; after Walker, Kubik (1996)

In region D, from the laws of geometrical optics, we have no signal reception except for signals diffracted at the obstacle. The increased signal path of the diffracted signal may produce a phase error of up to several centimeters or even decimeters. The effect can hence be considered as one of the dominant error sources in rapid static or kinematic GPS positioning (Wanninger, 2000).

A powerful means for the detection of diffracted signals is inspection of the signal-to-noise ratio (SNR). A proper weighting of the undifferenced phase observables, based on the SNR values, can be used for minimizing the diffraction effect on coordinate estimates.

Interferences with artificial signals from HF-transmitters occur for frequencies in or near the bandwidths of the GPS signals. The reason is that GPS signals are not transmitted at a discrete frequency but, due to the code modulation, they are spread over a certain bandwidth, namely 2.046 MHz for the C/A-code and 20.46 MHz for the P-code (spread spectrum technique, see [7.1.4]). The effect of disturbances from signals at nearby frequencies can be minimized by adequate filter technology, however

it cannot completely be avoided. The effects on GPS signals, for strong disturbances, are

- decreased SNR level,
- more difficult or impossible acquisition of the GPS signal, and
- loss of signal in phase tracking loop.

Interferences mainly occur for L2. L1-only receivers are usually not affected. Modern receivers, with enhanced filter technology, are better protected than older receiver types. Possible sources for signal interferences are

- VHF, UHF, TV transmitters with strong radiation power within a distance of 100 to 500 m,
- digipeater directional transmission of amateur radio (in Germany just 3 MHz off L2), and
- radar installations of aviation control services, up to 20 km distance.

The influence of high voltage power cables is small. GPS receiver antennas should stay distant by about 10 m, as from all other transmitters, to avoid direct disturbances. Cellular phones seem to have little influence. Possible interference with the forthcoming *ultra-wide-band* technology is under discussion (Akos et al., 2001). For more details on signal interference with GPS see Johannessen (1997), an excellent review in German is given by Kolb (1999).

A particular effect is *foliage attenuation* for stationary and mobile users. Depending on the type of trees and the length of foliage penetration, the attenuation can vary significantly. A detailed treatment of the subject is given by Spilker (1996a).

7.4.5 Receiving System

The main error sources in the receiving system are

- antenna phase center variations,
- receiver noise,
- interchannel bias, and
- oscillator instability.

7.4.5.1 Antenna Phase Center Variation

Positioning in navigation and geodesy refers to the *electrical phase center* of the antenna, that varies with the intensity and direction of the incident signals. For precise applications, the phase center positions of all antennas involved in a project have to be known exactly. This is of particular importance for determination of the height component because in the GPS adjustment the elevation-dependent effects are highly correlated with the height and the tropospheric scale parameter.

The mechanical center of an antenna is usually defined to submillimeter precision. It often coincides with the intersection of the vertical mechanical axis of symmetry and the ground plane (Fig. 7.57). The mean electrical phase centers for the L1 and L2 signals may be a few mm off from the mechanical center. The *antenna reference point* (ARP) is also defined mechanically, usually as the intersection of the vertical

mechanical axis with the lowest part of the antenna housing. For most antenna types, the 3D-coordinates of the offsets of the L1 and L2 mean electrical phase centers with respect to the ARP are given by the manufacturers. The actual electrical phase center depends on the azimuth and elevation of the observed satellites. The deviations of the actual phase centers from the mean electrical phase center are the *phase center variations* (PCV). They can reach millimeters to a few centimeters.

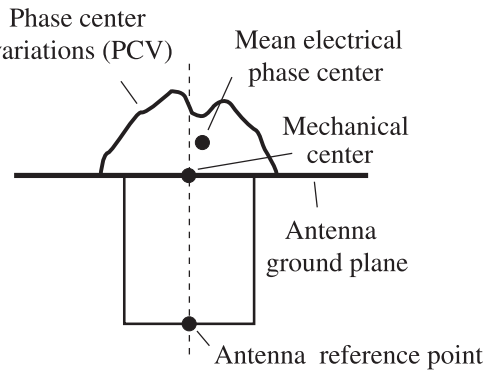

Figure 7.57. Antenna phase center variation and reference points

If antennas of the same type are used within one observation session over short baselines, the remaining phase center offsets and variations are eliminated in the differencing process. In cases where the phase center variation is azimuth dependent, all antennas have to be orientated prior to the survey. For this reason, some antenna types have an orientation mark directed to magnetic north.

If different antenna types are involved within the same project, as is often the case for precise DGPS with reference stations [7.5.1], the observations have to be corrected for the PCV. The same is true when identical antennas are used with very large baselines, because the satellite signals are observed under different elevation angles due to Earth's curvature. Note that different antennas of the same type may also show differences in PCV. For highest accuracy requirements only calibrated antennas should be used. Three major GPS antenna calibration methods are presently available (Rothacher, 2000a):

- anechoic chamber calibrations,
- relative field calibrations, and
- absolute field calibrations.

The *anechoic chamber calibration* (Schupler, 1994; Schupler, Clark, 2001) is a laboratory method, and it is rather seldom applied because not many anechoic chambers exist. A GPS antenna is tilted and dislocated with respect to an artificial GPS signal, generated in the chamber. Absolute PCV are determined under the assumption that they are also valid for observations in the field.

In *relative field calibration* the PCV and the mean offset of a specimen antenna is determined with respect to another antenna, the reference antenna. The PCV of the reference antenna (often the *Dorne Margolin T* choke ring antenna) are set to zero or taken as known. Both antennas are mounted close together on pillars with very precisely known coordinates. The calibration is based on single or double difference residuals. The elevation- and sometimes azimuth-dependent PCV model uses polynomials or spherical harmonics. The method has been widely used to calibrate all major GPS antenna types (e.g. Rothacher et al., 1995; Mader, 1999).

Methods of *absolute field calibration* have only been developed recently (Wübbena et al., 1997; Menge, Seeber, 2000; Wübbena et al., 2000). The basic idea is to elim-

inate multipath effects by either using sidereal differences between observations on two consecutive days, or to use a high-precision robot (Fig. 7.58) that rotates and tilts, at rather high speed, the antenna to be calibrated. Observations from a nearby stationary reference antenna are required to eliminate distance dependent errors; the results are, however, absolute PCVs independent on the type of the reference antenna. Fig. 7.59 shows one example. For details of the method see the cited literature.

Figure 7.58. Robot for the absolute field calibration of GPS antennas

The advantages of absolute PCV values, from field calibrations, when compared with the traditional techniques are, among others

- they are available in real-time (robot technique) for L1, L2, GPS, GLONASS, future GNSS,
- they are independent of a reference antenna and reference coordinates,
- they are free of multipath,
- they cover the whole hemisphere and are independent of the "northern hole" [7.6.1.1],
- they support the absolute calibration of GPS reference stations, and
- they facilitate the separation from other error sources like troposphere, and satellite antenna phase center offset.

Absolute PCV of modern antennas are mostly below 10 mm, but they can also reach much higher values. The influence on height determination can be several centimeters. It is hence advisable to only use absolutely-calibrated antennas for active reference stations, and for all tasks where high accuracy is required.

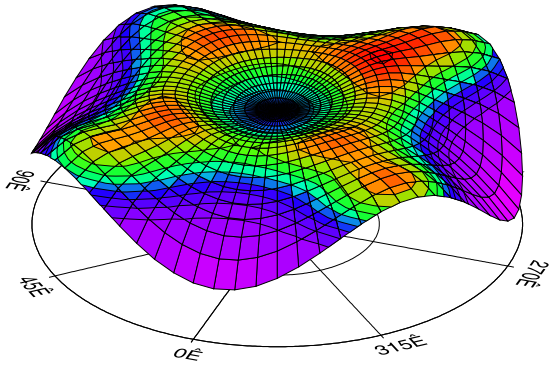

Figure 7.59. Absolute PCV of a GPS antenna

In order to facilitate the use of reference data, a "zero-antenna" can be introduced, i.e. an antenna where all observations are corrected for the PCV. The rover then only has to apply its own PCV corrections. A particular RINEX format, *ANTEX*, for the distribution of antenna PCV information has been developed (IGS, 2002b).

At some sites, it is advisable to protect the antenna set-up with a radome against hostile environmental influences. Such radomes may change the antenna PCV characteristics; hence the antennas have to be calibrated together with the radome (Kaniuth, Stuber, 1999; Schupler, Clark, 2001).

Note that the azimuth dependent variation of PCV can also be used vice-versa to determine the orientation of the antenna and of the related platform (Tetewsky, Mullen, 1997). The resolution, however, is only in the order of several degrees.

With an absolutely calibrated reference antenna at hand, it is also possible to calibrate satellite antennas. A first result has been reported for the Block IIa satellites (Mader, Czopek, 2002).

7.4.5.2 Other Error Sources Related to the Receiving System

The *receiver noise* results from the fact, that GPS phase and code observables cannot be measured perfectly but are subject to random influences. For example, the observations are affected by unwanted disturbances in the antenna, amplifiers, cables, and the receiver itself. For details see e.g. Langley (1997b).

As a rule of thumb the observation resolution for classical receivers is about 1% of the signal wavelength. For the GPS signals we obtain:

C/A-code	$\lambda \approx 300$ m,	noise ≈ 3 m,
P-code	$\lambda \approx 30$ m,	noise ≈ 30 cm, and
carrier	$\lambda \approx 20$ cm,	noise ≈ 2 mm.

Modern receiver technology tends to bring the internal phase noise below 1 mm, and to reduce the code-resolution to the 10 centimeter level [7.2.4.2]. Low noise code measurements are important for real-time ambiguity resolution.

Multichannel receivers exhibit different *signal propagation delays* for each hardware channel, since each satellite signal travels along a different electronic path. The instrument makers try to calibrate and to compensate these *interchannel biases*. Multiplexing and software receivers [7.2.1], [7.2.5] are free of interchannel biases. It is recommended that parameters for satellite and receiver hardware delays are included in the parameter estimation models, in particular if the concept of original undifferenced phase data is used (cf. [7.3.2.2]).

Oscillator instabilities play only a minor role in carefully designed receivers because the timing signal is taken from the satellite clock. They can be modeled in the adjustment process. For highest accuracy requirements, and in precise navigation, the use of external precision oscillators (rubidium or cesium) is recommended.

Further error sources that can be counted to the receiving system, are the stability of the station ground and of the pillars, as well as the quality of the station mark. These items are of particular interest in geodynamic networks.

7.4.6 Further Influences, Summary, the Issue of Integrity

Several more aspects exist that influence the achievable accuracy. Among them are the process noise and the tidal upload. *Process noise* means that some liberty exists in

the data analysis approach. We can identify

- *software noise*, the agency is free in the selection of a particular software package,
- *operator noise*, the operator is free in the selection of particular options of the software package, and
- *reference frame realization noise*, there exist various possibilities to select a set of fiducial stations, e.g. from the IGS, to connect a project with a given reference frame.

As a consequence a given data set will lead to slightly different results, when different operators work with different software packages. This is also true for the same package used in different laboratories. An impressive example based on the analysis of a 40-days-long data set from about 50 stations, analyzed with 4 different software-packages at 7 laboratories is given by Dietrich et al. (2001). The mean differences between solutions are 1 cm in horizontal position and 2 cm in height.

Tidal upload means that GPS stations show a vertical displacement due to the crustal deformation caused by oceanic and solid Earth tides (Dach, 1999). The effect can reach several centimeters but it is the same over large areas and hence will be cancelled by relative GPS. Considering today's high accuracy potential of GPS observations, corrections for tidal upload should be applied whenever highest accuracy of the results is attempted. For details see e.g. Dach (1999); Zahran (2000).

In summary, the accuracy achievable with GPS for geodesy, surveying, and navigation, depends on various conditions, for example

- single or multi-receiver operation,
- single or dual-frequency data,
- L2 high quality access under AS available or not,
- receiver noise level,
- static or kinematic positioning,
- real-time or post-processing results,
- accuracy of orbits used, and
- extent of data modeling.

Because of the many options and influences, and the eminent progress in error modeling during the last years, it is not possible and not meaningful to describe the accuracy potential of GPS with a single distance dependent formula, as has frequently been done in the past (e.g. Lichten, 1990). The statement of today is that 1 cm accuracy, at a global scale over all distances, is achievable with appropriate instruments, observation design, and data analysis models. For selected examples, see the section on applications [7.6.2].

A major issue when using GPS in navigation is the *integrity* of the system. For navigational purposes integrity is defined as the "ability of a system to provide timely warnings to users when the system should not be used" (Brown, 1990). The timely warning is, in particular, required for the navigation of civil aircraft. The GPS control segment does not provide sufficient warning when a component of the system fails. Different solutions to the problem have been discussed. With *internal methods of integrity monitoring*, GPS integrity is achieved using information available inside the receiver, such as redundant measurements to additional satellites. This technique is

known as *receiver autonomous integrity monitoring* (RAIM). Using *external methods of integrity monitoring*, the GPS signals are controlled in real-time through a network of ground monitoring stations.

The information is broadcast to users through a *GPS integrity channel* (GIC) via geostationary satellites such as INMARSAT. A further approach to assuring the integrity of the GPS navigation solution is possible by integrating GPS data with data from other sensors. These can be, for example, inertial navigation systems, Loran-C receivers, GLONASS and future GALILEO receivers. For more information see the discussion in the navigation literature, e.g. the journal *Navigation*, and also [7.7.2]. A good introduction to the topic of integrity is Langley (1999b).

7.5 Differential GPS and Permanent Reference Networks

The absolute position determination with GPS is, in general, much less accurate than relative positioning between two stations. This is due to the fact that most of the acting errors (biases) are highly correlated. Error sources can be grouped into three categories [7.4]:

(1) errors decorrelated with distance,
(2) errors decorrelated with time, and
(3) uncorrelated errors.

Errors of type 1, mainly ephemeris and propagation errors, are nearly the same for neighboring stations, as long as they are sufficiently close, and hence disappear in the differences. Errors of type 2 are coped with by synchronized or nearly simultaneous observations. Errors of type 3 affect both participating stations and need a calibration.

To minimize the effect of errors of type 1, instead of absolute coordinates, coordinate differences are determined with respect to a known reference station. Several concepts are in use; the basic strategies are

(a) use of the data of one or more reference stations for post-processing,
(b) use of corrections in position or range from code observations at the reference station in real-time,
(c) use of code-range and carrier phase data from the reference station in real-time, and
(d) use of reference data from a network of reference stations in real-time.

The option (a) is often referred to as *relative GPS*, whereas options (b) through (d) are called *Differential GPS* (DGPS) with different attributes. Option (b) is *ordinary DGPS*, in its proper sense, whereas option (c) is called *precise DGPS* (PDGPS) or also *Real-Time Kinematic* (RTK) GPS. Option (d) is known as the concept of *Multiple Reference Stations*, *Networked Reference Stations*, or also *Network RTK*. The wording, in general, is not uniform.

In this book the term "relative GPS" is used in a general sense, including all concepts where data from more than one station are processed simultaneously, either in post-processing or in real-time. The term "Differential GPS" means that processing of data from more than one station is performed in real-time, or near real-time. Usually,

the original measurements and/or correction data are transmitted in real-time from one or more reference stations to one or more user stations, also called *rovers*. Differential GPS is mostly applied in navigation (see Fig. 7.60). In the following the various concepts are discussed in more detail.

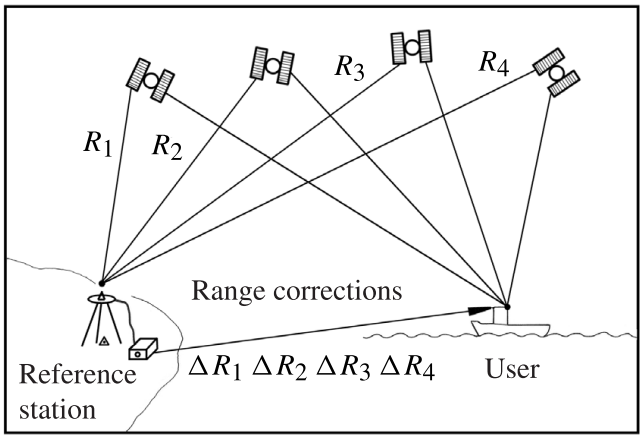

Figure 7.60. Differential GPS; range corrections are transmitted to the mobile user

7.5.1 Differential GPS (DGPS)

7.5.1.1 DGPS Concepts

Differential GPS (DGPS) is a technique that is used to improve the determined position of a roving station by applying corrections provided by a GPS monitoring station, also called *reference station*. Different procedures are in use for generating corrections:

(i) Corrections in the *position domain*

The GPS-derived position of the reference station is compared with its a priori known position. Position corrections Δx, Δy, Δz or $\Delta\varphi$, $\Delta\lambda$, Δh, are transmitted and used to correct the rover position.

(ii) Corrections in the *measurement domain*

The observed pseudoranges to all visible satellites are compared with ranges derived from known satellite and station positions. The differences are transmitted to the rover to correct its observed pseudoranges.

(iii) Corrections in the *state space domain*

Observations from several reference stations are used to estimate the state vector of the biases within the working area.

The first option (i) is rather simple, but also not very flexible. It only works if the same satellites are used at the reference and rover stations, and it is only efficient over short distances. It is therefore seldom applied.

Option (ii) is the ordinary DGPS procedure and is explained in more detail later on in this chapter. It is very flexible and works well within a radius of several hundred

kilometers about the reference station. Due to decorrelation of the biases with distance (orbit, ionospheric and tropospheric delay) the accuracy decreases roughly by about 1 m per 100 km.

Option (iii) is the most flexible procedure, and allows the use of DGPS over larger distances (Wide Area Differential GPS (WADGPS)) and for precise applications in surveying and geodesy (networked reference stations). It is explained in more detail in section [7.5.3]. Whereas option (ii) is based on observed pseudorange corrections (scalar corrections), option (iii) is based on correction vectors.

Several more classifications of DGPS are in use. Following the achievable accuracy we have:

Ordinary DGPS with code range corrections; accuracy 1 to 3 m, depending on the distance from the reference station.

Carrier smoothed DGPS; at the rover station, the carrier observations are used to smooth the coarse code observations with a suitable filter (cf. (7.108)) without solving for ambiguities. The achievable accuracy is < 0.5 m.

Precise DGPS (PDGPS); carrier phase observations, or carrier phase corrections, from the reference station are transmitted to the rover and are used to resolve ambiguities. This procedure is identical to the *Real Time Kinematic* (RTK), see [7.5.2]. Fig. 7.61 shows the accuracy potential of the different options.

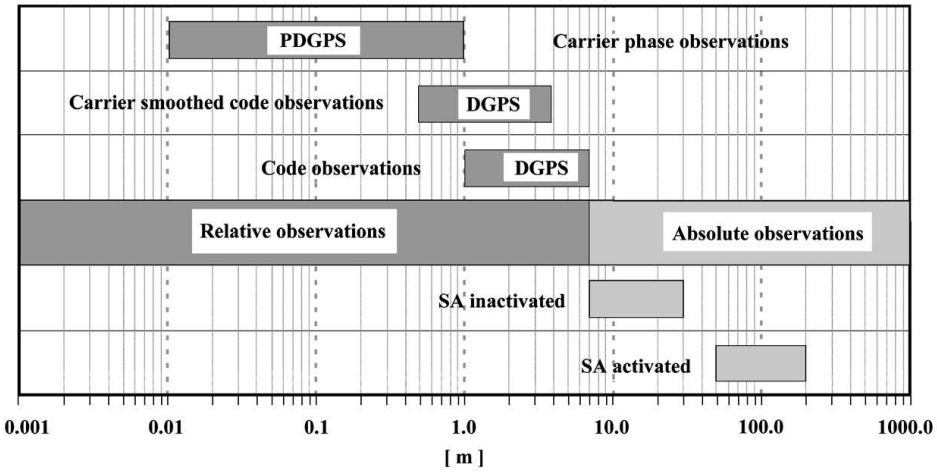

Figure 7.61. Accuracy potential with different modes of DGPS

Another classification is into

- Local Area DGPS (LADGPS),
- Wide Area DGPS (WADGPS), and
- Carrier Phase DGPS (CDGPS), or precise DGPS (PDGPS).

Local Area DGPS (LADGPS) corresponds to the procedures (a) and (b), above. In option (b) a scalar correction to code-phase measurements is applied for each satellite. The corrections are used for areas up to 1000 km radius. However, with the deactivation of SA single receiver accuracy meets the same level of accuracy as DGPS

for distances larger than several hundreds of kilometers. In LADGPS, corrections of all the different error influences are put together in one value. If more than one reference station provides corrections, they can be weighted to form a mean value.

Wide Area DGPS (WADGPS) uses vector corrections for each satellite, derived from observations in a continental or global network of reference stations. The concept corresponds to option (c). The vector consists of individual corrections for the satellite clock, satellite position, and ionospheric delay model. Compared to a scalar correction, a vector correction is valid over much greater areas (Parkinson, Enge, 1996). The concept is discussed in more detail in [7.5.3.1].

Carrier Phase DGPS (CDGPS) is used for surveying applications (see RTK [7.5.2]) and also for attitude control of vehicles [7.6.2.7].

The basic model of the ordinary DGPS concept with code phases is as follows (cf. Misra, Enge, 2001). Starting from equation (7.44), we find expressions for the observed pseudoranges at the user station, PR_u, and at the reference station, PR_r:

$$PR_u = R_u + c\,(dt_u - dT) + d\,\mathrm{ION}_u + d\,\mathrm{Trop}_u + d\,\mathrm{Eph} + \varepsilon_{PR_u}, \quad (7.150)$$

$$PR_r = R_r + c\,(dt_r - dT) + d\,\mathrm{ION}_r + d\,\mathrm{Trop}_r + d\,\mathrm{Eph} + \varepsilon_{PR_r}, \quad (7.151)$$

with

$dt_u,\ dt_r$	the receiver clock errors,
dT	the satellite clock error with respect to GPS system time,
$d\,\mathrm{ION}_u,\ d\,\mathrm{ION}_r$	the ionospheric delays,
$d\,\mathrm{Trop}_u,\ d\,\mathrm{Trop}_r$	the tropospheric delays, and
$d\,\mathrm{Eph}_u,\ d\,\mathrm{Eph}_r$	the effect of the ephemeris error, at both stations.

The geometric range,

$$R_r = |\boldsymbol{x}_s - \boldsymbol{x}_r|, \quad (7.152)$$

is calculated from the known satellite position (from broadcast ephemerides) and the predetermined position of the reference station. Any variation with time in the above equations has been neglected for simplicity.

The error in the pseudorange observation at the reference station, the *Differential Correction DC*, is given by

$$DC = R_r - PR_r = -c\,(dt_r - dT) - d\,\mathrm{ION}_r - d\,\mathrm{Trop}_r - d\,\mathrm{Eph}h_r - \varepsilon_{PRr}. \quad (7.153)$$

In addition to DC, also the range rate of the correction, the *Differential Correction Rate, DCR*, is being determined and transmitted. DC and DCR refer to a reference epoch, t_k; they arrive at the user station with a certain delay or *latency*. Latency is hence defined as the elapsed time from the epoch of the measurement at the reference station until the use of the correction at the remote user site. At the user station, the corrections are predicted for the actual observation epoch, t, as

$$DC(t) = DC(t_k) + DCR \cdot (t - t_k). \quad (7.154)$$

The prediction of DC was critical with activated SA, because its changes with time were rather large. Today, a latency of up to 10 seconds would not be harmful. Typical

latencies are in the order of 3 to 4 seconds. In general it holds that DGPS corrections decorrelate with time if the application is delayed. With (7.150), (7.153) and (7.154) at hand, the corrected pseudoranges $\widehat{PR}_u$ for the user at the epoch of observation are, again omitting a time index, written as

$$\begin{aligned}\widehat{PR}_u &= PR_u + DC \qquad (7.155)\\ &= R_u + c\,(dt_u - dt_r) + (d\,\mathrm{Ion}_u - d\,\mathrm{Ion}_r) + (d\,\mathrm{Trop}_u - d\,\mathrm{Trop}_r) + \\ &\quad (d\,\mathrm{Eph}_u - d\,\mathrm{Eph}_r) + (\varepsilon_{PRu} + \varepsilon_{PRr})\\ &= R_u + c\,(dt_u - dt_r) + \delta_{\mathrm{Ion}} + \delta_{\mathrm{Trop}} + \delta_{\mathrm{Eph}} + \varepsilon_{PRur}. \qquad (7.156)\end{aligned}$$

For little or no latency, the satellite clock error, dT, is identical and hence disappears in (7.155) for no or only small latency. The biases, δ_{Ion}, δ_{Trop} and δ_{Eph}, are negligible for small station separations (a few kilometers) and grow (decorrelate) with increasing interstation distances.

For coordinate determination at the rover station the corrected pseudoranges from equation (7.155) are applied and lead to an improved positioning result. For nearby stations, the only remaining bias term in equation (7.156) is the combined user clock error, $c\,(dt_u - dt_r)$. With a latency of zero the equation is identical to the single difference equation in relative positioning (7.57). Note, that the terms DC and DCR are frequently named *PRC* (pseudorange correction) and *RRC* (range rate correction) in literature.

7.5.1.2 Data Formats and Data Transmission

Relative positioning requires the availability, at the rover station, of data from the reference stations. Several data formats have been developed for the transmission of such data. The two most important formats are RINEX, for post-processing or near-online purposes, and RTCM for real-time applications.

The *Receiver Independent Exchange Format* RINEX is described in [7.3.3.2]. It provides the complete data set and is accepted by most software packages. RINEX data can be obtained from the reference station provider via Internet and ftp, via mobile and fixed phone, CD ROM, or other data storage devices, depending on the service. Data from the IGS stations, for example, are available in the RINEX format [7.8.1].

For applications in real-time, in general, the transmission of the complete raw data from a reference station to the rover is not possible, because of the limited capacity of available transmission channels. Instead, the data are pre-processed at the reference station, and only a set of corrections is transmitted. Depending on the accuracy level, different types of corrections are required. For ordinary DGPS, the transmission of code-corrections is sufficient, for PDGPS carrier phase data are also required.

The standards established by the *Special Committee 104 of the Radio Technical Commission for Marine Services* (*RTCM SC-104*, or in short, *RTCM*) are today internationally accepted, and supported by nearly all receiver types. The preliminary version RTCM 1.0, from 1985, was replaced in 1990 by RTCM 2.0. This format

provides pseudorange and range rate corrections and is sufficient for ordinary DGPS (b) with an accuracy level of few meters or better. All necessary information can be transmitted with a bandwidth of 1200 bps (bits per second) or less. Depending on data rate and number of satellites, the required bandwidth can be reduced to 100 bps.

Version RTCM 2.1, from January 1994, additionally includes carrier phase data and hence provides the possibility to resolve ambiguities at the rover station. This version is the required standard for PDGPS and RTK. The necessary data rate is at least 4800 bps. Version RTCM 2.2, from January 1998, includes still further information, in particular the option to transmit correction data from more GNSS systems, e.g. GLONASS. Version RTCM 2.3, from May 2001, is a further refinement, allowing, for example, antenna phase center variation (PCV) data to be included. A new version, RTCM 3.0, is under discussion and will include capabilities for network RTK.

The RTCM message format is very similar to the format of the GPS navigation message. The messages consist of 30-bit words. Each word consists of 24 databits and 6 parity bits. Each message starts with a two-word header containing information such as the reference station identification, the message type, and a reference time for the parameters. In total, 64 message types are reserved, the majority not yet defined. Table 7.16 shows some of the message types valid in RTCM 2.3. For detailed information see e.g. Parkinson, Enge (1996), Kaplan (1996), Hofmann-Wellenhof et al. (2001) or the official document of the RTCM Special Commission 104 (RTCM, 2001).

Table 7.16. Selection of message types in the RTCM 2.3 format

Message type number	Current status	Title
1	Fixed	Differential GPS Correction
2	Fixed	Delta Differential Corrections
3	Fixed	Reference Station Parameters
18	Fixed	RTK Uncorrected Carrier Phases
19	Fixed	RTK Uncorrected Pseudoranges
20	Fixed	RTK Carrier Phase Corrections
21	Fixed	RTK High precision Pseudorange Corrections
31	Tentative	Differential GLONASS Corrections
32	Tentative	Differential GLONASS Reference Station Parameters
37	Tentative	GNSS System Time Offset
59	Fixed	User Defined

For precise DGPS, either message types 18/19 or 20/21 can be used. One advantage of types 20/21 is a higher compressibility when compared with the raw carrier data 18/19. Proprietary formats have been developed to transmit compressed – and possibly decoded – corrections together with other information in the user defined message, type

59. This compression allows the transfer of reference data for all satellites in view (up to 12 SVs) with a bandwidth of just 2400 bps (Wübbena et al., 1996).

The possible data links for the transmission of DGPS data may be categorized as follows:

- ground-based radio links,
- cellular phones,
- satellite communication, and
- internet.

For ground-based radio links there exists a set of general rules:

- the lower the frequency the larger the range,
- the higher the frequency the higher the bandwidth and data rate, and
- high frequency and short range installations are cheaper than low frequency and long range installations.

In the *Low Frequency* (LF) domain, < 300 KHz, data can be transmitted over several hundred kilometers because the waves follow Earth's curvature. One example in Germany, is the system ALF (Accurate Positioning by Low Frequency) where one transmitter near Frankfurt covers all of Germany and beyond (600 to 800 km) with DGPS data at a rate of 3 seconds.

Medium frequency (MF), 300 KHz–3 MHz, transmitters are cheaper than LF transmitters and also cover several hundred kilometers. They are often used as marine radio beacons in coastal areas. The data capacity is sufficient for ordinary DGPS services (several meters accuracy).

Very High Frequency (VHF), 30–300 MHz, and *Ultra High Frequency* (UHF), 300–3000 MHz, radios can communicate over short distances, limited by the line of sight. The data capacity is in the range of 2400 bps and hence suffices for PDGPS with a compressed data transmission format (see above). A further possibility is to use a frequency modulation subcarrier in the *radio data system* (RDS) from broadcasting services. The capacity is then in the order of 100 bps, which is sufficient for ordinary DGPS.

Mobile UHF radio systems, operating at 400 MHz and higher have a range of several kilometers, depending on their power, and can transmit 9600 bps, which is sufficient for RTK applications. A frequently-used wavelength is 70 cm (428 MHz) which, for example, at low power (0.25 W) can be operated without permission in Germany.

Cellular phone is spreading across densely populated areas to provide telephone and data services. With decreasing user fees cellular phones are an attractive alternative to radio frequencies. One disadvantage, however, is that with cellular phones the number of simultaneous users is limited by the number of modems at the reference station, whereas the broadcast system works for an unlimited number of users.

Globally-operating DGPS services use geostationary communication satellites like INMARSAT, or a network of Low Earth Orbiters like *Globalstar*, to transfer DGPS data. The DGPS corrections provide accuracies at the few-meters level. Some of the services use the WADGPS concept [7.5.3.1]. The capacity of INMARSAT also offers PDGPS applications.

A very powerful technique is the distribution of DGPS data via Internet. GPS data from global or regional permanent arrays like the IGS or EUREF networks are already available in the RINEX format on a routine basis (de Jong, 2001). Since quite recently data for real-time applications are also accessible via internet. The JPL is building up an "Internet-Based Global Differential GPS System" (Muellerschoen et al., 2000). Regional and local RTK services via Internet are under development (Weber, 2002). Table 7.17 gives an overview on some DGPS data channels.

Table 7.17. Frequently used data links for DGPS transmission

Name	User	Range	Capacity	Type
radio 2 m	unlimited	tens of km	2400 bps	P-DGPS
radio 70 cm	unlimited	few km	9600 bps	P-DGPS
radio LF	unlimited	hundreds of km	300 bps	DGPS
radio MF	unlimited	hundreds of km	100 bps	DGPS
radio UHF/RDS	unlimited	tens of km	100 bps	DGPS
satellite	unlimited	global	$>$ 2400 bps	DGPS
mobile phone	1 per channel	variable	9600 bps	P-DGPS
internet	unlimited	global	$>$ 9600 bps	P-DGPS

7.5.1.3 Examples of Services

During the last years a large number of GPS reference station services with different architecture and performance has appeared. We can distinguish between global, regional, national, and particular services, as well as between public and commercial, or post-processing and real-time services. In the following, some examples are given.

The most important global reference network is maintained by the *International GPS Service* (IGS) [7.4.3.2][7.8.1], a non-governmental scientific organization. More than 300 stations worldwide are continuously operating and provide, among other products, information on position and observation data, via regional and global data centers. The IGS network (Fig. 7.101, p. 399) is closely related to the ITRF reference frame [2.1.2.2], hence it is possible to connect new GPS observations everywhere in the world directly to the ITRF. The IGS stations generally do not transmit DGPS data in real-time. The IGS is a passive global reference network mainly used for post-processing purposes.

Another global service under development is the *Global Differential GPS System* (GDGPS) of the NASA JPL (Bertiger et al., 1999; Muellerschoen et al., 2001). Based on observations from about 60 stations in the NASA global network, state parameters are modeled and provided to users via Internet. The quasi real-time accuracy is estimated to be 10 cm for the horizontal position and 20 cm in height with a latency of about 1.5 to 3 seconds. The GDGPS approach belongs to option (iii) in [7.5.1.1], see also [7.5.3].

Examples of global commercial services are *Skyfix* and *Omnistar*. Skyfix maintains a network of about 80 reference stations within the reference frame ITRF92. Correc-

tion data, in the RTCM 2.0 format, are transmitted via INMARSAT communication satellites. The achievable accuracy is about 2 m. Omnistar runs about 70 reference stations, covering about 95% of the world. The correction data are distributed via 9 different geostationary satellites, also in the RTCM 2.0 format, and submeter accuracy is promised. Both services apply some state space modeling [7.5.3] in order to obtain the indicated accuracy over large distances.

At the continental level, the *EUREF Permanent Network* (Fig. 7.75, p. 358) can be considered to be a densification of the IGS in Europe. It consists of about 140 (status July 2002) permanent stations and has a similar structure to the IGS. The main purpose is maintenance and control of the European Reference Frame, ETRF 89 [2.1.2.2] (Ádám et al., 2000). Similar so-called "regional networks" are operated in other parts of the world. For example, in South America, in the SIRGAS project [7.6.2] a permanent network of about 30 stations, related to IGS, is continuously operated (Seemüller, Drewes, 2000).

National networks are being established worldwide. Existing maritime radiobeacons are used to broadcast DGPS data, in the standard RTCM format, to marine users in the LF- and MF-bands (285–325 kHz maritime radiobeacon band) with a data rate of 100 to 200 bps (Parkinson, Enge, 1996; Mangs et al., 2001). Beacon networks are mainly located along coastlines or large navigable inland waterways, but are in some regions also expanded inland. In the U.S. the beacon network is the responsibility of the Coastguard. There are plans to cover the whole U.S. with about 80 stations in the *Nationwide Differential Global Positioning System* NDGPS (DOD/DOT, 2001a). Complete coverage is expected for 2003. In Europe, the beacon network nearly covers the complete coastline of the European coastal states (see Fig. 7.62). Within the *EUROFIX* project, tests are underway to use existing Loran C stations for the transmission of DGPS data (Helwig et al., 1997).

The U.S. National Geodetic Survey (NGS) runs the CORS network. CORS stands for *Continuously Operating Reference Stations*. It consists of about 200 stations and is still growing. CORS will meet the post-processing requirements of positioning users by providing code phase and carrier phase observation data in the RINEX format. The data are freely accessible via Internet or anonymous ftp. Depending on the station, the data are recorded at 1, 5, 15, or 30 seconds intervals. For details see Snay (2000).

Canada is establishing the *Canadian Base Network* (CNB), with a spacing between 200 km and 1000 km, depending on the area. CNB consists only of pillar monuments, and is connected to the *Canadian Active Control System* (CACS). CACS basically consists of a number of active stations (14 in 2002), partly remotely controlled. The dual frequency pseudorange and carrier phase data are transmitted to a processing center in Ottawa and are being used as a backbone for estimating state vectors in wide area DGPS systems (e.g. the *Canada-wide Differential GPS* (CDGPS)), see [7.5.3].

In Japan, a nationwide GPS control network, also named *GEONET*, is being established under the responsibility of the Japanese Geographical Survey Institute (GSI). The network consist of about 900 sites, equipped with dual-frequency GPS receivers and additional sensors like tiltmeters and meteorological stations. The spatial density

Figure 7.62. DGPS Beacons in Europe

is very high; the mean distance between stations is about 25 to 30 km. The primary purpose of the network is the determination of crustal strain for earthquake monitoring and prediction [7.6.2.2]. The stations equally provide reference code and carrier data, for surveying purposes, both in real-time and for post-processing. In addition, the data can be used to map tropospheric zenith delay, and to contribute to weather forecasts [7.6.2.9].

In Brazil, a continuously growing network of active reference stations, the *Rede Brasileira de Monitoramento Continuo* (RBMC), coordinated with respect to SIRGAS, provides reference data for precise post processing (Fortes et al., 1997).

Relatively dense permanent networks are already running, or being established, in European countries. Examples are SWIPOS in Switzerland, SWEPOS in Sweden, SATREF in Norway, and SAPOS in Germany. In Great Britain the *Ordnance Survey National GPS Network*, consisting of 30 active stations, provides RINEX data via Internet for users who thereby can directly access the national coordinate system. Additionally, over 1000 passive GPS points are available to support surveying measurements (Cruddace, 2001).

The above examples demonstrate the many activities and different approaches for DGPS services all over the world. In the following, the SAPOS project in Germany

will be explained in more detail because it provides a large variety of services.

SAPOS stands for *Satellite Positioning Service*, and is organized by the German State Surveying Agencies (AdV). The final objective is to cover the complete area of Germany with a network consisting of about 250 permanent stations at a separation of about 40 to 70 km. The rationale behind SAPOS is to provide services to many users who have different requirements concerning accuracy of position results, required observation time, and coverage. SAPOS runs different services, providing different accuracy levels, namely (Hankemeier, 1996)

EPS Real-Time Positioning Service,
HEPS High Precision Real-Time Positioning Service,
GPPS Geodetic Precise Positioning Service, and
GHPS Geodetic High Precision Positioning Service.

EPS is similar to many commercial and national DGPS services, and provides an accuracy of 1–3 m, sufficient for a broad variety of navigational applications. The correction data are available free of charge via different communication channels.

HEPS is the principal precise real-time service, and it can be used for many applications in surveying and GIS, including cadaster. The position accuracy is between 1 cm and 5 cm, and depends on several influences, in particular on the behavior of distance dependent errors. In order to obtain 1 cm accuracy in real-time over distances larger than a few kilometers, it is necessary to model the error state in the working area (see [7.5.3]). For the application of HEPS a special decoder and payment of user fees are required.

GPPS provides 1 s data from the reference stations for a limited time (e.g. 10 days; thereafter the data are reduced to 15 seconds). Via mobile phone, the data can be transmitted directly to the user in the field for precise near real-time positioning. GHPS requires precise ephemerides, and is a post-processing service. An overview of the SAPOS services is given in Table 7.18.

Table 7.18. SAPOS products

DGPS Service	Positioning Accuracy	Positioning Mode	Data Format	Data Transmission	Data Rate
EPS	1 – 3 m	Real-time	RTCM 2.0	LF, UHF 2 m Band	3 – 5 s
HEPS	1 – 5 cm	Real-time	RTCM 2.1 Modified	2 m Band Cellular Phone	1 s
GPPS	1 cm	Quasi Real-time Post-processing	RINEX	Cellular phone Fixed Phone Data Network	1 s (15 s)
GHPS	< 1 cm	Post-processing	RINEX	Data Network Data Storage	1 s (15 s)

7.5.2 Real Time Kinematic GPS

Real Time Kinematic GPS (RTK) is another name for carrier-phase differential GPS (option (c) in [7.5]). Its eminent characteristic is that users can obtain centimeter-level positioning accuracy in real-time over short distances with an easy-to-handle and highly integrated instrumentation. It is the RTK technology that makes GPS a universal surveying tool, replacing traditional surveying techniques. RTK technology is based on the following features:

- transmission of pseudorange and carrier phase data from a reference station (base station) to the user station (rover) in real-time,
- resolution of ambiguities at the rover station "on the way" or "on the fly" (OTF), and
- reliable determination of the baseline vector in real-time or near real-time.

In a typical RTK configuration (see also Fig. 7.84, p. 369, in [7.6.2.4]), a local (usually stationary), GPS reference receiver transmits pseudorange and carrier phase data over a radio link to the roving station. The GPS receivers may be single- or dual-frequency receivers; dual band equipment facilitates ambiguity resolution (faster, more reliable). The data are transmitted via a data radio (modem). The software runs in the receiver or an external data processor. In general, for the sake of flexibility, identical equipment is used at both stations. Most manufacturers provide highly integrated systems for both the base and rover stations (see [7.2.4.2]).

For the transmission of RTK data, new message types were defined in the RTCM SC-104 version 2.1, in 1994. Message types 18 and 19 contain raw carrier phase and pseudorange information. Alternatively, message types 20 and 21 can be used, containing corrections to the measurements at the reference station.

The transmission of RTK data requires a much higher capacity than does broadcasting of pseudorange corrections. For a set of raw phase corrections from 12 satellites, in format RTCM 2.1, more than 4800 bits are necessary. For a data rate of 1 second the usually available transmission channels with 2400 bps are not sufficient, so that alternative channels with higher data capacity, or particular (mostly proprietary) data formats are required [7.5.1.2]. Depending on the legal situation in a given country, VHF and UHF channels for the required data capacity, but with rather low power, can be used, sometimes without permission. As a consequence, the range of commercial RTK systems is often reduced to a few kilometers.

Another reason for the short range of VHF and UHF transmission is its limitation to the line of sight. The theoretically achievable maximum distance, d, in kilometers between base and rover is (Langley, 1998c) calculated according to

$$d = 3.57\sqrt{k}\,(\sqrt{h_t} + \sqrt{h_r}). \tag{7.157}$$

h_t and h_r are the heights in meters of the transmitting and receiving antennas above the average terrain level. k is a factor depending on Earth's curvature and the atmospheric refractivity. A mean value for moderate climates is 1.33. To give an example, for a transmitting antenna at 25 m, and a rover antenna at 2 m, above the terrain, the

theoretical maximum signal range is 26 km. In practice signals may be blocked by hills, trees or buildings (see [7.4.4.4]).

Data transmission for RTK applications via Internet is still in the experimental stage (Weber, 2002).

A key factor for RTK is the ability of the rover to resolve ambiguities while the receiver antenna is in motion. This feature is named ambiguity resolution "on the way" (OTW) or, more frequently, *on the fly* (OTF) (for details see [7.3.2.3]). As long as the ambiguities are estimated as real values, (the so-called *ambiguity float-solution*), the achievable accuracy ranges from the meter to the decimeter level, depending on the tracking time. With resolved (fixed) ambiguities, the accuracy numbers go down to the centimeter level. In order to prove the correctness of the ambiguity fixing, the algorithm is initialized at least twice.

Another important factor is the necessary *time to fix ambiguities* (TTFA). Many OTF algorithms use the wide-lane linear combination to accelerate their ambiguity search procedures. The search algorithms are more effective with a large number of satellites and a limited search space. The requirements for a suitable RTK receiver are hence:

- dual frequency data for ambiguity resolution, also if the baseline is derived from single frequency data,
- low-noise code pseudoranges to narrow down the ambiguity search space, and
- all in view capability, to use as many satellites as possible for the search algorithms.

The necessary TTFA strongly depends on the behavior of the distance dependent errors [7.4.4]. For short distances (a few kilometers), and under favorable conditions, the TTFA can be as short as only one epoch.

For the estimation of the baseline vector between the base and the rover station, powerful real-time software is required. The data processing can either follow the concept of parameter elimination (single and double differences) or parameter estimation (undifferenced observables) [7.3.2.2]. Either raw data (RTCM message types 18, 19) or carrier phase corrections (RTCM message types 20, 21) can be used. The carrier phase correction, CPC_r, at the reference station, is given by

$$CPC_r = \Phi_r - \mathrm{Frac}\left(\frac{R_r}{N_r \lambda}\right), \tag{7.158}$$

with

Φ_r the raw phase at the base station,
N the resolved ambiguity at the base station,
R_r the geometric range at the base station, and
λ the wavelength.

Furthermore

$$N = \mathrm{Int}\left(\frac{R_r}{\lambda}\right).$$

The corrected phase, $\hat{\Phi}_u$, at the rover station is given by

$$\hat{\Phi}_u = \Phi_u + CPC_r. \tag{7.159}$$

In analogy to (7.154), the phase correction rate *PCR* is also transmitted and applied. The prediction, from the reference epoch, t_k, to the observation epoch, t, is given by

$$CPC(t) = CPC(t_k) + PCR \cdot (t - t_k). \tag{7.160}$$

Carrier phase corrections have several advantages when compared with transmission of raw phase data (Wübbena et al., 1996, 2001b). Corrections are less receiver dependent and more flexible than raw data; this minimizes problems arising from the use of unequal receiver equipment. In addition, corrections for local errors like antenna phase center variations can be applied. Corrections can also be derived from several reference receivers, which makes network solutions possible (cf. [7.5.3]). Another important aspect is that phase corrections require much less bandwidth for transmission than the raw phase data. A capacity of 2400 bps is sufficient.

Most RTK algorithms are proprietary solutions and have a key impact on the performance of commercial RTK equipment. With modern RTK sets a wide variety of surveying tasks can be solved. A short list of possible applications is:

- GIS, cadaster,
- detailed surveying,
- staking out,
- machine control, and
- precision farming.

More applications are discussed in [7.6.2]. A significant limitation of RTK solutions is the fact that the errors and TTFA grow with increasing distance from the base station. A general rule of thumb for the achievable accuracy is

10 mm + 1 to 2 ppm for horizontal coordinates, and
15–20 mm + 2 ppm for the height component.

RTK applications are therefore limited to a range of a few kilometers (mostly below 2 km) with a TTFA of just a few seconds. Only in times of low ionospheric disturbances can the range be larger, up to 10 km; however the TTFA will significantly increase. For larger distances, the use of multiple reference stations solves the problem (see [7.5.3.2]).

For more details about RTK see e.g. Langley (1998c), Hofmann-Wellenhof et al. (2001) or the information brochures of receiver manufacturers.

7.5.3 Multiple Reference Stations

One of the serious drawbacks of DGPS is the fact that the influence of some error sources, such as orbit, ionosphere and troposphere, grows with increasing distance from the reference station. In other words, the *error correlation* decreases with station separation, and the *error decorrelation* grows. The effect is roughly 1 meter increase in the DGPS positioning error per 100 to 150 km, for single frequency code receivers.

It would require a very large number of individual DGPS installations to cover a single country or even a whole continent.

A solution to this problem is the idea to interconnect several reference stations, and to transmit their measurement data in real-time to a central processing station. All data are used in a common filter to estimate the error state for the whole area and to separate the error components. The state vector can then be applied to improve the corrections for the complete area, as a function of the geographic user location, resulting in much better accuracy and a much sparser density of reference stations. Alternatively, as a first step, simple interpolation algorithms can be used.

This concept was developed early on under the name *Wide Area Differential GPS* (WADGPS) for the use of code-range measurements in continental networks [7.5.3.1]. Only recently, a similar concept has been developed for high precision DGPS using carrier phase data in local, regional or national networks. The concept is known as *Networked Reference Stations*, *Virtual Reference Stations*, or *Area Correction Parameter Approach* [7.5.3.2]. Note that the terms are not uniformly used in the literature.

7.5.3.1 Wide Area Differential GPS

The term *Wide Area Differential GPS* (WADGPS) was coined by C. Kee and others, in 1991 (Kee et al., 1999). The basic idea was to establish a sparse network of reference stations, over an area as large as the continental U.S. (CONUS), to provide high quality DGPS corrections to navigation users at land, on sea, in the air, as well as in the near-space. A WADGPS system consists of (see Fig. 7.63)

- a sparse network of reference stations, equipped with dual frequency GPS receivers, high precision clocks and optional meteorological sensors,
- a master control station that receives all measurements and estimates the differential corrections,
- an upload station, to broadcast the corrections, possibly via GEO satellite, to the users, and
- monitor stations, to control the system.

The main objective is to overcome the decorrelation of the distance dependent errors by using suitable network algorithms. An excellent overview of the basic algorithms is given by Mueller (1994). A rough separation is into measurement and state-space domain algorithms.

Measurement domain algorithms do not estimate the individual error components, but form a weighted mean of all corrections from the participating reference stations. The weighting scheme may use

- a distance weighting (the nearest station gets the highest weight),
- an elevation angle weighting (higher elevation satellites get more weight),
- an age weighting (lower latency gets higher weight),

or other criteria.

State space domain algorithms try to identify the individual error sources and transmit the information to the user in a suitable form. The components are, in particular,

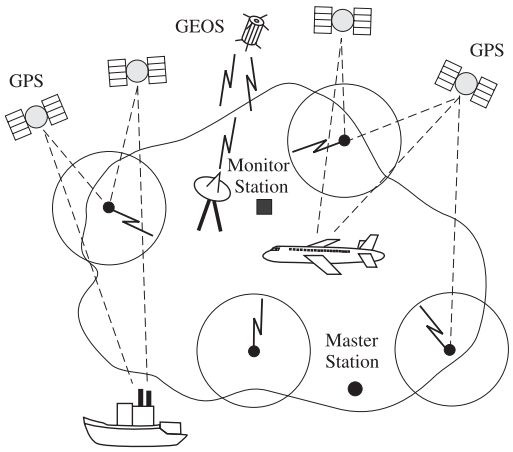

Figure 7.63. Architecture of an Wide Area DGPS installation

a 3-D ephemeris error, a satellite clock offset, ionospheric delay parameters, and tropospheric parameters. User apply this information as a function of their geographical location.

The measurement domain algorithms are the simplest, and hence cheapest, solutions. The corrections, however, are not independent of distance, but degrade with growing separation from the center of the network. The concept is hence seldom used. The state-space domain approach is baseline independent and provides the highest accuracy. The error components in the state vector also vary rather slowly, so that the data transmission rate can be low. For details about the algorithms see also Kee (1996); Kaplan (1996).

The advantages of using WADGPS when compared with a single DGPS reference station are obvious:

- coverage can be extended over inaccessible regions like water areas,
- the number of reference stations can be reduced, and
- the biases are nearly or completely distance independent.

The WADGPS concept has been realized in various services. In the U.S., the Federal Aviation Administration (FAA) is establishing the *Wide Area Augmentation System* (WAAS) to meet safety-related integrity requirements and to support as well the en route as the precision approach phases of flight (FRNP (2001), see also [7.7.2]). In Europe, the *European GPS Navigation Overlay System EGNOS* has similar objectives [7.7.2]. The same is true for the Japanese service (MSAS). All three augmentation systems will be interoperable to provide seamless global coverage. In Canada the *Canada-Wide Differential GPS* (CDGPS) provides accurate differential corrections, via communication satellite, for the whole country to support positioning and navigation at the meter level. The service is, however, not intended for commercial aviation (no integrity channel).

The aforementioned commercial services, *Skyfix* and *Omnistar*, like other global or regional commercial services, also apply the WADGPS concept. The approach of

Omnistar is very similar to the idea of "virtual reference stations" VRS, see [7.5.3.2].

7.5.3.2 High Precision Networked Reference Stations

With the installation of reference stations providing carrier phase data for precise DGPS applications in real-time (e.g. SAPOS in Germany [7.5.1.3]) the problem of distance dependent errors became evident. When 1 cm accuracy is required, the number of reference stations with the necessary density would be unrealistically high, in particular during periods of strong ionospheric disturbances. A solution to the problem comes from interconnection of the reference stations, and the estimation of the error state in the working area in real-time (state-space domain approach, option (iii) in [7.5.1.1]). Fig. 7.64 demonstrates the problem and its solution. For "ordinary" baseline RTK (without network), the achievable accuracy decreases with increasing distance; at the same time the TTFA increases. In the networked solution the error state in the area is estimated and transmitted to the rover, where the measurements can be corrected accordingly. As a result, the accuracy and the TTFA remain at a constant level independent from the distance.

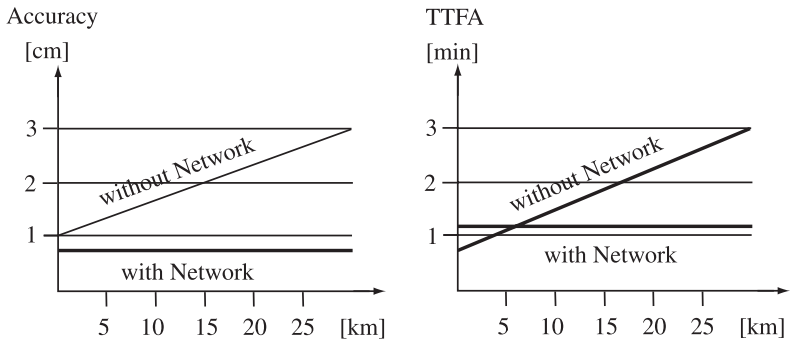

Figure 7.64. Modelling the distance dependent errors in an interconnected reference network

The current data formats for DGPS corrections do not allow to transmit the complete state vector to the roving station; hence the state vector needs to be represented by a simplified model. The basic idea of this approach is as follows (Wübbena, Willgalis, 2001). The observation equation for carrier phase observations between the antenna phase centers of satellite, i, and receiver, j, is

$$PR_j^i = |\boldsymbol{R}_j^i| + \delta B_j^i + \lambda N_j^i + \varepsilon_j^i. \tag{7.161}$$

The equation can be written for each particular signal, s; the index, s, is omitted for clarity. The bias term δB_j^i comprises the terms

δC_j^i for clock related errors,
δD_j^i for distant dependent errors, and
δS_j^i for station dependent errors,

hence

$$\delta B_j^i = \delta C_j^i + \delta D_j^i + \delta S_j^i. \tag{7.162}$$

The clock related errors δC_j^i contain components originating in the satellite and the receiver clock, and signal delays in the hardware of the satellite and the receiver. The distance dependent errors, δD_j^i, are composed of the ionospheric delay, δI_j^i, the tropospheric delay, δT_j^i, and the orbit error vector, $\delta \boldsymbol{o}^i$, hence

$$\delta D_j^i = -\delta I_j^i + \delta T_j^i + \frac{\boldsymbol{R}_j^i}{|\boldsymbol{R}_j^i|}\delta \boldsymbol{o}^i. \tag{7.163}$$

The station dependent errors, δS_j^i, finally, are composed of multipath and receiver antenna phase center variation (PCV). For completeness, multipath and PCV at the satellite antenna can be included.

Precise positioning with carrier phases requires the correct determination of the phase ambiguities, N. As has been outlined in [7.3.2], two approaches are possible:

- parameter elimination, and
- parameter estimation.

In the parameter estimation approach all biases have to be estimated together with the coordinates and ambiguity terms. The parameter estimation procedure with undifferenced observables has some advantages when compared with the parameter elimination process, for example (see also [7.3.2.2]):

- biases can be constrained by specific models,
- precise clock models can be used,
- absolute information is maintained,
- more flexibility with changes in network design, and
- different receiver types and different GNSS signals can be adapted more easily.

It is obvious that the parameter estimation concept is particularly well suited to be used in the state space approach, because all biases can be separately modeled and the distance-dependent biases can be applied as corrections. Table 7.17 gives an overview of possible functional and stochastic models for the above mentioned error sources.

Once all state parameters are estimated with sufficient accuracy, they can be transmitted to the user, who can eliminate the corresponding error terms from the observation equation. As a result, precise absolute coordinates for the user antenna can be determined. This is basically the approach for precise point positioning (PPP) in global and regional networks (see e.g. [7.3.4] and Muellerschoen et al. (2001)). For operational multi-station networks, for the time being, a simplified *state representation* is used instead of the complete state vector.

In a first step a network solution with ambiguity fixing is established for the participating reference stations, and all states are properly estimated. The measured ranges at the reference stations are filtered using the state space model of the network. The differences between filtered and computed ranges, for all satellites, give residuals which

Table 7.19. Functional and stochastic description of GPS error sources, after Wübbena, Willgalis (2001)

Bias	Functional Model	Stochastic Model
Satellite clock	2nd order polynomial	white noise process
Signal delay (SV)	constant	white noise process
Satellite orbit	Cartesian elements	3D Gauss–Markov process
Ionospheric delay	single layer model	3D Gauss–Markov process
Tropospheric delay	modified Hopfield model	2 scaling parameter/station
Receiver clock offset	–	white noise process
Signal delay (rcv)	constant	white noise process
Satellite PCV	–	–
Receiver PCV	calibration	–
Multipath (rcv)	elevation dependent weighting	1st order Gauss–Markov process
Measurement noise	–	white noise process
Carrier phase ambiguity	constant after fixing	–

are separated into ionospheric, orbit and tropospheric residuals. This separation is possible because of the proper state estimation. In some cases, the orbit and tropospheric residuals are combined as *geometrical residuals*.

In a second step, the residuals are interpolated between the reference stations. Fig. 7.65 demonstrates that a rover experiences an error, δ_1, by using the range correction, ε_1, only, and an error, δ_2, by using the range correction, ε_2. With a linear interpolation between the reference stations, RS_1 and RS_2, the interpolation error is just $\delta\varepsilon$. Investigations show that up to distances of 100 km a linear representation is sufficient. Larger spacing between reference stations requires use of a polynomial of 2nd or higher order, depending on the spatial decorrelation characteristics of the particular error sources.

From the various proposals of how to correct the measurements at the rover station, two frequently applied procedures are the concepts of

- Area Correction Parameters (ACP), and
- Virtual Reference Stations (VRS).

The interpolation between two stations, as in Fig. 7.65, only models the errors along one baseline. For three reference stations, the state residuals can be represented by a plane (Fig. 7.66). For each epoch, the time variable parameters, $a_\varphi(t)$ and $a_\lambda(t)$, describing the inclination of the plane are determined. The parameters $a_\varphi(t)$ and $a_\lambda(t)$ are called *area correction parameters* (ACP). They are estimated at a rather slow rate, of about 10 seconds, separately for the two distance dependent error components, the *ionospheric component* and the *geometric component*, and transmitted in addition to the conventional PDGPS range corrections (e.g. RTCM 2.1) from the reference station

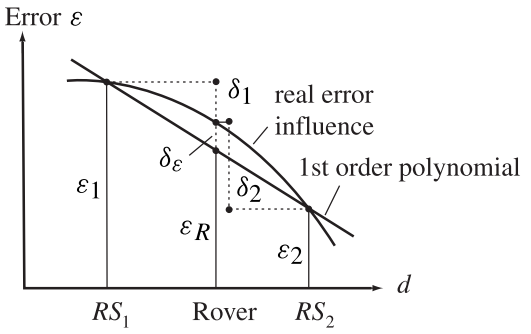

Figure 7.65. Interpolation of distance dependent errors (Wuebbena1998)

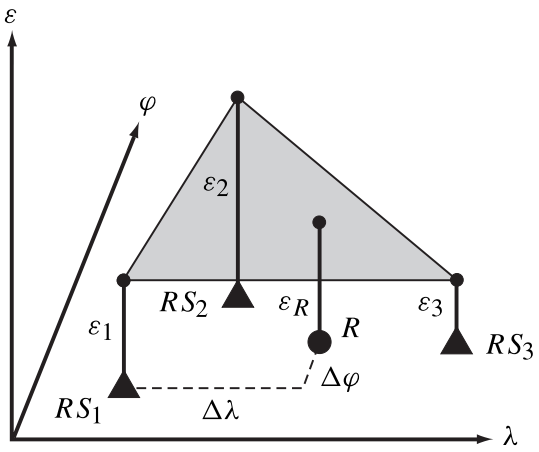

Figure 7.66. Linear modeling of area correction parameters (ACP) for three reference stations

located at φ_0, λ_0 to the rover. With

$$\varepsilon_R(t) = a_\varphi(t)(\varphi - \varphi_0) + a_\lambda(t)(\lambda - \lambda_0), \tag{7.164}$$

the user can then compute the corrections, $\varepsilon_R(t)$, valid for his approximate position, φ, λ.

Experiences show (Wübbena et al., 2001b; Willgalis et al., 2002) that the approach works well for distances up to about 50 km between reference stations. For larger station separation, more sophisticated representation techniques have to be developed.

An alternative approach is the use of so-called *virtual reference stations* VRS (Weber, Tiwari, 1995; Wanninger, 1998, 2000). Here, the user communicates his approximate position to the analysis center of the reference network. Based on a state estimation of the network, as discussed before, the analysis center computes a set of range corrections valid for the approximate rover positions and transmits these "virtual observations" to the rover. The rover accepts this data set as PDGPS corrections from a nearby reference station, and applies conventional RTK algorithms.

Both procedures have advantages and disadvantages. In the VRS concept only one data link is required between reference and rover station, and conventional RTK software can be applied. Disadvantages are that only a limited number of users can work at the same time, and that for moving rovers the virtual reference station is also moving. Most RTK software, however, only accepts fixed reference stations. The main advantages of the ACP concept are its unlimited number of users, higher flexibility and correctness in the error state modeling when undifferenced phase data are used, and the possibility to include moving reference stations.

As a result of either procedure, the rover is able to determine its position in real-time with an accuracy of about 1 cm, independent of the distance from the reference station, see Fig. 7.67. This situation helps to support many tasks in detailed surveying [7.6.2].

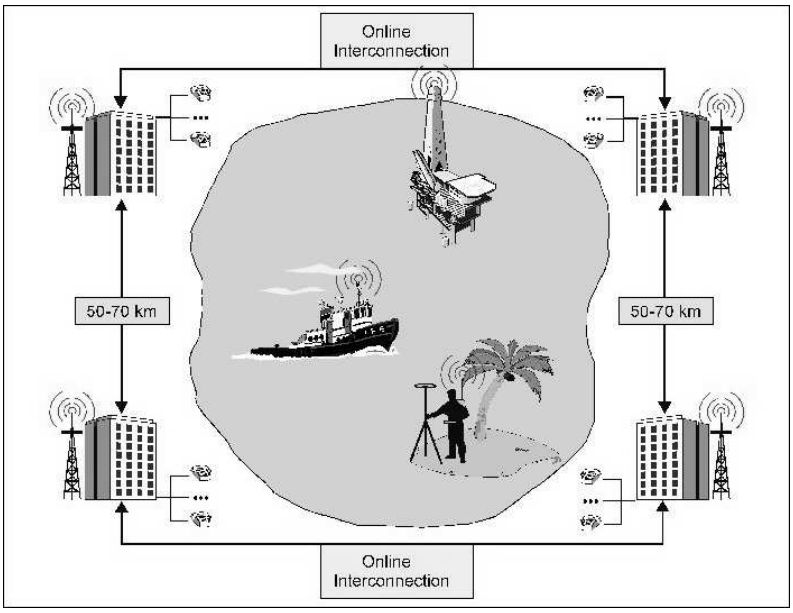

Figure 7.67. Concept of precise real-time positioning in an interconnected network

The concept of multiple reference stations for precise positioning in real-time is already applied in some areas, for example in Germany within the SAPOS service ([7.5.1.3], Jahn, Winter (2002)). Other developments are reported in Vollath et al. (2000) or Raquet, Lachapelle (2001). In total, however, development is still in its early stages. A next step will be the unification of existing networks, and an extension of services (Wübbena et al., 2001b; Wübbena, 2002). Global state parameters (orbit, global ionosphere, and troposphere) can be taken from global networks, like the IGS, and introduced into the state modeling of regional or local networks. The same is true for a combination of sparse national and dense regional networks. The state parameters of the national network, including solved ambiguities, are forwarded to regional or local networks that are only established in more densely populated areas. This concept, of an *adapted PDGPS network*, is in particular suitable for large countries where a uniform dense coverage is neither feasible nor required (Willgalis et al., 2002).

7.6 Applications

7.6.1 Planning and Realization of GPS Observation

In the early days of GPS, the planning and execution of field projects resembled, in many aspects, the execution of Doppler projects with TRANSIT [6.6]. Experience gained in the preparation, organization, and execution of Doppler projects could be transferred for the most part to GPS work, and has influenced GPS practice. It is hence of interest to study chapter [6] and to read some of the original TRANSIT publications. Due to the much broader field of applications, however, the issue of

GPS project planning and realization has its own significance, and is widely discussed in the literature, e.g. Jäger (1990); Santerre (1991), Hofmann-Wellenhof et al. (2001, chap. 7). The following gives the most important aspects.

7.6.1.1 Setting Up an Observation Plan

As long as the GPS system was not yet complete, a pre-computation of satellite coverage was an indispensable preparatory step in project planning. With the system completely deployed in 1995, sufficient satellites are visible above the horizon at any time; hence field campaigns can be planned independently of the constellation. For analysis purposes, and for kinematic observations, in particular in areas with obstructions, a pre-computation of the satellite constellation can still be of importance.

These so-called *ALERT-lists* can be computed with data from the satellite almanac. *Almanac data*, that is, low-accuracy orbit data for all available satellites, are transmitted in the fourth and fifth subframes of the navigation message [7.1.5.4]. These subframes have 25 "pages" each 30 seconds long, so that the complete almanac information can be read in 12.5 minutes.

With the aid of the almanac data, satellite positions can be precomputed over several months with sufficient accuracy for planning purposes. One must, however, occasionally expect larger orbit maneuvers, so that a regular check of the almanac data is recommended. With the almanac data, visibility diagrams (Fig. 7.68) and PDOP values [7.4.2] can be generated. Most manufacturers provide suitable software packages (*mission planning software*) on a PC basis. The almanac data are listed in Table 7.20. Almanac data are available from various internet sources, for example from the U.S. Coast Guard.

Table 7.20. Almanac data

PRN	Space vehicle identification number [–],
000	Health status (000 = healthy) [–],
e	Orbit eccentricity [–],
$\sqrt{A}$	Square root of the semi-major axis [$\sqrt{m}$],
Ω_0	Right ascension of the ascending node [degrees],
ω	Argument of perigee [degrees],
$\overline{M}_0$	Mean anomaly [degrees],
t_{0a}	Reference time for almanac data [s],
δ_i	Difference of orbit inclination from 54° [degrees],
$\dot{\Omega}$	Nodal rate [degrees/s$\times 10^{-3}$],
a_0	Clock correction [s$\times 10^{-9}$],
a_1	Drift of clock correction [s/s$\times 10^{-9}$], and
XXX	GPS week [–].

From the almanac data the current position vectors of the satellites can be calculated in the CTS coordinate system, using the formulas in [7.1.5.3]. With known approximate absolute coordinates (φ, λ, h) of the observation site, the satellite's azimuth and elevation can be found as a function of time with standard formulas, and used for the construction of visibility diagrams. A corresponding visibility diagram in stereographic projection (sky plot) is shown in Fig. 7.68. The related bar diagram is given in Fig. 7.15, p. 231.

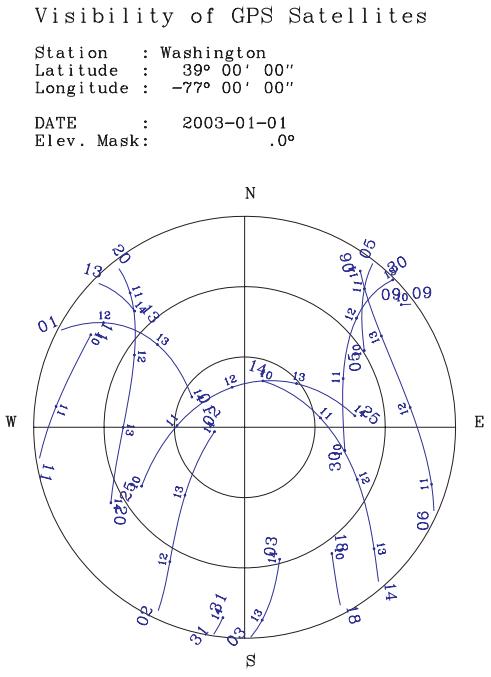

Figure 7.68. Visibility diagram (*Sky Plot*), 4 hour period, for Washington DC

The sky plot shows a certain lack of symmetry in the distribution of satellite tracks. This comes from the fact that the inclination of the GPS orbits is 55° and hence defines an area of the observer's sky (shadow area) where it will not be possible to make observations. The shadow area is a function of the observer's latitude and is equal for any longitude. Fig. 7.69 gives an example for equatorial, polar and mid-latitude observers (Santerre, 1991). For observers at northern mid-latitudes the shadow area is also called the "northern hole".

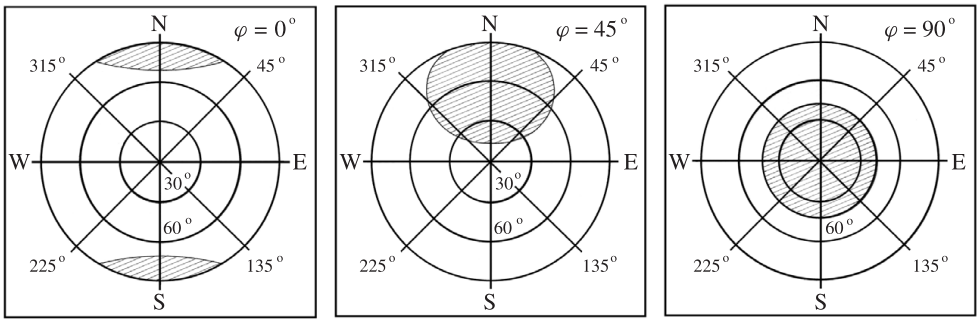

Figure 7.69. Shadow area as a function of the observers geographic location

Additional information for planning purposes is given by the computation of PDOP values that reflect the geometrical strength of the satellite configuration [7.4.2]. As long as the complete satellite coverage had not yet been installed the PDOP values indicated the best time periods for observations (*observation window*). With the current status of GPS, the importance of the PDOP criterion should not be over-emphasized; most

geodetic receivers can track all visible satellites and do not require any pre-selection, whilst for most geodetic applications the observation period is long enough to average out the influence of geometry. With the full GPS constellation the PDOP is sufficiently low for most of the day. Hence the PDOP criterion is of interest only for

- navigational purposes,
- kinematic surveying, and
- applications with satellites obscured by obstructions.

The necessary length of observation depends on the purpose of the survey, the instrument type, the desired accuracy, the software capacity, and logistic aspects. The basic requirement for precise surveys is the resolution of phase ambiguities. Once the ambiguities are resolved the observations can be finished. Over short distances (up to 10 km), with sufficient satellites (six or more), dual frequency receivers, and advanced software, this period can be as short as a few minutes or even less [7.3.2.3]. In kinematic surveying the few centimeter level can be achieved continuously (see [7.5.2]).

Over larger distances and under difficult environmental conditions (such as ionospheric disturbances, multipath), several hours of observations are required to obtain a precise ambiguity float solution [7.3.2.3]. For the establishment of national or continental fundamental networks, and for geodynamic purposes the observations can last 24 hours or even several days, to average out orbital, meteorological, multipath, and other time-variable effects.

Under difficult logistical conditions (e.g. in areas with difficult access) it is advisable to increase the usual observation time in order to avoid the need for re-occupation of sites in cases of poor data. Hence the following observation scenarios can be distinguished:

24 hours up to several days	fundamental networks, geodynamics,
several hours	highest accuracy over larger distances or under difficult conditions,
15 to 30 minutes	control surveys with short distances up to 10 km, and
continuous measurements	rapid methods and navigation.

With the installation of permanent networks and dense arrays of GPS receivers, continuous reference observations are available for many tasks. The number of specifically organized “GPS-projects” will decrease. Most observations in applied geodesy and surveying will be done with respect to existing reference stations or networks.

7.6.1.2 Practical Aspects in Field Observations

Advance local *reconnaissance* can be essential for successful observations. The observation sites should have unobstructed visibility and should be accessible to vehicles. As a general rule, a free line of sight down to the horizon is required in all directions.

In forested areas or near buildings, a satellite visibility diagram (sky plot, cf. Fig. 7.68) helps in the site selection. However, if the sites are to remain usable for later observations with other satellite constellations, it is recommended that the horizon be generally open, at least down to a 10° angle of elevation. Note that for precise

height determination observations down to 5° are advantageous [7.6.2.3]. Existing obstructions should be documented in the reconnaissance sheet in a *shadow diagram* (Fig. 7.70).

The GPS technique requires and permits selection criteria other than those of classical triangulation techniques. Control points no longer have to be installed on topographic elevations or towers with mutual station intervisibility, but rather where ever they are needed, on easily accessible sites with a minimally obstructed horizon. Also places near high buildings, towers, power lines, and transmitting antennas are not suitable. Nearby walls or other reflecting surfaces can cause multipath effects [7.4.4.3].

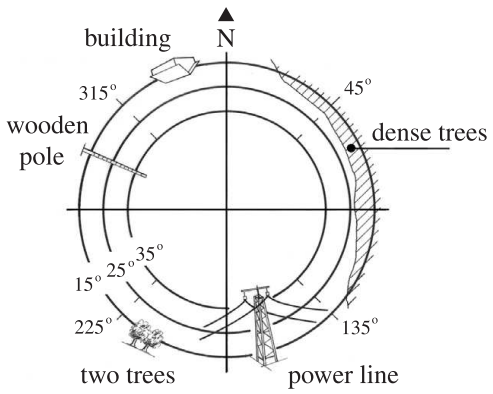

Figure 7.70. Shadow diagram, indicating obstructions with elevations $\geq 10°$

In wooded areas, the antennas can be mounted on light masts. In the case of non-centric observations, however, centering and plumbing must be done with the same accuracy with which the GPS measurements can be evaluated, that is centimeters to millimeters. Note that eccentricity calculations have to be done in the 3-D space (cf. [2.1.7]). In the interest of convenient future use for topographic and surveying purposes, points should be selected where centric observations can be made.

In many cases the follow-up surveying is done with conventional equipment, for example with electronic tacheometers. Points should be selected such that either a free sight is available to a nearby surveying mark, or an intervisible second GPS point has to be installed a few hundred meters away.

The *monumentation* of station marks usually follows the general rules of the responsible surveying and mapping authorities. Regarding the high accuracy potential of GPS the monuments should be established on stable ground, if possible on rock, or concrete blocks with sufficiently deep foundation. The station marker should be defined to at least 1 mm, for example with a fine grid mark on a corrosion-resistant metallic rivet. In such cases, the GPS stations can also be used for control purposes and engineering surveying. In addition, the station markers should be suitable as an exact vertical reference. The central survey marker is usually controlled by eccentric reference marks.

All essential information should be documented in a *reconnaissance sheet*. Possible elements are

- station name and identification code,
- description of site,
- approximate coordinates and height,
- accessibility (car, road conditions, walking distance),
- necessary antenna height (tripod, mast),

– orientation marks, and
– shadow diagram.

Power supply is no longer a major problem in practical field work as it was for older receiver types, such as the TI 4100 [7.2.4.1]. Modern instruments have a very low power consumption. Internal batteries usually last for a whole working day or even longer. For security reasons it is advisable to recharge batteries every other day.

Whereas older receivers could only be operated by skilled people, modern receivers work completely automatically. Dialogue with the receiver is possible, but not required in standard operation. Usually all visible satellites are tracked, and no pre-selection, or change of constellation, is necessary. The personnel should be able to carefully

– mount the tripod on the station mark,
– measure the antenna height,
– control the receiver operation,
– work according to a given time schedule,
– run the station control sheet (station log), and
– measure additional data if required (meteorological data, eccentric elements).

In most projects the measured GPS data have to be stored on suitable recording media for subsequent computation, e.g. for multistation adjustments. Modern receivers have built-in solid state memories, or plug-in memory cards. Depending on the memory capacity and the amount of data, the recorded measurements have to be downloaded to a computer once a day or at the end of a campaign; or the memory cards have to be exchanged. In larger field projects it is recommended that the data are transferred to suitable mass storage in the observation area, making a data check at the same time. Data can also be transferred via cellular phone and/or internet from the field to a central processing facility.

The amount of incoming data is enormous if the full data rate of modern receivers is exploited, i.e. once or twice per second. For most static applications a much lower data rate is completely sufficient, for example every 15 or 30 seconds. It is, however, essential that all receivers in one project sample at the same data rate. This condition may cause problems if different receiver types are used.

In some cases, the meteorological data are used in the subsequent multi-station evaluation. The data have to be recorded at adequate intervals, for example every 30 minutes: pressure (±1 mm), temperature (±1° C), and relative humidity (±1 %). However, note that meteorological station data can introduce biases into the multi-station adjustment because they are not representative for the working area.

It is advisable to keep a *station log*, entering not only the weather data but also the station identification code, the receiver and antenna identification numbers, the antenna position and height, the observation schedule, operation problems, and other significant information of relevance to future data processing.

7.6.1.3 ● Observation Strategies and Network Design ●

Three basic observation strategies can be distinguished (cf. [7.3.4]):

– point positioning concept (single receiver),

– baseline concept (relative observations at two stations), and
– multistation concept (three and more receivers operating simultaneously).

Various procedures can be selected within the last category. Of particular relevance are observations in connection with active multiple reference stations [7.5.3].

The choice of the observation concept depends on the objective of the survey, the required accuracy, the number and type of receivers available, and the logistic conditions. Hence a general classification is difficult and not appropriate. As regards the necessary accuracy, the following user classes may be defined, though the boundaries are debatable (Table 7.21).

Table 7.21. GPS user classes

	Category	Average required relative accuracy	Corresponding accuracy in [m], distance dependent
A:	Exploration geophysics Georeferencing low accuracy GIS	$1 \cdot 10^{-4}$	$1 \ldots 50$
B:	Topographic map surveys Small scale engineering Vehicle control systems	$1 \cdot 10^{-5}$	$0.2 \ldots 1$
C:	Cadastral surveys Engineering surveys of mean accuracy	$5 \ldots 1 \cdot 10^{-6}$	$0.01 \ldots 0.2$
D:	Geodesy, Control surveys High precision engineering surveys	$5 \cdot 10^{-7} \ldots 1 \cdot 10^{-6}$	$\leq 0.01 \ldots 0.05$
E:	Geodynamics Highest precision engineering surveys	$1 \cdot 10^{-7}$	$0.001 \ldots 0.02$

With a *single receiver*, an absolute position determination can be achieved continuously (navigation mode) with an accuracy of 5 to 15 m, without SA, under the Standard Positioning Service (SPS), cf. [7.1.6], [7.4.1]. Even after several hours of observation the achievable absolute accuracy is not better than several meters. Therefore, only group A activities can be undertaken with a single receiver. A new situation evolves with the use of precise ephemerides and clocks in the concept of *Precise Point Positioning* (PPP). In essence this is, however, an implicit form of differential GPS [7.3.4].

For all other user groups, only *relative observation techniques* with at least two simultaneously operating GPS receivers are worth considering. The terms *differential GPS* and *translocation observations* are also used equivalently.

The concept of relative observations is extensively discussed in chapter [7.5]. It applies for moving and static antennas. The essential strength of relative techniques lies in the fact that a part of the error influences at neighboring stations is strongly correlated and is therefore cancelled out when a difference is taken (cf. [6.5.3] for the TRANSIT system). This is especially true of orbit errors, errors of the satellite clock, and errors in the ionospheric modeling.

Comparison of Table 7.12 and Table 7.6 makes it clear that the systematic model errors and the observation noise of the code have more or less the same order of magnitude, namely 1 to 10 m. Hence, in the navigation mode, a relative navigational accuracy of ±2 to 3 m is successfully achieved using code phase measurements and corrections from a reference station (*differential GPS* [7.5.1]). For a static receiver, the extremely low observation noise of the carrier phase measurement, which is three to four orders of magnitude less than the systematic error effects, can be used to advantage only if the systematic components are eliminated by relative measurements.

In this way, an accuracy increase by a factor 10^3 to 10^4 is brought about in the geodetic relative mode with at least two simultaneously operating receivers, as compared with the single receiver mode. Relative techniques are particularly effective when the station distance is small compared with the satellite range (∼ 20 000 km). The amount of correlation decreases as the distance increases; however, the correlation is effective up to several thousand kilometers. The adjustment models for relative observations are discussed in [7.3.4] and [7.5].

● If two receivers are available, a point field or network can be set up by the observation of *baselines*. One possibility is to operate one instrument at a central station, and occupy the adjacent points in a star-shaped pattern (Fig. 7.71). Adjacent central stations A, B, C, ... are linked through baseline observations. The baselines between the non-simultaneously occupied stations can then be derived by computation. For control purposes, some of those "trivial" baselines (cf. [7.3.4], Fig. 7.42, p. 284) can be independently observed.

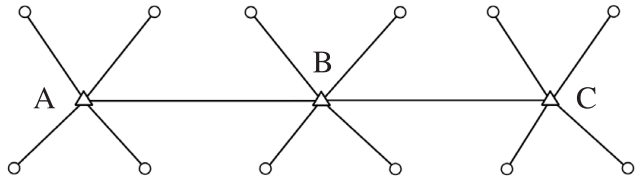

Figure 7.71. Baseline observations with two receivers

Another possibility is to occupy neighboring points and form triangles or quadrangles (Fig. 7.72). This method leads

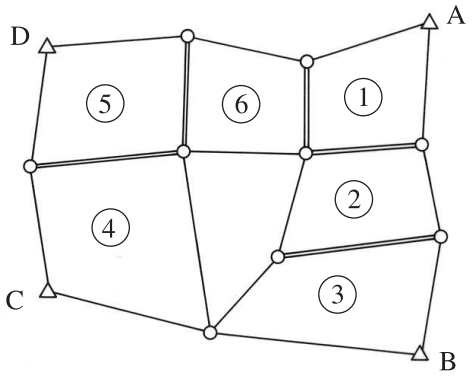

Figure 7.72. GPS network

to a high relative accuracy, in particular if the quadrangles in Fig. 7.72 are subdivided into triangles, but it is very laborious. It is clear that the use of more than two instruments is much more economical, even for small sized networks. The configuration in Fig. 7.71 can be worked with two instruments in 14 observation sessions, and with three instruments in seven sessions.

All observations made simultaneously during a given time period in the course of a GPS project are called a *session* (cf. [7.3.4]). Each session has to be connected to at least one other session of the network through one or more identical stations where observations have been carried out in both sessions. An increasing number of identical stations increases the stability, accuracy, and reliability of the total network.

- When three or more receivers are used in a multi-session project, the design of an observation plan becomes an optimization problem between efficiency (economy), accuracy, and reliability. Some basic considerations are discussed here. We define

  r number of simultaneously operating receivers,
  n number of stations,
  m number of stations with more than one observation in two different sessions, and
  s number of sessions.

We know already from (7.100) that

- $r(r-1)/2$ number of possible baselines in one session, and
  $(r-1)$ number of independent baselines in one session.
- The number of sessions required for a given network is:

$$s = \left[\frac{n-m}{r-m} \right], \tag{7.165}$$

with s being the next larger integer number. With two or more reoccupied stations in each session, some of the baselines are determined twice. In the total network we have

- $$\begin{aligned} s(r-1) &\quad \text{number of independent baselines, and} \\ (s-1)(m-1) &\quad \text{number of double determined independent baselines.} \end{aligned} \tag{7.166}$$

Let the network example in Fig. 7.72 be observed by four receivers with two connecting points between consecutive sessions. From (7.165) and (7.166) we find:

13 stations,
6 sessions,
5 double determined baselines, and
9 repeatedly determined stations.

A detailed inspection of Fig. 7.72 shows that four stations are observed once and seven stations are observed three times. From the economic point of view, a homogeneous distribution of reoccupations would be favored, because it provides an equally distributed redundancy with the least number of sessions required.

Some software packages include the possibility of executing simulation calculations with a given station and receiver configuration. The importance of these features

must not be overemphasized because network configuration is only one aspect of a GPS mission.

- As regards logistic and practical limitations, the choice of an observation strategy will often be guided by experience, with formal optimization criteria providing valuable aid. Since the accuracy of a local GPS network is only little dependent of the station distance, the design aspects are mainly governed by logistic, economic, and reliability factors. Some general rules from experience are that
  - each station should be occupied at least twice, under different conditions, to identify blunders,
  - neighboring stations should be occupied simultaneously because the ambiguity resolution works best over short distances,
  - for medium-sized projects the use of 4 to 10 receivers is a good compromise with respect to logistics, production rate, and reliability, and
  - a certain number of baselines should be observed twice for accuracy checks.

These rules are valid for independent projects. In active multiple reference station networks the situation is different, insofar as a new station is always determined by a single receiver with respect to the whole network [7.5.3].

Besides accuracy, the *reliability* of a GPS network is an important issue of network quality. Reliability means the ability of a network to self-check against blunders or systematic errors. Fig. 7.73 gives an example (Augath, 1988). Stations A and B are used

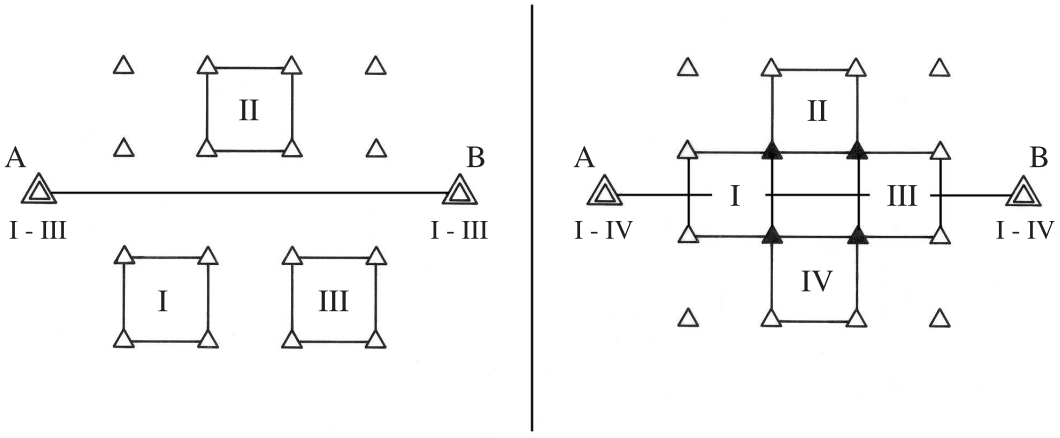

Figure 7.73. Network with accuracy criteria (left, three sessions) and reliability criteria (right, four sessions)

as reference points, and are occupied during all sessions. Four receivers are mobile. The left part of Fig. 7.73 demonstrates a network design with accuracy criteria only. The right part yields more or less the same accuracy, but in addition offers reliability, because each point (solid triangle) is used in two sessions. Additional constraints come through the permanent stations. An even more controlled network is shown with Fig. 7.74.

The data set which results from a GPS multi-station adjustment process has a high relative accuracy. The absolute coordinates, however, may have standard deviations of several meters because of the uncertainty in the realization of the satellite datum through observations (cf. [6.6.1] and [7.3.4]). As a rule, therefore, newly determined "GPS networks" must be tied to previously existing known points, either from the particular national control network, or from fundamental stations which are determined by precise global techniques such as VLBI, Laser, or GPS tracking networks. Examples of the latter group are

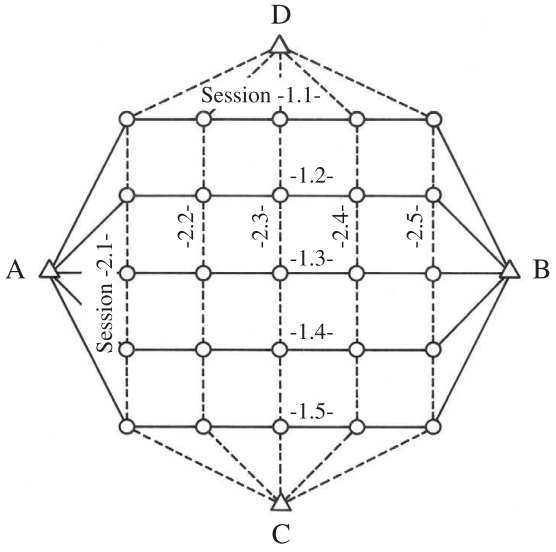

Figure 7.74. Generic network densification with GPS

- International Earth Rotation Service Terrestrial Reference Frame (ITRF) [2.1.2], [12.1.2],
- International Global Positioning System Service (IGS) [7.4.3], [7.8.1],
- National or regional GPS tracking networks like the Canadian ACS, the U.S. CORS, the Brazilian RBMC, or the German SAPOS [7.5.1], and
- Continental or national fundamental GPS networks like EUREF, SIRGAS, and DREF [7.6.2].

The tie can be made over one or several identical points, for example A, B, C, D in Fig. 7.72, or nearby permanent stations. For smaller working areas, a single connection point may be sufficient. The control points can be used as *fixed points* with minimum variances, or as *fiducial points* with a predefined, non-vanishing dispersion matrix. The network datum is derived from the pre-existing control points rather than from the GPS observations. In active reference networks the datum comes from the network datum. For detailed discussion see [12.1.1] and [7.6.2.1].

When planning GPS projects in remote areas, careful attention must be paid to connection to reference points with known precise geocentric coordinates. Otherwise, the errors in the absolute coordinates, inherent in the actual GPS observations, will propagate into the relative coordinates of the network solution [7.6.2]. A good solution to the problem is to connect new measurements with IGS stations.

With the evolution of sufficiently dense global, continental, and national fundamental networks, based on precise space techniques as well as on GPS, the reference point or fiducial point concept will be the technique usually applied when establishing GPS networks. In other words, GPS will be mainly used as an interpolation technique for network densification in the working area.

A final generic example (Fig. 7.74) highlights some of the essential items that have been discussed in this chapter. Stations A, B, C, D are points of the existing network,

either from previous GPS campaigns, or from a precise classical terrestrial network. They are used as fixed or fiducial points. The datum is completely defined through these stations. Seven receivers are applied. Two of them are operated on the fixed stations, and five are moving. Five sessions are observed each project section. During the first section the fixed stations A and B are occupied, and the roving receivers move in five sessions according to the solid lines. During the second section stations C, D are occupied, and the receivers move according to the dashed lines. Depending on the interstation distances, the receiver types, the available software, and the project objective, the individual sessions can last minutes, hours, or even days. Note that the following principles are fulfilled for the newly determined points:

- *high accuracy*, caused by a sufficiently long observation period in each session (following the project goals),
- *highly economic*, because the session number for double occupancy, and the interstation travel times are minimized, and
- *high reliability*, because each new point is derived from two completely independent determinations (new antenna installation), tied to different control points, and (mostly) observed under a different satellite constellation.

For observation strategies using active reference networks see [7.5.3] and [7.6.2.1].

7.6.2 Possible Applications and Examples of GPS Observations

Since GPS is an all-weather, real-time, continuously available, economic, and very precise positioning technique, almost unlimited possibilities are opened up for its use in geodesy, surveying, navigation, and related fields, including

- control surveys,
- geodynamics,
- altitude determination,
- cadastral surveying and GIS,
- monitoring and engineering,
- precision navigation,
- photogrammetry and remote sensing, and
- marine and glacial geodesy.

Some typical fields and examples of GPS application will be discussed in the following. The use of satellite methods is further reviewed in [12].

It was recognized early on that GPS is a multipurpose system. One major advantage is its capability of forming a powerful building block in integrated systems. GPS together with a coordinate system and geographic information produces a map. GPS together with a map facilitates navigation. GPS together with a digital geometric data base, a geographic information system (GIS), and a communication link produces a command and control system (Gibbons, 1991).

With the establishment of continuously operating reference stations covering a whole country [7.5], the acceptance of GPS as a basic positioning tool will further grow. The availability of position information in real-time at any level of required

accuracy and at any place will be taken for granted as today is the availability of precise time or of communication links.

Because of the fast growing application market, only some basic concepts are described here. For more information on the current discussion see journals like *GPS World* or symposia proceedings like *ION GPS*. The statements in this chapter refer to NAVSTAR GPS. They are, however, also valid for other GNSS systems like GLONASS or the forthcoming European GALILEO [7.7].

7.6.2.1 Geodetic Control Surveys

The following objectives can be identified:

(a) setting-up of a completely new field of control points,
(b) densification or extension of existing networks,
(c) inspection, analysis, and improvement of existing networks, and
(d) establishment of a network of active reference stations.

The terms "network" and "control point field" are used as synonyms.

(a) New network

The installation of a completely *new network* can be performed in three steps. Since all densification work will be done with GPS techniques, it is advisable to select a global geocentric datum compatible with the World Geodetic System WGS 84 [2.1.6]. WGS 84 is now defined with an accuracy level of about ±1 cm (Merrigan et al., 2002) and corresponds at that level with the *International Terrestrial Reference Frame* ITRF [2.1.2.2], [12.4.2]. Areas with an insufficient coverage of ITRF sites, for example Africa, or some parts of Asia (see Fig. 2.4) are densified by stations of the IGS service [7.4.3.2], [7.8.1] with the same accuracy standard. For most practical purposes, the global network ITRF2000 and the IGS network can be considered as equivalent.

Starting from the global network, three basic levels of "Geodetic GPS Networks" may be distinguished, all with the same high accuracy standard, namely about 1 cm:

– Level A: Continental (or Sub-Continental) Reference Frames,
– Level B: National Fundamental Networks, and
– Level C: All other GPS networks.

At *Level A*, a continental or sub-continental GPS network is installed, with the ITRF/IGS sites as fiducial points. The interstation distances are between 300 km and 500 km. The station coordinates have to be determined with the highest achievable accuracy, in general ±1 cm. This is possible with the fiducial point concept [7.4.3.2], about one week of observations, dual frequency receivers, precise orbits and advanced software.

As an example, see the EUREF (European Reference Frame) project. EUREF has been built up since 1989 by successive GPS campaigns. The existing ITRF stations in Europe (Laser and VLBI) were used as fiducial points. The first campaign was performed in May 1989 with about 60 dual-frequency receivers. In 1990, some 30 stations were added during the EUREF North campaign. After 1990, in several campaigns stations from Eastern Europe were included. The European Reference System

was defined as ETRS, in agreement with the ITRS, for the epoch 1989.0. Its realization is ETRF89 that coincides with the ITRF89 for stations in Europe. The basic idea is that ETRF89 rotates with the stable part of the European plate and hence can remain unchanged for a long time period. About 90 stations of the more than 200 EUREF sites form the permanent EUREF network [7.5.1.3] (Fig. 7.75), with the objective to maintain the ETRS and to densify the IGS network in Europe.

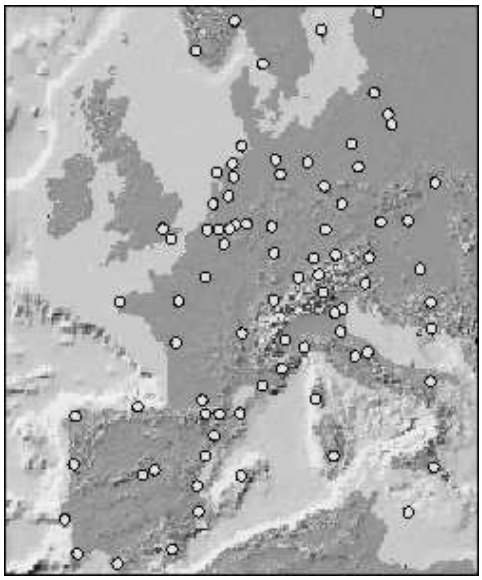

Figure 7.75. EUREF Permanent Network, source: BKG

Figure 7.76. SIRGAS 1995, source: DGFI

Similar basic reference frames have been or are being built up in other continents or subcontinents. In South America, the *Sistema de Referencia Geocentrico para America del Sur* (SIRGAS) was created in 1995 by 10 days of simultaneous GPS observations at nearly 60 stations (Fig. 7.76), Hoyer et al. (1998). The network was tied to ITRF94. Some of the stations continue as permanent stations, provide reference data, and maintain the frame. The data are processed in the IGS *Regional Network Associate Analysis Center* for SIRGAS (RNAAC SIR) [7.8.1] and contribute to a densification of the IGS global network. In 2000, the SIRGAS network was re-observed and enlarged including sites in Central and North America. In North America, the already mentioned U.S. CORS network [7.5.1] and the Canadian Active Control System (CACS) play a similar role. A continental network for Africa, AFREF, is under discussion.

At *level B*, nationwide, or statewide fundamental networks are installed with a spacing of 50 to 100 km, depending on the size of the country and the objectives. The stations of level A are kept fixed for use as fiducial points. The accuracy of the individual GPS station with respect to the neighboring stations is again ± 1 cm, hence providing a homogeneous set of coordinates for the whole country.

One example is the DREF campaign (Fig. 7.77) in Germany. DREF was observed early in 1991 with 83 dual frequency receivers. The network contains 109 stations with a mean spacing of 70 to 100 km. Some 20 stations are EUREF sites from level A.

It is advisable to use as many receivers as possible to provide a homogeneous set of observations. In most cases it will not be possible to occupy all stations of a national network simultaneously. The total network has then to be broken down into sub-networks and sessions. The single sub-networks and sessions are interconnected via fiducial stations (from level A), and by selected "identical" points at the rim of the individual sub-networks. Most countries have established fundamental networks of this type or will do so within the near future.

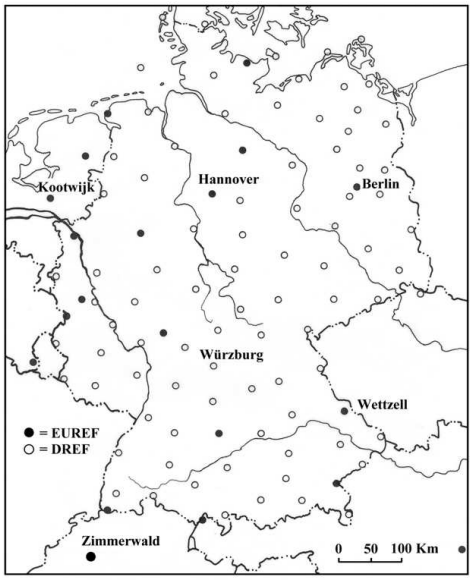

Figure 7.77. The DREF (German Reference Frame) network

Before using station coordinates from level A as a reference frame for densification at level B, the coordinates have to be corrected for crustal deformation, if applicable. Even small motions of say 2 cm/year will lead to a 10 cm deformation already after 5 years, which is not tolerable in precise geodetic networks. The procedure is as follows (Drewes, 1998). *Step 1*: Transformation of level A coordinates of the stations, S, used as connecting points (fiducials) from the epoch, t_0, of the reference frame (level A) to the epoch, t_i, of the new observations. Station velocities, $\boldsymbol{v}_S$, derived either from repeated observations or from crustal deformation models, are applied according to

$$\boldsymbol{X}_S(t_i) = \boldsymbol{X}_S(t_0) + \boldsymbol{v}_S(t_i - t_0). \tag{7.167}$$

Step 2: Network adjustment of the new stations, N, (level B) using the observations at epoch t_i together with the coordinates, $\boldsymbol{X}_S(t_i)$, of the fiducial points from level A. *Step 3*: Transformation of the new station coordinates, $\boldsymbol{X}_N$, from observation epoch, t_i, back to the epoch, t_0, of the reference frame (level A), using

$$\boldsymbol{X}_N(t_0) = \boldsymbol{X}_N(t_i) - \boldsymbol{v}_N(t_i - t_0). \tag{7.168}$$

This procedure ensures a homogeneous network of level B stations in the datum of level A. Since station velocities of the new stations are not always available, it is advisable to develop continuous deformation models for all continental plates [12.4.1].

At *level C*, all other control points have to be connected to stations of level B, again at the 1 cm accuracy level. One advantage, when compared with classical techniques, is that no systematic densification is necessary. Work can be done, following a priority schedule, where coordinates are required. The densification procedure can follow the scheme of Fig. 7.74. The classical division into geodetic networks of 1st to 4th order, within a country, will disappear, and be mostly replaced by two levels:

– the fundamental national reference frame (level B), and
– all other control points (level C).

(b) Densification of an existing network
This can be treated in different ways.

(1) A precise classical terrestrial network of 2nd or 3rd order exists. In this case GPS is used as a modern surveying tool for precise network densification. GPS is in nearly all cases much more economical than the classical methods. The procedure is as outlined in Fig. 7.74, see [7.6.1.3]. The existing control points are taken as fixed reference points. The existing national datum is maintained. This approach is used in many countries as a tool for rapidly providing precise geodetic control.

(2) A terrestrial network of medium or low accuracy exists; the old coordinates shall be maintained. In this case the distortion of the traditional network is introduced into the precise GPS results. GPS is only used as a method of cost-effective interpolation into the existing national framework. The GPS observations should be preserved for a rigorous adjustment once a fundamental GPS network has been established in the area at some later date. This procedure is acceptable as an intermediate solution, in particular in developing countries, until a completely new network and datum, based on satellite techniques, can be established.

(3) The existing terrestrial network is combined with new GPS observations. In this case the existing network datum is maintained; however the complete network is readjusted and strengthened with the inclusion of GPS measurements. New points are linked to the existing network in an optimal way. All network coordinates are slightly changed. The method only works if sufficient stochastic information on the existing network is available, see [12.1] and related literature, (Leick, 1995; Strang, Borre, 1997).

A particular problem arises when multiple reference stations [7.5.3.2] are established in an area with existing geodetic control. Even if the traditional network is of highest quality, discrepancies at the several centimeter level have to be expected when "distortion-free" new GPS points, derived from references stations at about 30 to 50 km distance, are established in the direct neighborhood of existing "distorted" surveying points. Such discrepancies are often not acceptable in cadaster or engineering projects. Two solutions are possible:

In a first step, for the whole area, local transformation parameters have to be derived from GPS observations at a sufficient number of existing control points. These parameters are either used to transform all existing surveying points into the distortion-free reference frame defined by the GPS reference stations, or they are used to transform the GPS determined coordinates of new object points into the existing distorted local frame, realized through the conventional surveying points. In the latter case the GPS results should be maintained in order to use them for a new coordination as soon as the former solution can be realized.

(c) Analysis of an existing network.
This procedure is of particular importance in countries where little information on the original observation and computation is available, for example in developing countries

(e.g. Campos et al., 1989). The analysis, however, also offers a very important insight into the present official networks of countries with an advanced cartographic tradition, such as Germany. For the analysis a certain number of existing stations is re-occupied with GPS. The residuals, after a seven-parameter Helmert transformation (2.46) are inspected.

Fig. 7.78 shows residual vectors between an early GPS campaign in Germany (DÖNAV) and the official terrestrial network DHDN (Seeber, et al., 1987). The residuals reach up to 1 m. A similar analysis is used to derive detailed expressions for a transformation formula between the datum of the GPS network and the existing local network.

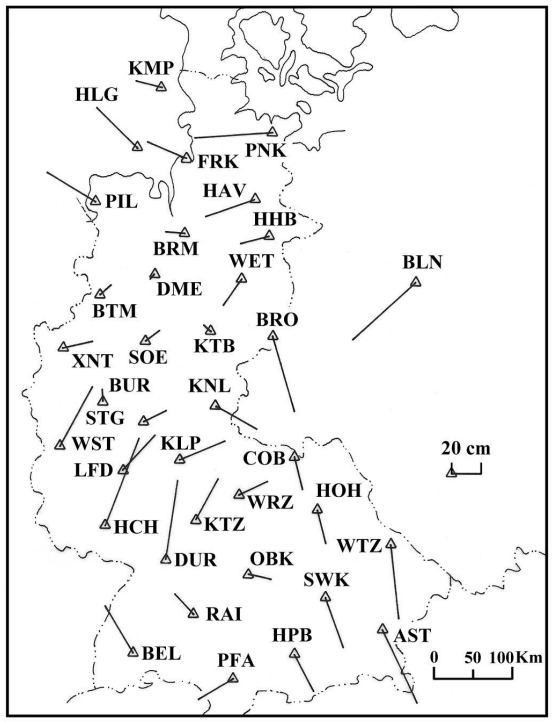

Figure 7.78. Residuals after a 7-parameter transformation between the DÖNAV network and the classical German network DHDN

(d) Active Reference Stations.

A modern tendency is to represent the fundamental reference frame in a country by a network of *active control points* that provide relative information, for any authorized user, on a routine basis (see [7.5.3.2]). This service can be operated under the responsibility of the national surveying authorities (for example SAPOS in Germany). The long-term rationale behind this concept is to substitute the reference frame exclusively through the active reference stations and to considerably decrease the number of monumented points.

Fig. 7.79 shows some possible concepts of active reference station networks at level C. In version (a) GPS data are collected at the reference stations and distributed to users via a control station. In option (b) range corrections in the RTCM 2.0 or RTCM 2.1 format are broadcasted to users from the nearest reference stations. In

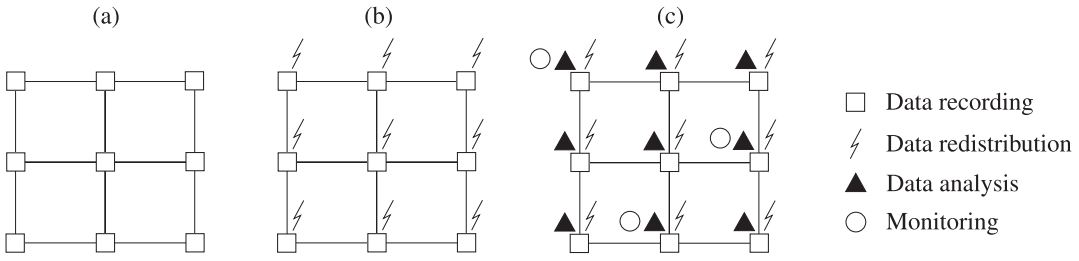

Figure 7.79. Different concepts of permanent reference stations

version (c) all stations are interconnected and work as monitor stations and analysis stations; they all transmit range corrections and ACP to the users [7.5.3.2]. The subject of network densification no longer arises.

7.6.2.2 Geodynamics

The very high accuracy potential associated with comparatively easily transportable equipment makes GPS a suitable technique for determining recent crustal movements [12.4.1]. Until about 1985, crustal movements were mainly analyzed with Very Long Baseline Interferometry (VLBI) [11.1] and satellite laser ranging (SLR) [8]. With VLBI, long-range baselines can be determined precisely; a few millimeters accuracy and precision over several thousand kilometers are achievable.

The main disadvantage of the VLBI method is the enormous technical expenditure and the limitation to a comparatively small number of fundamental stations; only very few transportable systems are available. With satellite laser instruments, very precise and reliable movement rates have been derived from many years' observation, for example, in the area of the San Andreas fault (Watkins et al., 1990), or along with the WEGENER/MEDLAS project in the Mediterranean region (Ambrosius et al., 1991). Transportable satellite laser ranging systems are also in use [8.3.3]; still, the use of SLR technology involves high costs and long mobilization times. For many areas of interest, in particular if a large number of points are to be determined for higher spatial resolution, GPS offers considerable advantages. This is why since about the late 1980s, besides VLBI and SLR, GPS is the technology of preference for the operational determination of crustal deformation and global plate motion.

In the early days of GPS one of the most important limiting factors in the error budget for precise baseline determination over large distances was orbit accuracy. Following the rule of thumb (7.134) an orbit error of about 2.5 m would propagate 1 cm error per 100 km into the baseline. In view of the known motion rates of a few cm/year or only mm/year, station spacing should then not be much greater than 100 km. With today's orbit accuracy of 5 cm or better for IGS products [7.4.3.2], the orbit is no longer a critical factor in crustal motion studies, even over large distances. Key factors of the error budget are rather [7.4.4][7.4.5]:

- modeling of atmospheric propagation effects,
- antenna phase center variations (PCV), and
- multipath effects.

Much research has been invested in recent years into the modeling of tropospheric and ionospheric propagation effects [2.3.3]. The use of data from GPS observations in LEO (GPS-MET, see [7.6.2.9]) projects will further help to improve the models. In addition an attempt can be made to raise the accuracy level through the use of *water vapor radiometers* [2.3.3.2], [7.4.4.2].

Tectonically active areas near the geomagnetic equator or in high latitudes will experience large ionospheric disturbances [7.4.4.1], e.g. Wanninger, Jahn (1991), Völksen (2000). The use of dual-frequency receivers is hence essential. Long observation periods, over at least 24 hours, help to average out residual effects.

Site dependent effects can be minimized with absolutely calibrated antennas and multipath reducing observation techniques [7.4.5.1]. Even for identical antennas, the PCV variation will not be cancelled in relative observations over very long baselines, because of Earth's curvature (Menge, Seeber, 2000).

Two strategies are being used to determine station velocities:

(i) repeated observations within dedicated campaigns, and
(ii) continuous observations at permanent installations.

Strategy (i) was mainly used during the development phase of GPS, and it is still applied for smaller independent projects or in remote areas with difficult access. A first epoch measurement establishes a network of well demarcated stations, and repeated epoch observations are performed after one or several years. A typical example is given with the Iceland campaigns in Fig. 7.80.

With the availability of fully automatic, low power consumption GPS receivers and the possibility to transfer data over large distances, strategy (ii) is more and more applied. One main advantage, compared with option (i), is, that data are continuously available and sudden events, like displacements due to earthquakes, can be directly analyzed. Two eminent examples are the IGS network and the GEONET in Japan [7.5.1.3].

The following main fields of application for crustal motion monitoring can be identified:

(a) global and continental plate motion and deformation analysis,
(b) regional crustal motion analysis, and
(c) local monitoring of deformation and subsidence.

Projects of *group (a)* show very impressive results after a couple of years of observations. Comparisons between GPS and other space techniques like VLBI and SLR demonstrate an agreement at the centimeter-level, and hence prove the capability of GPS for global geodynamics (Boucher et al., 1999).

A major break-through came with the establishment of the IGS [7.8.1]. More than 300 globally distributed stations deliver data on a permanent basis and as such provide a continuous monitor of deformation. The station velocities can be used to compute global stress maps and to determine a kinematic model of the individual plate rotation vectors (see [12.4.1], Fig. 12.13, p. 529, Tab. 12.3, p. 528).

Two examples are given for continental projects. The motion of the Antarctic plate was determined with two epoch measurements in 1995 and 1998 (Dietrich et al., 2001). Three weeks of observations, each time at about 45 stations on the Antarctic continent and the adjacent tectonic plates, were taken to establish a precise reference network linked to the ITRF96 reference frame, and to determine, besides of local deformations, the rotation of the Antarctic plate. Based on a data analysis with four different software packages at seven analysis centers, the combined solution yields an accuracy of 1 cm for the horizontal and 2 cm for the height components. For details see Dietrich (ed.) (2000). With horizontal velocities of about 2 to 3 cm per year an epoch difference of three years gives reliable results. For the detection of height changes the situation is more critical. A longer time span, and even more sophisticated modeling is required (see [7.6.2.3]).

The SIRGAS network in South America was observed in 1995 with 58 stations and again in 2000 with, in total, 184 stations. The results from the repeated stations are used to derive their velocity vectors. This information is also of high importance for follow-up geodetic work because SIRGAS 95 was adopted as a national datum by some of the participating countries. About 20 of the SIRGAS stations deliver data on a continuous basis. These data are included in the data set of the Regional IGS Network RNAAC SIR and continuously provide information on the motion of the South American Plate (DGFI, 2001).

An extremely challenging endeavour in this project group is the connection of continental control points with submarine control points near plate boundaries or subduction zones, because GPS measurements on floating platforms have to be integrated with underwater acoustic measurements [12.3.2], (Chadwell et al., 1998).

Projects of *group (b)* already show significant results. Investigations and epoch or continuous measurements have been started in nearly all tectonically active parts of the world. Well known examples are, among many others control networks in California, the CASA (Central and South America) and SAGA (South American Geodynamic Activities) GPS project, the GEODYSSEA (Geodynamics of South and South-East-Asia) project (Wilson et al., 1998), projects in the Mediterranean area (Kaniuth et al., 2001), and the neo-volcanic rifting zone in Iceland. Usually, displacement vectors are derived from the comparison of two or more epoch measurements if no continuous measurements are available. Fig. 7.80 shows the results derived from two early epoch measurements in 1987 and 1990 in the Northern Volcanic Zone of Iceland. About 50 stations were controlled with seven TI 4100 dual frequency P-code receivers. The epoch accuracy of adjacent stations is about 1 to 2 cm. The identified displacements in a post-rifting period reach about 3 to 5 cm/year. Subsequent epoch measurements in 1992, 1993, and 1995 provided a deeper insight into the mechanisms and enabled geophysical modeling and interpretation (Hofton, Foulger, 1996; Völksen, Seeber, 1998). Fig. 7.81 shows deformations after the $M = 8.1$ Antofagasta Earthquake on July 30, 1995 (Klotz et al., 1996).

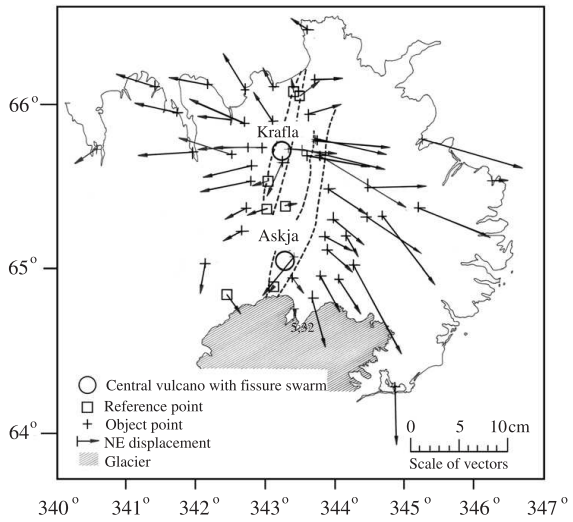

Figure 7.80. Displacement vectors from two consecutive epoch measurements in Iceland 1987–1990

One major difficulty in the analysis of a displacement field is the identification of "stable" reference points. Powerful methods have been developed in the field of network deformation analysis to address this problem (e.g. Mayer et al. (2000);

Niemeier et al. (2000)). One effective procedure is to relate all epoch measurements to ITRF. In order to demonstrate the local deformation behavior it can be helpful to select stations in the center of the deformation field, e.g. Fig. 7.80, Völksen (2000).

In areas of high risk (e.g. of earthquake, volcanic activities) like the San Andreas Fault in California, or in Japan, continuously monitoring GPS arrays have been installed (Bock et al., 1997). A fixed network of GPS receivers tracks all GPS satellites 24 hours a day. The data from all sites are transmitted via high-speed communication lines to the central facility, and are analyzed to obtain accurate "snapshots" of the relative positions of the network stations. Significant variations in these positions may indicate deformation caused by seismic or volcanic pre-event, co-event, or post-event activities.

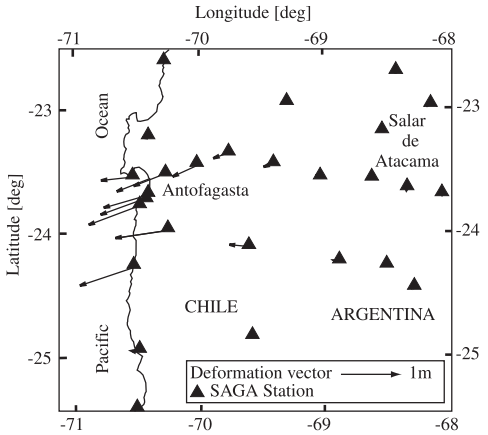

Figure 7.81. Deformation after the $M = 8.1$ Antofagasta Earthquake July 30, 1995, after Klotz et al. (1996)

Projects of *group (c)*, i.e. the monitoring of local deformation, belong in most cases to the field of deformation analysis in engineering surveying. Possible applications are the monitoring of

- land subsidence, e.g. in mining areas and oil fields,
- hang sliding, and
- local geotectonics.

In most cases the point distances are very small (about 1km), hence an accuracy of a few millimeters can be achieved, and very small deformations can be detected. Depending on the objectives of the control, and the expected rate of motion, the measurements have to be repeated after a given time period, for example days, weeks, or months. At least one stable reference station is required. In many cases, rapid methods can be applied [7.3.5]. In future, more and more continuously monitoring arrays will be built up. The data of the remote operating receivers have to be transmitted to the central station via cable, radio data link, or the internet.

A rather new and very promising field of GPS application in geodynamics is *Earth orientation monitoring*, in particular the variation of LOD and polar motion [2.1.2], [12.4.2]. Error analysis and comparison with other space techniques demonstrate the high potential of GPS to monitor daily and subdaily variations at the accuracy level of a few millimeters. Earth rotation monitoring, together with the delivery of precise orbits and station coordinates, is one of the major objectives of the *International Geodynamics GPS Service* (IGS) [7.4.3], [7.8.1].

7.6.2.3 Height Determination

GPS, like most other geodetic space techniques, yields three-dimensional coordinates X, Y, Z that can be transformed into ellipsoidal longitude, λ, latitude, φ, and height, h, see (2.37, p. 24). The vertical component is particularly sensitive to the geometrical configuration of the GPS satellites and to unmodeled errors in atmospheric refraction (cf. [7.4.4], e.g. Santerre (1991)). Simulation studies and experiences show that the error in the vertical component is about twice as high as in the horizontal components (Görres, 1996). Nevertheless, with improved receiver technology and careful data modeling, GPS is a powerful means for rapid and precise height determination. Considering the error budget, primarily height differences are of interest. For an elementary introduction into the GPS altimetry problem see also Schwarz, Sideris (1993).

It is obvious that the ellipsoidal height, h, coming from a GPS solution, is a purely geometric quantity. For most practical purposes, heights related to the gravity field rather than to the ellipsoid are required, namely *orthometric heights* or *normal heights*; for details on height systems see e.g. Torge (2001), also [2.1.5]. Note that orthometric heights are defined with respect to the geoid whereas normal heights refer to the quasigeoid. In the following no difference is made between orthometric and normal heights. The relation between

h ellipsoidal height from GPS observations,
H orthometric or normal height from spirit leveling, and
N geoid height from a geoid computation,

can be written, according to Fig. 7.82, as

$$h_1 = N_1 + H_1; \quad h_2 = N_2 + H_2,$$
$$\Delta H = H_2 - H_1; \quad \Delta h = h_2 - h_1; \quad \Delta N = N_2 - N_1,$$
$$\Delta H = \Delta h - \Delta N, \quad \Delta N = \Delta h - \Delta H, \quad \Delta h = \Delta H - \Delta N. \tag{7.169}$$

If two types of information are known, the third one can be determined, namely

- with precise geoidal heights the orthometric or normal heights can be derived from GPS, in order to control or to substitute spirit levelling, and
- with precise levelling information and ellipsoidal heights from GPS, the geoid can be determined or controlled.

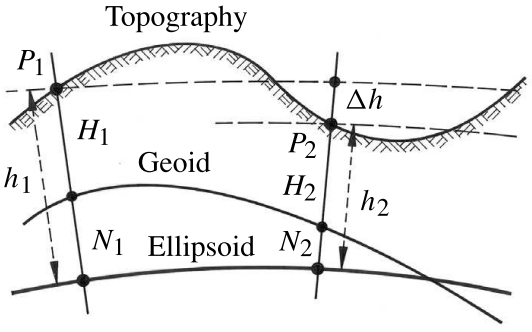

Figure 7.82. Relation between geoid height, N, orthometric height, H, and ellipsoidal height, h

If only height changes have to be analyzed the repeated determination of GPS heights without reference to the geoid is completely sufficient. Hence, three basic applications of GPS can be identified:

(a) height changes from repeated GPS control,
(b) transfer of orthometric or normal heights with known geoid, and
(c) determination of the geoid.

Height changes are mainly of interest for engineering purposes [7.6.2.6] or the detection of vertical crustal movements [7.6.2.2]. Particular engineering applications can be seen in the monitoring of subsidences in mining areas or for offshore drilling platforms. Vertical crustal movements are also of interest in connection with tide gauges to control sea level rise (Liebsch, 1996) [12.3.1].

The *determination of orthometric heights* (or normal heights) with GPS is a long-term goal in surveying and geodesy, in order to substitute time consuming and expensive spirit leveling. Fig. 7.82 and equation (7.169) demonstrate that we need two quantities for a height transfer from P_1 to P_2: the ellipsoidal height difference, Δh, determined with GPS, and the geoidal difference, ΔN, stemming from a geoid model.

In view of the high density of global and regional GPS reference networks [7.6.2.1] ellipsoidal height information is available at the centimeter-accuracy level over distances of several hundred kilometers or, in some areas, also several tens of kilometers. As a consequence, only ellipsoidal height differences over comparatively short distances have to be determined. This can be achieved with cm-accuracy by applying careful error modeling. For distances up to several tens of kilometers, even subcentimeter accuracy can be obtained (Görres, Campbell, 1998). The main limitations for precise GPS altimetry obviously come from

– modeling of the vertical error budget, and
– the requirement for precise geoid information.

The *vertical error budget* is mainly governed by the tropospheric propagation delay and antenna phase center variations (PCV) [7.4]. PCV, delay and the height component are strongly correlated with each other. Absolute PCV can be determined through calibration. The tropospheric delay can be separated from the height component, when a tropospheric scale bias is estimated from several hours of observations, including measurements at low elevation angles (Rothacher et al., 1998). Nonetheless, the troposphere remains a critical factor when the highest accuracy in GPS altimetry is required (Kaniuth et al., 1998).

Local and regional geoid models reach the accuracy level of a few centimeters, as can be verified by comparison with GPS levelling. For Europe, the EGG97 currently gives the best solution (Denker, Torge, 1998). For the area of the United States the GEOID 96 (Smith, Milbert, 1999) gives a similar accuracy level. In general, however, for worldwide applications, the available geoid information is still not satisfying due to the lack of data. In order to exploit the potential of GPS for altimetry it is necessary to improve knowledge of the geoid. Here, the current and forthcoming gravity field missions CHAMP, GRACE, and GOCE [10] will contribute considerably, in particular for the geoid's long wavelength components, (see also [12.2]).

For the time being, local solutions and approximation techniques have to be applied, for example the use of mathematical interpolation algorithms between GPS stations with known leveled heights. In particular, for small areas with a good coverage of

control points, the method delivers satisfying results (Zhang, 2000). Very good results have also been obtained with the use of finite elements to represent a height reference surface (Jäger, Schneid, 2002).

Where heights in the gravity field are known from levelling lines it is possible to directly derive *geoid heights* from GPS results. This method can contribute considerably to the determination of a precise geoid. Other major problems to be solved with GPS altimetry are the connection of separated tide gauges, e.g. Kakkuri (1995), Liebsch (1996), and the establishment of a *global height datum*. This includes the determination of a precise marine geoid and of the *sea surface topography* [9.5.1]. In coastal areas, a precise geoid strongly supports the height determination for near-shore engineering and shore protection activities (Seeber et al., 1997b).

Very precise geoid profiles can be determined with a transportable digital zenith camera using the concept described in [5.2], in combination with GPS. The camera provides the direction of the plumbline in near real-time, and the GPS receiver generates geodetic coordinates as well as precise time. Using the technique of *astronomical levelling* (Torge, 2001), a high resolution geoid profile and orthometric heights are provided on-line (Hirt, 2001).

7.6.2.4 Cadastral Surveying, Geographic Information Systems

Because of the high accuracy in connection with short observation time, GPS can also be employed economically for *detailed surveying* in rural or urban environments. Main fields of applications are in connection with the installation or maintenance of multi-purpose cadaster or geographic information systems.

One major problem in detailed surveying is signal shadow caused by buildings, trees, towers, bridges etc. This is why the exclusive use of GPS in cadastral surveying will be restricted to open areas. With the presence of such obstructions, GPS will be mainly used to determine rapidly the standpoints for electronic tacheometers or other conventional surveying instruments. Fig. 7.83 illustrates the situation.

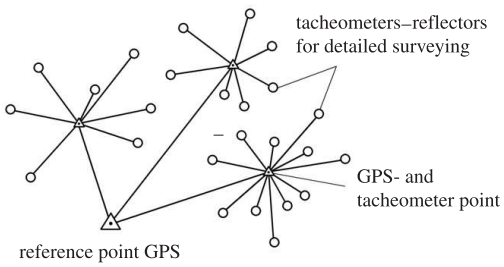

Figure 7.83. Combination of GPS with an electronic tacheometer

In areas of free sight, like most rural areas or urban areas with broad streets, low buildings, and low vegetation, rapid GPS methods can be used [7.3.5], in particular the RTK technique [7.3.5.4], [7.5.2]. Fig. 7.84 gives an artist's view of a detailed survey with GPS. The data can be stored in the moving receiver, or transmitted via a data link to the reference receiver, or vice versa.

With a continuously working data link, the setting-out of coordinates, or a re-identification of existing points or lost monuments will also be possible. The precise coordinates of the moving antenna are calculated in the field in real-time, and it is

indicated to the surveyor how far the antenna has to be moved to the final destination. Integrated systems of this type are available from most major GPS manufacturers.

The procedure depicted in Fig. 7.84 can be realized with a local, temporarily established reference station (case (a)), or with respect to a continuously operating reference station (case (b)).

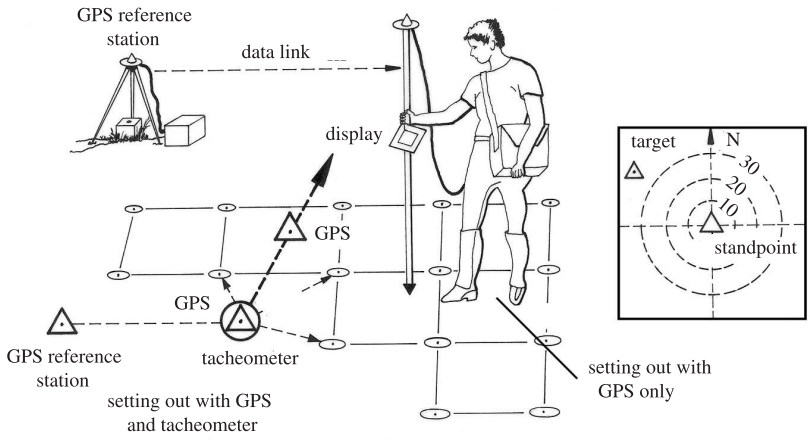

Figure 7.84. Use of GPS in real-time detailed surveying

Case (a) can be realized with conventional RTK equipment consisting of two GPS receivers and a radio. The reference receiver has to be installed on an existing demarcated surveying point, or the coordinates of the reference station have to be determined with respect to existing stations in the neighborhood. This can be realized when the roving receiver occupies two or three of such stations along with the survey. In modern surveying concepts it is no longer necessary to demarcate the temporary reference stations because the local field of surveying points is only represented by a strongly limited number of demarcated stations.

Case (b) has the advantage that only one GPS receiver is required in the field. Again, in most cases GPS will be used to establish standpoints for a tacheometer, whereas the object points (boundary marks or house-corners) are determined with conventional surveying tools. Another advantage is that all coordinates are immediately given with respect to the official reference frame and that no additional time is needed for the reconstruction of existing surveying marks. For high accuracy requirements it is necessary to work with networked reference stations [7.5.3.2]. For reliability purposes it is advised to occupy each object point twice.

GPS is a powerful means to support *Geographic Information Systems* (GIS). The role of GPS in this context is manyfold:

- it contributes to a uniform basic geometric frame, for example a coordinate system, a digital map, or a digital terrain model,
- it contributes to the geometric location of objects that enter the GIS, for example streets, buildings, power lines, proprietary boundaries,

– it allows the GIS to be taken out into the field with GPS direct-entry, and
– it forms an integrated building-block in a command and control system, for example for moving vehicles or machines that are navigating based on a digital terrain model.

In the following, only some examples are given. For all enterprises that provide services like energy, water supply or traffic information a geographic information system forms the basis of most decisions. As a first step, all spatially related data and object data have to be collected. Traditional maps are in many cases not sufficient. Here GPS provides an economic and efficient tool for an automatic data flow into the GIS. Vice versa, all objects that are selected in a GIS can be immediately identified in the field (e.g. Barrett, 1997). Integrated GIS - GPS concepts are offered by many manufacturers. The market is rapidly growing. Application examples are inventories for pipelines, power lines, fresh and waste water, streets, traffic signs, railway tracks, trees, contaminated locations, and so on.

Depending on accuracy requirements, GPS provides continuous position information at all scales of interest. In some cases, the accuracy of a single receiver (5 to 15 m) is sufficient. In most cases, ordinary DGPS will be applied (0.5 to 2 m). If highest accuracy is required (few centimeters), the services of multiple reference stations can be used [7.5.3], or even established for the purpose. Another advantage is that 3 D information is available. In connection with a digital geoid, gravity field related height information (e.g. orthometric heights) can be supplied for applications involving the direction of water flow.

Rapid digital data acquisition is possible with a *car driven survey system* for mobile mapping (Fig. 7.85). The positioning problem is solved by GPS in connection with an inertial sensor, or alternatively, wheel sensors, barometer and magnetic sensors. The data are acquired and analyzed automatically with several video cameras (Benning, Aussems, 1998; El-Sheimy, 2000).

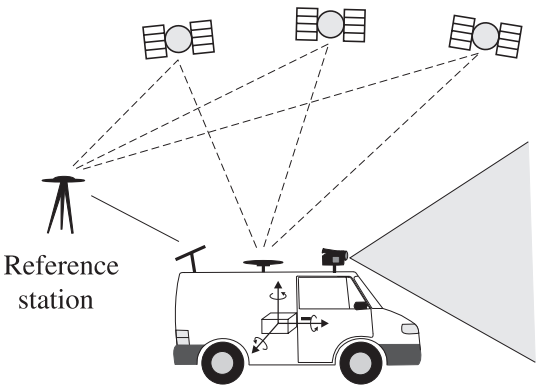

Figure 7.85. Car driven survey system

Another fast growing field of application is *precision farming*. Based on a GIS including the topography, soil quality, and actual state data, all steps in farming can be performed in an optimized way, like fine-tuned fertilization or spraying of infested areas. Computerized controllers and GPS-guided navigation form an optional part of farming equipment like sprayers or harvesters. Table 7.22 gives an overview of accuracy requirements (Demmel, 2000).

Many more examples could be given. The integration of GPS and GIS together with a communication link is increasing and widely discussed in the GPS literature (e.g. *GPS World*), as well as in the general surveying and GIS literature.

Table 7.22. Accuracy requirements for the use of GPS in precision farming

Task	Example	Required accuracy
Navigation	Search working area Search deposit place	± 10 m
Execution of work Information Documentation	Work in the field with – determination of returns – fertilization – plant protection – soil samples Automatic data recording	± 1 m
Vehicle guidance	Connected tracks Harvester-thresher	± 10 cm
Equipment guidance	Mechanical weed removal	± 1 cm

7.6.2.5 Fleet Management, Telematics, Location Based Services

These services present important new challenges, with a focus on real-time positioning, communication and information. They are mainly related to motorized vehicles like cars but also may concern pedestrians. The denomination is not yet clearly defined; all three terms are sometimes used for the same service. *Fleet management* means the control of a large quantity of vehicles like trucks, trains, police- and emergency cars, public buses and so on. *Telematics* is a new word composed from "telecommunication" and "informatics" and means the use of traffic-related information. *Location Based Services* (LBS) are mainly related to the use of cellular phones and mean the real-time availability of all kinds of position-related information to individual customers. The backbone of all services is composed of these elements:

- knowledge of the position of the client,
- knowledge of the position of other participants in the system, if required,
- a geographical information system,
- a personal digital assistant (PDA, palmtop) with the client, or a computer in the control center, and
- a communication link.

The positions can either be provided by GPS (GNSS), or another positioning device like the cellular phone identification. The GIS is either available in the PDA of the client, or via cellular phone from a provider. Considering the rapid development of communication technology, and the high number of vehicles in industrial countries, the market promises to develop fast. Some examples follow.

Car navigation systems for individual users, based on a digital map and a location service, are already well established. The inclusion of information on congestion, snowfall, or roadwork, for instance, will improve the service. Additional features are automatic location transmission in case of emergency, or theft.

Fleet management is essential for shipping agencies, train and bus systems, police and emergency services, and fire brigades. In connection with a traffic management system, traffic light priority can be given to public transportation and emergency vehicles.

At large construction sites a logistic system can be installed to organize and guide the construction vehicle traffic. Each vehicle gets a certain time slot when entering the site, and a GPS based local navigation equipment is deployed in each car as long as it operates inside the construction area.

A particular application is the mobility of blind people. A precise DGPS system and a precise and detailed specific digital map, connected to a voice generator enables a user to navigate in an unknown environment aided perhaps only by a stick.

A large potential market is developing for location based services. Tourists can request information on nearby touristic highlights, restaurants and public transportation. Parents can supervise their children, and persons with a critical health status can be remotely monitored by a medical center.

A particular application will be the automatic location of a mobile phone in connection with the emergency calls E-911 in the U.S., or E-112 in Europe. A further step will be a combination of outdoor and indoor navigation within a single hybrid location device.

7.6.2.6 Engineering and Monitoring

Almost unlimited possible uses and applications may be conceived in this field. The corresponding observation and evaluation methods are as discussed in the previous sections. Since the distances are usually small it is possible to achieve mm accuracy with routine methods. Rapid methods [7.3.5], real-time solutions, and integration with electronic tacheometers may be required. Some fields of application are:

(1) *Determination of geodetic control points*
- Geographic Information Systems (GIS),
- cartography,
- photogrammetry,
- geophysical surveys,
- inertial surveys,
- antenna location in hydrographic surveying,
- expeditions of all kinds, and
- archaeological mapping,

(2) *Monitoring object movements by repeated or continuous measurements*
- ground subsidence (mining, ground water withdrawal),
- land slides,
- construction of dams,
- subsidence of offshore structures, and
- settlement of buildings,

(3) *Setting out local networks for the control of engineering projects*
- tunnel construction,

– particle accelerators,
– bridge construction,
– road construction,
– pipelines, and
– waterways,

(4) *Real-time guidance and control of vehicles*

– construction vehicles,
– large excavators in opencast mining, and
– forklifts in open storage areas (e.g. container yards).

If two antennas (and receivers units) are used, GPS can also be employed as a method of determining directions. Usually the direction is derived from the coordinates of the two antenna phase center positions, hence precise carrier phase resolution and carefully designed and calibrated antennas are required. Table 7.23 shows the relation between station spacing, azimuth accuracy and required GPS relative position accuracy. If 2 mm relative position accuracy is considered to be the accuracy limit, it is possible to determine a 1 arcsecond azimuth over 400 m distance. This may be of interest for setting out a tunnel axis.

Table 7.23. Azimuth reference control with GPS

	azimuth accuracy in seconds of arc				
	1	2	4	6	10
station spacing (m)	GPS relative position accuracy in mm				
100	–	1	2	3	5
200	1	2	4	5	10
300	2	3	6	9	14
400	2	4	8	12	19
500	3	5	10	14	24
600	3	6	12	18	29

For operational use, a much shorter baseline can be selected. With a 1 m antenna separation a directional accuracy of a few arcminutes can be achieved, even in kinematic mode. GPS can hence be used for compassing. With three antennas the attitude of a moving platform can be controlled.

From the above list of possible applications two examples are given, a control network for tunnel construction and a network for dam control. The advantage of GPS can, in particular, be demonstrated for the *tunnel network*. The main purpose of such a network is the setting-out of the bearing of the shaft center line at both entrances, P_W (portal west), and P_E (portal east), cf. Fig. 7.86. In classical engineering, both portals had to be connected via a precise network, covering the whole area. This could be an extremely difficult task in mountainous or heavily forested areas. With GPS it is sufficient to determine two control points each at both entrances for setting out the

bearing of the center line. For security reasons, it is advisable to establish a second target pillar at each portal for reference bearings. The distance should not be too large to enable sights under unfavorable atmospheric conditions.

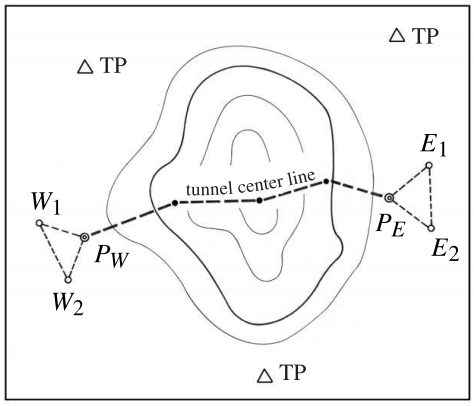

Figure 7.86. Generic tunnel network with GPS

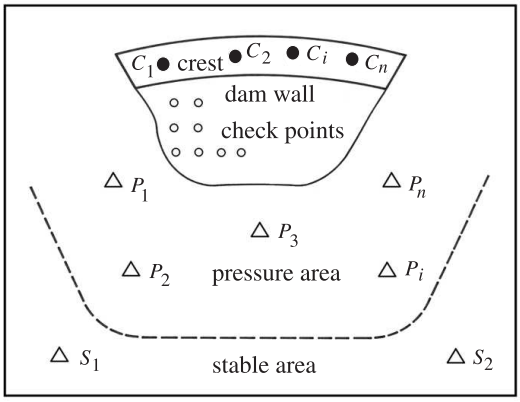

Figure 7.87. Dam control with GPS

The relative location of the two *portal networks* can be determined with an accuracy level below 1 cm for distances up to 10 km. The results are given in three dimensions. In order to provide "levelled heights" via GPS, it is necessary to include a precise local geoid [7.6.2.3]. If required, the tunnel network can be easily connected with the nearest control points (TP) of the geodetic network via GPS techniques.

The second example refers to the permanent control of a dam, during construction and after completion (Fig. 7.87). A difficult task is the selection of stable control points, and the delimitation of the pressure zone from the stable area. Usually, the advice and support of experts is required. One advantage of GPS is that the stable control points S_i can be placed well away from the influence zone, and that no direct sight connection to the near-construction control points, P_i, is required. GPS can be used to

- establish stable control points (S_i),
- establish and monitor control points in the pressure zone (P_i), and
- establish control points on the dam crest (C_i).

Checkpoints attached to the dam wall remain to be controlled by other techniques, either with electronic tacheometers or photogrammetry. GPS is suited to determining and monitoring the coordinates of the tacheometer or camera standpoints, P_i, in the pressure zone. GPS is also capable of identifying and analyzing point motion within the pressure zone. Regarding the high potential of GPS concerning accuracy and cost-effectiveness, it is possible to install a dense network of control points in the potential pressure zone.

Deformations can be derived from repeated observation, in intervals of days, weeks, or months, depending on the situation. In cases where there is suspicion of impending

structural distress, the establishment of a continuous monitoring array can be taken into consideration.

One critical factor is the limited visibility of satellites from stations near the dam wall. The situation will improve with the inclusion of other GNSS like GLONASS and GALILEO [7.7], but drawbacks result from unbalanced geometry and multipath effects. Instead, control points near the dam wall can be related to better placed control points by tacheometry.

GPS can also be used for deformation monitoring at the one millimeter or sub-millimeter level, when all acting error sources are eliminated or considerably reduced. The most critical part, multipath, can be eliminated by forming sidereal differences [7.4.4.3], because the satellite geometry repeats after 24 hours in sidereal time. The technique has been successfully applied for the monitoring of deformation during the filling process of locks (Seeber et al., 1997a; Wübbena et al., 2001a).

7.6.2.7 Precise Marine Navigation, Marine Geodesy, and Hydrography

Because of the real-time capability, continuous availability, and the high accuracy potential, this field of use is very broad, continuously growing, and is developing fast. In this chapter only a short overview is given. For more information see [12.3] [7.5.1] and the ample literature in symposia proceedings like *INSMAP 94, INSMAP 98* or journals like *Navigation, GPS World, Sea Technology*.

The possible applications, and the related accuracy requirements, can be divided into three user groups:

(a) low accuracy requirements, about 10 to 100 m in position, and 1m/s in velocity,
(b) medium accuracy requirements, about 1 to 10 m in position, and 0.1 m/s in velocity, and
(c) high accuracy requirements, better than 0.1 m in position and height, and 0.01 m/s in velocity.

User inquiries indicate that highest interest is in the group (b), i.e. a position requirement of a few meters.

User group (a) can be fully satisfied with a single C/A-code navigation receiver aboard a ship. GPS will provide continuous two-dimensional position accuracy of about 10 to 30 m, or better, under the Standard Positioning Service [7.4.1]. Important areas of employment in user group (a) are, for example (cf. [12.3]):

(1) general navigation tasks on the high seas,
(2) research in oceanography,
(3) ship's positioning in small scale bathymetry with swath systems, and
(4) position and velocity in small scale gravimetric, magnetic, and seismic measurements.

For some applications of tasks (3) and (4), the accuracy of a single operating receiver is not sufficient. In these cases, and for the majority of applications *(user group (b))* in marine geodesy, hydrography, and precise navigation, GPS must be operated in the relative mode (*Differential GPS*, see [7.5.1]).

Typical fields of application in user group (b) are, for example (cf. [12.3]):

(1) precise navigation in coastal waters,
(2) harbor approach,
(3) sea floor mapping in the *Exclusive Economic Zone* (EEZ), for the delimitation of seaward boundaries and/or for scientific purposes (cf. Fig. 12.10, p. 524),
(4) hydrography,
(5) precise gravimetric and seismic surveys,
(7) positioning of underwater sensors and samplers in marine prospecting for mineral resources, and
(6) calibration of transponder arrays.

In cases where the data are not required in real-time, the final positions can be computed afterwards (post-mission) in a post-processing step. However, considering the huge amount of data it is advisable to determine the ship's position in real-time and not to store the original raw data.

A further option of the differential mode is to use the carrier phase data at the remote station to smooth the code phase observations [7.3.6] with an appropriate filter algorithm (7.108, p. 296). This method works on a routine basis if an appropriate receiver is used, and provides a continuous accuracy of 2–3 m for the moving antenna, or even better. The accuracy level satisfies most users of the above list, in particular in hydrography and precise surveying activities.

An increasing user market requires an accuracy level of better than 0.1 m, in particular in the height component (*user group (c)*). In this case, the carrier phase observable has to be used as the primary quantity, and the ambiguities have to be resolved. The *pure kinematic method* [7.3.5.4] with ambiguity resolution techniques "on the way" [7.3.2.3] has to be applied. The methods work well with postprocessing and also in real-time if a data link of sufficient capacity is available [7.5.1.2]. For larger areas, the concept of multiple reference stations can be applied to model the distance dependent errors [7.5.3].

Possible applications in user group (c) are:

(1) precise hydrographic surveying,
(2) monitoring silt accretion and erosion in rivers, lakes, estuaries, coastal waters, and harbor areas,
(3) real-time dredge guidance and control,
(4) support of coastal engineering,
(5) marine geodynamics.

Two further particular applications are:

(6) precise continuous height control, and
(7) attitude control of ships, buoys, floating platforms.

For precise echo-sounding and sea level monitoring a continuous height determination with an accuracy of a few centimeters is required and feasible (Goldan, 1996; Goffinet, 2000; Böder, 2002).

The actual sea level at the location of the surveying vessel must be referred to the height reference onshore (depth reduction). The conventional method is to estimate the *depth reduction*, dh, from tidal and hydrodynamic models with respect to tide gauges onshore. With GPS, the reduction can be determined directly (Fig. 7.88). The GPS

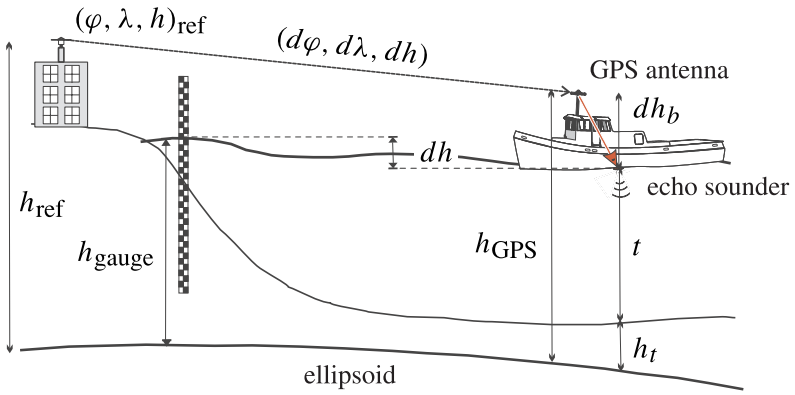

Figure 7.88. Depth reduction for echo-sounding; conventional and with GPS

antenna phase center does not coincide with the reference point of the echo-sounder (Fig. 7.89). The horizontal and vertical corrections are given by

$$dX = X - \sin(\beta + \gamma) \cdot S \tag{7.170}$$

$$dZ = Z - \cos(\beta + \gamma) \cdot S. \tag{7.171}$$

To minimize the effect of ship inclination on the depth correction, dZ, it is recommended to install the GPS antenna directly above the sounder ($\beta = 0$).

GPS onboard an anchored ship or a moored buoy can also be used to monitor tidal variation. The resolution is a few centimeters, depending on the size and behavior of the platform (a larger platform shows smaller noise), Goldan (1996). A challenging application is continuous height control in calibration areas for altimeter satellites [9.3.3], see Fig. 9.10

With three antennas/receivers on board a ship (Fig. 7.90) the time-dependent spatial behavior of the platform, its *attitude*, can be monitored in real-time (Seeber, Böder, 1998). The achievable accuracy depends on the baseline length between the antennas, and the noise in the GPS result (Table 7.24).

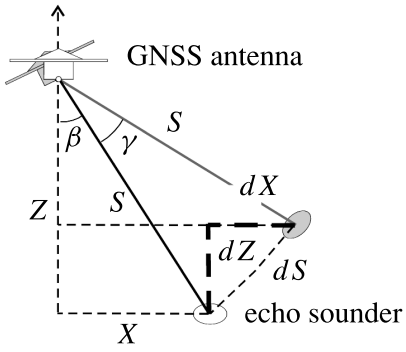

Figure 7.89. Inclination correction in echo-sounding

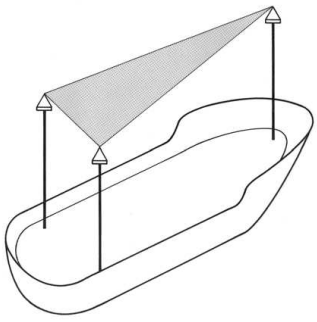

Figure 7.90. Attitude control with three GPS antennas

Table 7.24. Relationship between a height error, dx, baseline length, s, and GPS derived orientation accuracy

s \ dx	3 mm	1 cm	0.1 m
1 m	0.17°	0.57°	5.71°
5 m	0.03°	0.11°	1.15°
10 m	0.02°	0.06°	0.57°
30 m	0.006°	0.02°	0.19°

In general, a resolution of 0.1° is sufficient, for example for the correction in (7.171). Attitude control is of particular importance for the inclination correction of multibeam sonar systems in sea-bottom mapping ([12.3], Fig. 12.10, p. 524), and for the monitoring of floating GPS sensors at the sea surface in the precise location of submarine geodetic control markers (cf. [12.3.2], Fig 12.12, p. 526). For the mathematics of attitude determination see Kleusberg (1995); Cohen (1996).

Note that most developments in precise marine navigation with GPS can easily be applied in land navigation and remote vehicle control.

7.6.2.8 Photogrammetry, Remote Sensing, Airborne GPS

The use of GPS contributes in several different ways, for example:

(a) determination of ground control points in photogrammetry,
(b) navigation of sensor carrying airplanes, and
(c) determination of sensor platform coordinates and orientation.

The determination of *ground control points* (group a) for photogrammetric map production corresponds completely to the procedures discussed in [7.6.2.1]. The technique and effort required depend on the desired map scale. For cadastral purposes, centimeter accuracy can be achieved with carrier phase adjustment. Usually the photogrammetric products have to be related to the official reference frame via at least one control point with known coordinates [12.1].

The accuracy requirements are much less for control points and ground truthing in *satellite images* (e.g. SPOT, LANDSAT). The level of 1 to 5 m can be achieved by differential techniques using code or carrier-smoothed code observations only, without resolving ambiguities [7.5.1]. The remote receiver can be operated over distances up to several hundred kilometers. It is sufficient to collect only a few minutes of data on the site.

For the *precise navigation* (group b) of a survey aircraft the differential mode and a real-time data link are required. Usually the transmission of range corrections [7.5.1.1] is sufficient, to assure an accuracy of several meters, as long as at least four satellites are visible. Conventional DGPS services are well suited to the task.

The most promising contribution of GPS to photogrammetry is the determination of the *sensor orientation*, in particular the *precise camera position* (group c) in order

to support *aerial triangulation* (Li, 1992; Lee, 1996; Schmitz, 1998), Fig. 7.91. GPS determined camera positions are introduced as precise observations into the combined block adjustment. As a consequence the required number of ground control points can be reduced to about 10 percent, or even less, of those required in conventional aerotriangulation (Jacobsen, 1997, 2000).

In order to achieve the required accuracy level of about ±5 cm it is necessary to

- operate in the differential mode,
- use code and carrier phase data, and
- resolve the phase ambiguities.

Because of the cycle slip problem, in particular in the turns between individual survey strips, it is necessary to use ambiguity resolution techniques "on the fly" [7.3.2.3]. Receivers that provide sufficient channels for all satellites, both frequencies, and low noise code observations are particularly suitable.

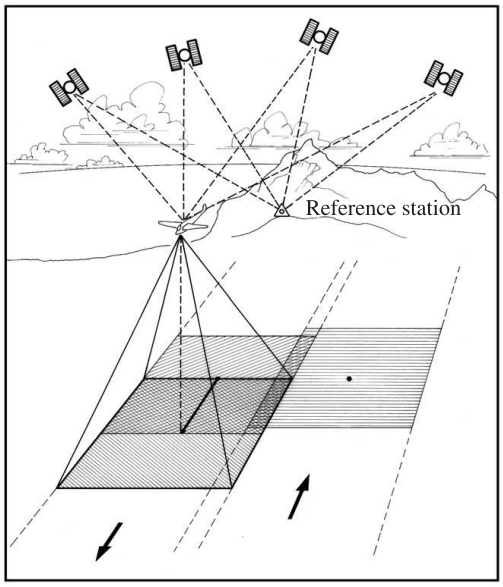

Figure 7.91. The use of GPS for camera positioning in aero-triangulation

The following problems or aspects have to be considered:

- simultaneity of receiver and camera operation,
- eccentricity between antenna phase center and camera projection center, and
- loss of satellite track or cycle slips in turns.

Modern GPS receivers and aerial cameras allow nearly *synchronous operation*. It is usually not possible for a GPS receiver to measure at arbitrary epochs, hence the camera shutter has to be triggered by an output signal from the receiver. For older aerial cameras it is advisable to operate the shutter manually, or by some external device, as near as possible to the GPS observation epoch, and to register the mid-open time of the shutter. Considering the average speed of a photogrammetric aircraft, asynchronous operation may introduce errors of up to several meters.

Another possibility is to interpolate the aircraft positions between the GPS positions with an *inertial platform* (INS). The integrated techniques of GPS and INS provide an accuracy of a few centimeters (Lee, 1996; Cramer, 2001).

The *3-D eccentricity* between the GPS antenna and the camera projection center (Fig. 7.92) includes the distance and the three orientation angles. The distance is invariable and has to be measured by conventional means. The orientation can be determined

- by a GPS based platform orientation unit with three GPS antennas (cf. [7.6.2.7]),
- as a by-product of an inertial package onboard, or
- with inclinometers.

If the camera is fixed to the aircraft body, the orientation angles serve at the same time for the eccentricity calculation and for the outer orientation of the photogrammetric camera.

Detailed investigations demonstrate that external information on the angles of orientation is of much less importance than the precise coordinates of the focal center. The orientation angles can be easily determined along with the bundle block adjustment (Jacobsen, 1992).

One particular problem is *signal loss* from some of the satellites, through the inclination of the aircraft whilst banking. If less than four satellites remain visible, a new cycle ambiguity has to be determined. Either powerful ambiguity resolution techniques *'on the fly'* have to be applied [7.3.2.3], or the data gap has to be bridged by inertial techniques. With the current constellation of more than 24 GPS satellites, signal loss is no longer a serious problem, because ambiguities can be easily fixed.

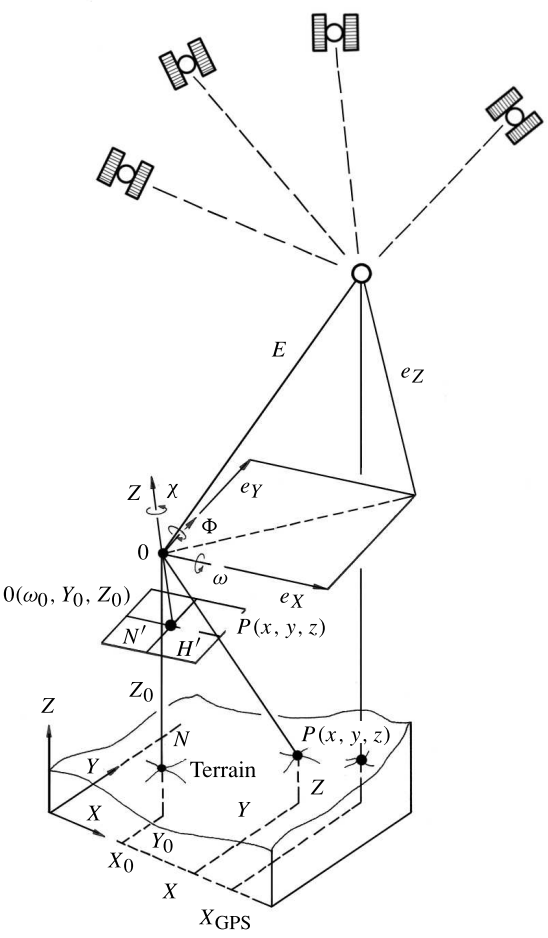

Figure 7.92. Eccentricity elements between camera, focal center, and GPS antenna (Li, 1992)

Because of the many restrictions in photogrammetric survey planning (weather conditions, vegetation period, etc.) the routine use of GPS requires the continuous availability of reference observations, for example from continuously operating reference stations [7.5.3]. Usually, post-processing techniques are applicable. Real-time results are also possible in areas with networked multiple reference stations like SAPOS in Germany [7.5.3.2].

The methods of precise platform positioning with GPS can also be used for related applications such as *laser bottom profiling* or *airborne gravimetry* (Schwarz et al., 1997).

7.6.2.9 Special Applications of GPS

As stated above, the possible applications of GPS in the field of engineering and geoscience are unlimited. Some further examples are

- glacial geodesy,
- time transfer,
- GPS carrying satellites, and

– GPS meteorology.

In *glacial geodesy*, and *Antarctic research*, GPS can be employed successfully to determine the movement of glaciers or ice sheets (Hinze, 1990). If motion parameters (velocity and azimuth) are to be derived from repeated measurements over the years, a quasi-online reading suffices, while a route is traversed with snow mobiles, or a helicopter lands for a short time. In the relative mode, to a fixed station, sub-decimeter accuracy can be attained with short observation times, depending on the distance to the reference station [7.5.1], so that correct results can be expected from repeated measurements after about 1 month in the same season. Fig. 7.93 gives an example from the Ekström Ice Shelf near the German Antarctic station "Georg v. Neumeyer".

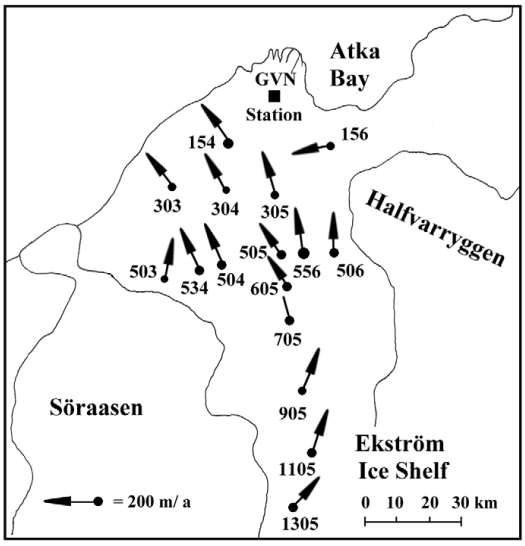

Figure 7.93. Ice flow from GPS observations

GPS is one of the most efficient means for operational *global clock synchronization*. Table 7.25 gives an overview of the achievable accuracy (Lombardi et al., 2001). In most cases, the so-called *common-view technique* is applied. The time of arrival of the same signal from one satellite is observed at two stations, and compared with the local reference clocks. Afterwards the data are exchanged. The signal travel time between the satellite and the station has to be calculated, based on precise coordinates for both stations and precise satellite orbits. *Single channel* technique means that the measurements follow a pre-determined tracking schedule. In the *multi-channel* common-view technique, data from all satellites in view are recorded without a schedule. The latter mode facilitates continuous comparison of standards with no gaps in the data.

Table 7.25. Accuracy of GPS time transfer

Technique	Timing Uncertainty 24 hours, 2 σ	Frequency Uncertainty 24 hours, 2 σ
One-Way	< 20 ns	$< 2 \times 10^{-13}$
Single-Channel, Common-View	~ 10 ns	$\sim 1 \times 10^{-13}$
Multi-Channel, Common-View	< 5 ns	$< 5 \times 10^{-14}$
Carrier-Phase, Common-View	< 500 ps	$< 5 \times 10^{-15}$

A station position error of 3 cm enters 100 picoseconds into the error budget. The effect of orbital errors follows the rule of thumb (7.134), hence an orbit accuracy of

0.1 m is required for 100 picoseconds time transfer over 5000 km. The *International GPS Service* (IGS) [7.8.1] considerably supports operational high precision global time transfer through its station network and products. A number of the stations are connected with external oscillators like H-masers, cesium and rubidium clocks (IGS, 2000)

Several manufacturers offer dedicated GPS receivers for time-transfer. For the nanosecond accuracy level all error influences including hardware delays have to be carefully modeled. Real-time relative time transfer at the 100 picoseconds level has been demonstrated within the "Internet-Based Global Differential GPS" project of the NASA-JPL (Powers, et al., 2002), see also [7.5.1.2]. For a topical treatment of GPS time transfer see for example Schildknecht, Dudle (2000).

Very powerful GPS Earth Science applications result from the deployment of GPS sensors on *near Earth orbiting satellites*, so-called LEOs [3.4.2]. The GPS data received at the orbiting platform may serve for

- precise orbit determination of remote sensing satellites, primarily altimeter satellites [9],
- precise position and orbit determination of satellites probing Earth's gravity field [10],
- attitude control of space vehicles, and
- analysis of GPS signals after passing the atmosphere (GPS-MET).

One of the first demonstrations for precise orbit determination with GPS (see [3.3.2.3]) was with the TOPEX/POSEIDON mission (Melbourne et al., 1994a). Since then, GPS receivers have been included in a number of missions, in particular on LEO satellites like CHAMP, GRACE, JASON-1, and ICESAT. Precise orbit determination (POD) [3.3.2.3] is supported by the orbit and clock products of the IGS. The accuracy level is in the order of a few decimeters and may reach sub-decimeter after tailored gravity field improvement (Wickert et al., 2001). On the other hand, LEO data are of interest to IGS. The IGS has started a pilot project to study the inclusion of LEO data into the regular IGS products (IGS, 2000).

If the satellite carries three or more GPS antennas it is possible to determine its attitude. Since the baseline between the antennas is always very small, and only the carrier phase difference is required, single frequency C/A-code receivers can be used. For details of the technique, see for example (Purivigraipong, Unwin, 2001).

GPS contributes with two different techniques to the improvement of global weather data. The continuous observations at more than 200 IGS sites are used to model the total zenith delay at a level of 3 to 5 mm, that corresponds to better than 1 millimeter in water vapor (IGS, 2000). The data are available as a regular IGS product and can be used by meteorological institutions in numerical weather prediction models. For details of the subject see also Schüler (2001). For very dense networks of monitor stations, for example in Germany, with a spacing of about 50 km the accuracy of the integrated water vapor was found to be 1 to 2 mm, with a delay of only 40 minutes (Reigber et al., 2002).

The second technique uses the observations made between GPS satellites and LEOs equipped with GPS receivers. Fig. 7.94 demonstrates how GPS contributes to atmospheric research.

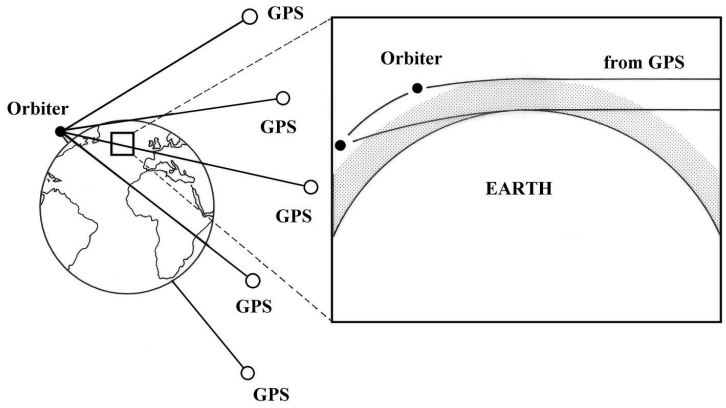

Figure 7.94. The use of GPS in atmospheric research (after Yunck, Melbourne (1990))

Due to the relative motion between the LEO satellite, and a GPS satellite setting behind Earth's disk, the tangential point of the radio link between the two space vehicles moves downward with a geocentric velocity of about 2.5 to 3 kilometers per second (Hocke, Tsuda, 2001) and scans the atmospheric layers from the high atmosphere down to Earth's surface. The signals are affected both by the ionosphere and the troposphere and can be used for ionospheric tomography as well as for mapping the integrated water vapor. The technique is known as *radio occultation* or *limb sounding*.

A first experiment (GPS/MET) was carried out with the launch of the MICROLAB satellite in 1995 (Hocke, Tsuda, 2001). Other suitable satellites for radio-occultations are ÖRSTED (launch January 1999), CHAMP and GRACE. For details on the technique see e.g. Kleusberg (1998). First results from CHAMP radio occultations are reported in Wickert et al. (2001). CHAMP measures at a rate of 50Hz and provides about 230 globally distributed vertical profiles of atmospheric parameters per day.

7.7 GNSS – Global Navigation Satellite System

GPS is not the only satellite-based navigation system. The Russian Federation is operating GLONASS, and the European Union together with the ESA is planning GALILEO. In addition, various augmentations to GPS are under preparation. The general name given to these systems is *Global Navigation Satellite System* (GNSS). Most of the material, outlined in chapter 7, is also valid for other GNSS systems. This is why, in many publications, instead of GPS the more general term GNSS is used.

The term GNSS was coined at the 10th Air Navigation Conference in 1991, when the *International Civil Aviation Organization* (ICAO) recognized that the "primary

stand-alone navigation system in the 21st century will be provided by the Global Navigation Satellite System (GNSS)" (Hein, 2000). As commonly understood GNSS includes more than just satellite-based positioning. Important features besides accuracy are *integrity*, *availability*, and *continuity of service*. GPS and GLONASS, being primarily military systems, do not guarantee these capabilities. On the way to establish GNSS, several steps have been defined.

GNSS-1 is based on GPS and/or GLONASS as backbone, and is augmented by additional civil components.

GNSS-2 is a second-generation satellite navigation system which fulfills the above requirements, such as GPS IIF or the European GALILEO.

In the following, some of the particular features of GNSS developments are explained.

7.7.1 GLONASS

The former Soviet Union (SU) has, since the 1970s, been developing a navigation system very similar in design to GPS under the name GLONASS (*GLO*bal *NA*vigation *S*atellite *S*ystem). The Russian denomination is *Global' naya Navigatsionnaya Sputnikowaya Sistema*. Today, GLONASS is continued by the Russian Federation. Like GPS, GLONASS is a military system, but it has been offered to civil use by several declarations of the Russian Federal Government (Slater et al., 1999). The system was officially put into operation on September 24, 1993 as a first-stage constellation of twelve satellites. By the end of 1995 the constellation was expanded to 24 satellites (standard constellation). Due to a lack of new launches, the constellation has since then decreased considerably. By the end of 2002 only 7 satellites were operational.

Similar to GPS with SPS and PPS [7.1.6], [7.4.1], GLONASS provides a *standard precision* (SP) and a *high precision* (HP) navigation signal. The SP signal is continuously available to all civil users world-wide. The specification for SP accuracy is 50 to 70 m horizontally and 70 m in height. Information for civil users is available via the *Coordinational Scientific Information Center* (CSIC) of the Russian Space Forces.

In this section, some basic information on GLONASS is given. For further reading with additional references see e.g. Kaplan (1996, chap. 10), Daly, Misra (1996), Habrich (2000), Roßbach (2001). A short introduction is given by Langley (1997a). Table 7.26 compares GLONASS with GPS and indicates similarities and differences. It is evident that the systems have strong similarities. The main characteristics and differences are as follows.

(a) *Satellite orbits*
In the baseline constellation, both systems consist of twenty-four satellites including three spares. Unlike GPS, the GLONASS satellites are arranged in 3 orbital planes 110° apart. Each orbital plane contains eight equally spaced satellites. Fig. 7.95 shows the complete configuration.

The ground tracks repeat every seventeen orbits, or eight sidereal days. The orbits are arranged in such a way that resonance phenomena do not occur and that in one

8-day period all satellite footprints pass through a given position. As a consequence, for an observer on Earth, a particular satellite will pass the same point in the sky after eight sidereal days, and one of the satellites in each orbital plane will appear at the same location in the sky at the same sidereal time each day (Forsell, 1991; Kaplan, 1996).

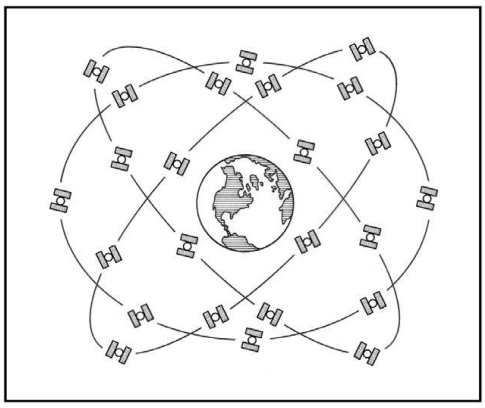

Figure 7.95. GLONASS satellite configuration

Both systems provide similar coverage for the 24 satellite constellation. Between 6 and 11 satellites are visible at any place on the Earth for either system, if fully deployed. Hence, both systems together can generate a coverage of about 12 to 16 satellites simultaneously for a given spot on Earth. Due to the differing inclination angle, the position of the "shadow area" (see [7.6.1.1], Fig. 7.69, p. 347) is also slightly different and leads to a better overall coverage of the sky.

(b) *Satellite navigation signals*

GLONASS is, like GPS, a one-way ranging system. The radio-signal structure of both systems is very similar. Two carrier signals in the L-band are broadcast, and the signals are modulated by two binary codes and the message. In contrast to GPS, all GLONASS satellites transmit carrier signals at different frequencies. The L1 frequencies are

$$f_{L1} = f_0 + k \cdot \Delta f_{L1} \qquad k = 0, 1, 2, \ldots, 24, \tag{7.172}$$

where $f_0 = 1.602$ MHz, and $\Delta f_{L1} = 0.5625$ MHz. L1 and L2 are related by

$$f_{L1}/f_{L2} = 9/7. \tag{7.173}$$

Here, k means the frequency number of the satellite. Number 0 is designated as "technical frequency" and reserved for testing purposes. The procedure is called *frequency division multiple access* (FDMA). Satellites need not be distinguished by their unique satellite modulation, hence all GLONASS satellites use the same codes. The code frequencies are about half the corresponding GPS values, hence the range resolution may be slightly lower for GLONASS. The frequency ranges of both systems are close, thus permitting the use of a combined antenna and common input amplifiers in the user equipment, and allowing the development of combined receivers. The signal processing, however, is different (Roßbach, 2001).

Originally, each of the 24 GLONASS satellites should have transmitted on its own unique frequency. Since some of the GLONASS transmissions cause interference to radio astronomy in the frequency bands 1610.6–1613.8 MHz and 1660–1670 MHz, as well as to some communication satellites, it was decided to shift the GLONASS

Table 7.26. Comparison of GLONASS and GPS

Parameter	GLONASS	NAVSTAR GPS
Satellites		
Number of satellites in the baseline constellation	21 + 3 spares	21 + 3 spares
Number of orbital planes	3	6
Inclination	64.8°	55°
Orbital altitude	19 100 km	20 180 km
Orbital radius	25 510 km	26 560 km
Orbital period (sidereal time)	11 hours 15 min.	12 hours
Repeat ground tracks	every sidereal day	every 8 sidereal days
Navigation message		
Ephemeris representation	9 parameters (position, velocity, acceleration) in the ECEF Cartesian system	Keplerian elements and interpolation coefficients
Geodetic datum	PZ-90	WGS 84
Time base	GLONASS system time	GPS system time
Related system time	$UTC_{[SU]}$	$UTC_{[USNO]}$
Almanac transmission	2.5 minutes	12.5 minutes
Signals		
Satellite signal division	Frequency division	Code division
Frequency band L1	1.602–1.615 MHz	1.575 MHz
Frequency band L2	1.246–1.256 MHz	1.228 MHz
Codes	same for all satellites C/A-code on L1 P-code on L1, L2	different for all satellites C/A-code on L1 P-code on L1, L2
Code type	PRN sequence	Gold code
Code frequency C/A-code	0.511 MHz	1.023 MHz
Code frequency P-code	5.11 MHz	10.23 MHz
Clock data	clock offset frequency offset	clock offset frequency offset and rate

frequencies to a somewhat lower domain. In addition, the number of frequency channels is cut in half; so-called "antipodal satellites", i.e. satellites in the same orbital plane separated by 180 degrees in the argument of latitude, share the same channel. The re-organization of the frequency plan occurs in several steps. From 1998 to 2005 the frequency numbers $k = -7, \ldots, 12$ are applied, and after 2005 the numbers $k = -7, \ldots, 4$ (5, 6 for testing only). The bands will, hence, be finally shifted to 1598.0625–1604.25 MHz for L1 and 1242.9375–1247.75 MHz for L2.

(c) *Navigation message*
The navigation data are bi-phase modulated onto the carrier at 50 bits/s. The binary sequence has a total length of 2 seconds, called one line. The digital data structure is formed by navigation superframes of 2.5 minutes in length. A superframe consists of five frames of 30 seconds each, and each frame consists of fifteen lines (subframes).

As with GPS, the GLONASS message contains precise orbital information (ephemeris data) about the individual satellite's own position and status, and less precise "almanac" information about other satellite positions. Lines 1–4 of a frame contain the ephemeris data of the transmitting satellite, and line 5 general information concerning the entire system. Lines 6–15 contain the almanac data for five satellites. The almanac data of the complete system hence require one superframe, corresponding to 2.5 minutes.

Details of the data format can be found in the official *GLONASS Interface Control Document* (ICD-GLONASS) or in the literature cited above. The navigation message contains, for example

- coordinates for the i-th satellite in the ECEF reference frame for the reference time,
- speed vector components for the i-th satellite,
- acceleration vector components caused by Earth and Moon gravity,
- time scale correction to the GLONASS time scale for the i-th satellite, and
- satellite identification number, status information, reference time.

The GLONASS broadcast ephemerides are updated every 30 minutes and refer to the center of the 30 minutes time interval. For a measurement epoch in between these half-hour marks, the satellite position has to be interpolated using the position, velocity and acceleration data from the reference epochs before and after the observation epoch. These data are used as initial values for an integration of the equation of motion [3.3], (e.g. Roßbach, 2001).

(d) *Control Segment*
The ground-based control segment is responsible for (Kaplan, 1996)

- prediction of satellite ephemerides,
- uploading of ephemeris, clock correction and almanac data into each satellite,
- synchronization of the satellite clocks with GLONASS system time,
- estimation of the offset between GLONASS system time and UTC (SU), and
- spacecraft control.

The ground segment consists of

- the System Control Center,
- the Central Synchronizer,
- several Command and Tracking Stations, and
- Laser Tracking Stations.

The ground control center is in Moscow. The monitoring stations are uniformly distributed over the territory of the former Soviet Union, hence lacking global coverage. The navigation and control parameters are uploaded twice per day to each satellite. The central synchronizer forms the GLONASS system time and is related to the "phase

control system" (PCS), which monitors the satellite clock time and phase signals. Two laser tracking stations measure the distance and orientation of the GLONASS satellites with the objective to calibrate radio frequency tracking measurements. All GLONASS space vehicles are equipped with laser reflectors. The error specifications for GLONASS broadcast orbits are, following the Interface Control Document, 20 m (along track), 10 m (cross track), and 5 m (radial) for the position vector, and 0.05 cm/s (along track), 0.3 cm/s (cross track), and 0.3 cm/s (radial) for the velocity vector. The time scale synchronisation is within 20 ns.

(e) Geodetic Datum
The satellite coordinates are given in the PZ-90 (*Parametry Zemli 1990*) geodetic datum. Until 1993, the "Soviet Geodetic System 1985" (SGS 85) was in use. The main defining parameters of PZ-90 are given in Table 7.27. Note that slight differences exist compared to WGS 84 (see Table 2.1, p. 28).

Table 7.27. Main defining parameters of PZ-90

Parameter	Value
Semi-major axis	6 378 136 m
Flattening	1/298.257
Geocentric gravitational constant	$398\,600.44 \times 10^9$ m^{3}s^{-2}
Earth rotation rate	$7\,292\,115 \times 10^{-6}$ rad s^{-1}
2$^{\text{nd}}$ zonal harmonic	-1082.63×10^{-6}

For GLONASS only solutions, the user positions are determined in the PZ-90 frame. For combined GPS-GLONASS solutions a transformation between PZ-90 and WGS 84 (or ITRF) is required. Transformation parameters for a seven parameter transformation (2.46) [2.1.5] can be derived with collocated GLONASS/GPS observations at a number of stations and/or GLONASS observations at stations with known WGS 84 or ITRF coordinates (Table 7.28). An alternative approach is the computation of precise orbits with microwave and/or laser observation within the ITRF frame, and

Table 7.28. Transformation parameters from PZ-90 to WGS 84, 1: Misra et al. (1996), 2: Roßbach et al. (1996), 3: IGEX-98 (BKG), 4: IGEX-98 (ITRF-97)

Source	ΔX	ΔY	ΔZ	$\varepsilon_x[''] $	$\varepsilon_y[''] $	$\varepsilon_z[''] $	m
1	0 m	2.5 m	0 m	0	0	-0.39	0
2	0 m	0 m	0 m	0	0	-0.33	0
3	0.06 m	0.07 m	-0.57 m	0.035	-0.021	-0.358	$-1 \cdot 10^{-8}$
4	0.3 m	0.0 m	-0.9 m	0	0.012	-0.354	0

comparison with GLONASS broadcast orbits. Several solutions have been published in recent years. For an overview see Roßbach (2001).

Of particular interest are the results of the *International GLONASS Pilot Experiment* IGEX-98 campaign (Slater et al., 1999). In general, the transformation parameters are smaller than the related standard deviation. Significant values have only been found for a rotation of about -0.35 arcseconds around the Z-axis. Note that the quality of the parameters depends on the realization of the PZ-90 frame through the GLONASS broadcast orbits. Analysis of the IGEX-98 data revealed an accuracy of about 5 m for the broadcast ephemerides.

(f) *System Time*
Navigation signals of GLONASS and GPS are tied to slightly different system times. The GPS system is related to the UTC standard maintained by the U.S. Naval Observatory ($\mathrm{UTC}_{[USNO]}$), while GLONASS system time refers to the UTC standard in the former Soviet Union ($\mathrm{UTC}_{[SU]}$). Unlike the GPS time scale, GLONASS system time considers leap seconds, and it has a constant offset of three hours (difference Moscow time to Greenwich time). GLONASS system time is generated and controlled by the *GLONASS Central Synchronizer*, based on a set of hydrogen masers. The relationship between UTC and GLONASS time is

$$t_{\mathrm{UTC}} = t_{\mathrm{GLONASS}} + \tau_c - 3^{\mathrm{h}}. \tag{7.174}$$

The discrepancy, τ_c, comes from the different clock ensembles used and can reach several microseconds. τ_c is communicated to the GLONASS users in frame 5 of the GLONASS navigation message.

When using both navigation systems jointly, the difference in system time, $\Delta\tau$, depends on the clocks from both systems and has to be taken into account. The estimation of the clock offset, $\Delta\tau$, requires observations to one additional satellite. In other words, at least two GLONASS satellites must be added to the GPS configuration in order to contribute to the positioning solution.

(g) *System status*
The first satellite in the GLONASS System was launched on October 12, 1982. The satellites are carried into orbit three at a time by PROTON launchers from the Baijkonur Cosmodrome in Kazakhstan. The booster is first brought into a low altitude orbit, already with the final inclination of 64.8 degrees, and is then transported via an ascending ellipse (Hohmann transfer [3.4.3]) to the apogee height of 19 100 km. The orbital positioning of the three satellites is performed by their own thrusters. Launches 1 to 6 were pre-operational launches and also carried "dummy" satellites, without navigation payload. The operational deployment phase began with the seventh mission in 1985.

As of January 2001, with 30 launches, in total 74 GLONASS satellites were placed into orbit, along with two passive ETALON satellites [8.2]. The current operational satellites have a mass of 1400 kg, carry three cesium beam oscillators and have a design lifetime of three years. The physical configuration of the GLONASS spacecraft

is depicted in Fig. 7.96. The three-axis stabilized satellite is equipped with a propulsion system for station keeping and relocation, attitude control, and laser corner-cube reflectors. The antenna beamwidth of 35 to 40 degrees provides navigation signal reception up to 2000 km above Earth's surface.

The numbering scheme is manyfold. Besides the international satellite ID number, the GLONASS satellites are given numbers in the COSMOS series, a sequential GLONASS number, an orbital position number, and a channel number. The usual identification follows the channel number.

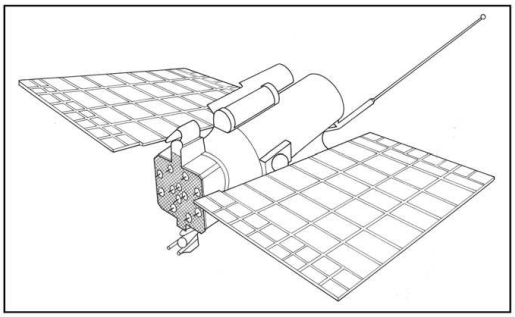

Figure 7.96. GLONASS spacecraft

A new generation of spacecraft, intended to replace older satellites, is under preparation and commonly referred to as GLONASS-M (M for modified). A first launch took place in December 2001, but as of December 31, 2002, no GLONASS-M spacecraft was operational. The main advantages of the GLONASS-M series are

- longer lifetime of five to seven years,
- enhanced clock stability,
- intersatellite communication,
- autonomous spacecraft operation, and
- modified structure of the navigation format.

For details on GLONASS satellites see Johnson (1994); Kaplan (1996).

Due to the short design lifetime of the current spacecraft generation, frequent launches are required to maintain the constellation. During the first months of 1996 the constellation was fully deployed with 24 satellites. Thereafter, several spacecraft were withdrawn from service and not replaced. Between January 1997 and January 1999 12 to 16 satellites were always available. Since then, the number has continuously decreased. As of March, 2003, the following eleven satellites were in service:

- Plane 1: SV channel 2, 7, 8, 9, 12
- Plane 3: SV channel 3, 5, 5, 10, 10, 11.

Note the use of the same channel on pairs of antipodal satellites.

(h) *Use of GLONASS*

During the period of full deployment, GLONASS showed a similar performance to GPS. The advantages of GLONASS are that there has been no artificial signal degradation like SA, and that the P-code is fully available. The user range error (URE) shows a standard deviation of 8 to 10 m (Roßbach, 2001).

After 1992, several commercial receiver makes entered the market. Two groups of user equipment can be distinguished:

- navigation receivers, L1 C/A code and L1 carrier phase, and
- geodetic receivers, L1 C/A- and P-code, carrier phase, L2 P-code and carrier phase.

Several advanced receivers offer the possibility to observe both GPS and GLONASS satellites. Examples are the *Ashtech Z-18*, the series of *JPS* receivers, and the *Novatel Millennium* board. Fig. 7.97 shows the JPS "Legacy" dual frequency 40 channel receiver with the RegAnt antenna.

Figure 7.97. Combined GPS-GLONASS receiver JPS Legacy with RegAnt antenna

As with GPS, plans were developed in Russia to establish a differential GLONASS (DGLONASS) service. Because of the large size of the country the implementation went slowly (Ganin, 1995). Instead, developments started to combine GPS and GLONASS into a combined DGNSS service (Chistyakov et al., 1996). An important prerequisite was fulfilled with the inclusion of DGLONASS correction data into the format RTCM 2.2 from January 1998 [7.5.1.2].

GLONASS carrier phase data can be used, either alone or together with GPS data, for precise geodetic applications. For observation equations, modeling of observables, and ambiguity resolution see e.g. Habrich (2000) or Roßbach (2001).

In 1998/1999 a major effort was undertaken to exploit the potential of GLONASS for the geodetic community. Under the auspices of the IAG and the IGS and also ION and IERS, the *International GLONASS Experiment* (IGEX-98) was initiated and realized. The major objectives were to collect GLONASS data for several months from a worldwide network of tracking stations, compute precise orbits, evaluate receivers, and resolve geodetic reference frame and time system issues (Slater et al., 1999). The campaign lasted six months, from October 1998 until April 1999. Over 60 GLONASS tracking stations and 30 Satellite Laser Tracking (SLR) observatories in 25 countries participated. Precise orbits were computed by several analysis centers using the SLR and GLONASS receiver data, with accuracies of 20–50 cm. Datum transformation parameters relating PZ-90, WGS 84, and ITRF were analyzed. The most interesting results were discussed at a meeting in September 1999 and are published in a comprehensive report (Slater et al., 1999).

After the termination of IGEX-98, a number of stations (32 as of December 2000) continued dual frequency tracking within the *International GLONASS Service* (IGLOS) *Pilot Experiment* under the auspices of the IGS. The goals and objectives are (IGS, 2000)

- establish and maintain a global GLONASS tracking network,
- produce precise (10-centimeter level) orbits, satellite clock estimates, and station coordinates,
- monitor and assess GLONASS system performance,
- investigate the use of GLONASS to improve Earth orientation parameters,

- improve atmospheric products of the IGS, and
- fully integrate GLONASS into IGS products, operations, and programs.

To support these goals, three GLONASS satellites are tracked with high priority by Satellite Laser Ranging (see [8.5.1]).

The long-term success of an International GLONASS Service certainly depends on the reliability and maintenance of the GLONASS constellation. Many of the potential applications do not require a full constellation but take advantage of GLONASS as an augmentation to GPS. This may, however, become a critical issue if not enough new launches take place and the number of usable satellites further goes down.

7.7.2 GPS/GLONASS Augmentations

GPS and GLONASS are systems under military control and do not fulfill the requirements for safe navigation, in particular coming from the international aviation community. These requirements are, in particular,

- accuracy,
- integrity,
- availability, and
- continuity of service.

Accuracy requirements depend on the particular application, for example 4 m vertical position accuracy in Category I aircraft approach landing (FRNP, 2001). Differential techniques are required to meet these demands.

Integrity is the ability of a system to provide timely warnings to its users when it should not be used for navigation (see [7.4.6] and Langley (1999b)). This service requires a network of control stations and channels to transmit the warnings in due time to the user.

Availability means the ability of the system to provide usable service within the specified coverage area, and *continuity of service* means the availability of service without interruptions for the intended operations (Hein, 2000).

In order to meet these requirements, augmentation systems to the existing satellite navigation systems GPS (and GLONASS) have been established or are under development. These are the

- *Wide Area Augmentation System* (WAAS) in the U.S.A. [7.5.3.1],
- *European Geostationary Navigation Overlay System* (EGNOS) in Europe,
- *Multi-functional Satellite-based Augmentation Service* (MSAS) in Japan, and
- *Satellite Navigation Augmentation System* (SNAS) in China.

All are contributions to a first generation of a Global Navigation Satellite System (GNSS-1) and intend to provide seamless coverage of the whole globe. They are also known as *Satellite-Based Augmentation Systems* (SBAS). An alternative solution are *Ground-Based Regional Augmentation Systems* (GBRAS), broadcasting corrections on VHF.

The generic architecture of a satellite-based augmentation system is as follows. A network of GPS (GLONASS) stations at surveyed locations collects dual frequency

measurements of pseudorange and pseudorange rate for all spacecraft in view, along with local meteorological data. The data are processed, and generate precise corrections to the broadcast ephemerides and clock offsets. These corrections, together with system integrity messages, are transmitted to the users via a dedicated package on geostationary satellites. This technique also supports an additional GPS-like ranging signal between GEO and user. Hence, in total three additional signals are provided, a ranging, integrity, and WAD (wide area differential) signal.

The European contribution to GNSS-1, EGNOS (Benedicto et al., 1999), includes augmentations to GPS and GLONASS. It is described in more detail as an example. EGNOS is part of the *European Satellite Navigation Program* (ESNP) and is initiated by the *European Tripartite Group* (European Commission (EC), European Space Agency (ESA), EUROCONTROL) since about 1993. The current EGNOS space segment is composed of transponders on two geostationary INMARSAT-3 satellites, positioned over the Atlantic Ocean Region East (AOR-E) and the Indian Ocean Region (IOR). These satellites provide extra ranging signals over Europe. For the full operational capability (FOC), expected for 2004/2005, additional GEO transponders are required (Soddu, Razumovsky, 2001).

The EGNOS ground segment consists of about 40 "Ranging and Integrity Monitoring Stations" (RIMS), mostly in Europe. These RIMS collect ranging measurements from the GPS, GLONASS and GEO navigation signals on L1 and L2 frequencies. The collected data are transmitted to a set of redundant "Mission Control Centers" (MCC) where the integrity information, differential corrections, GEO satellite ephemerides and ionospheric delays are estimated. These data, together with the GEO ranging signal, are uplinked to the GEO satellites from where they are transmitted on the GPS L1 frequency, as GPS like navigation signals, to the users. It will be possible to receive EGNOS navigation data over Europe, South America, Africa, Western Australia, and a large part of Asia.

The U.S. WAAS architecture is very similar to EGNOS. For details see FRNP (2001). WAAS, however, does not include GLONASS satellites. The system is projected to be fully operational by the end of 2003.

Augmentation systems like WAAS, EGNOS, MSAS, or others, will make it possible for many applications to obtain DGPS accuracy without the cost of extra reference stations or radio data links, and they offer continent-wide coverage. In the long term, augmentation systems are likely to replace the conventional DGPS services.

7.7.3 GALILEO

The European Commission, together with the European Space Agency (ESA) and European industry, is building up a European Satellite Navigation System under the name *GALILEO* as Europe's contribution to GNSS-2. The system will be controlled by civil authorities and be inter-operable with GPS and GLONASS. It offers dual-frequency as standard and will provide real-time positioning and timing services at different levels of accuracy, integrity, and availability. Other than the existing satellite navigation systems, GALILEO is a suitable system for safety critical applications,

such as landing aircraft, guiding cars, tracking hazardous materials, and controlling rail traffic.

The GALILEO schedule comprises several phases. The *definition phase*, from 1999 to 2001, included the initial definition of requirements and system architecture. Two major studies took place, the EC study *GALA* on the system architecture, and the ESA study *GalileoSat* on the space segment. Based on the results of these studies, a 4 years *design and validation phase* from 2002 to 2005 was initiated by the European Council (EC) on March 26, 2002. This phase includes a consolidation of the requirements, the development of satellites and ground based components, and the in-orbit validation. A first experimental satellite will be launched by the end of 2004. Up to four operational satellites will be launched thereafter in 2005 and 2006 for final validation of the space and ground segment. The remaining operational satellites will be launched in the *deployment phase*, from 2006 to 2007, to reach the full *operational phase* in 2008.

The information within this section is mainly taken from ESA documents and the cited literature, e.g. (Forrest, 2002; Eisfeller, 2002). Details are subject to changes during the design and validation phase. For updated information see the ESA homepage, the journal *Galileo's World*, and conference proceedings like ION-GPS.

(a) *Space segment*

The GALILEO space segment, when fully deployed, consists of 30 satellites (27 operational + 3 active spares) in three circular Medium Earth Orbits (MEO) (Fig. 7.98). This configuration is also called *Walker constellation 27/3/1*. The inclination angle is 56 degrees, and the orbital altitude 23 616 km. The orbital period is 14 hours 4 minutes, and the ground tracks repeat after about 10 days. The constellation is optimized for Europe and provides a good coverage up to a northern latitude of 75 degrees.

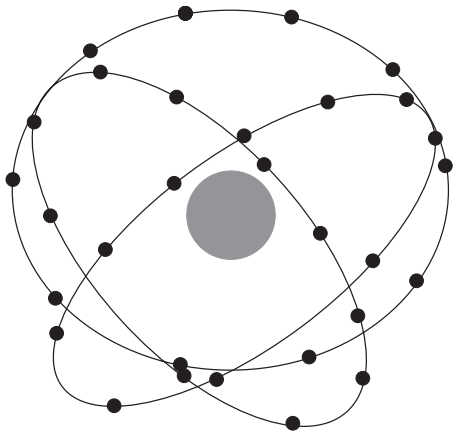

Figure 7.98. Probable GALILEO constellation

The GALILEO satellite (Fig. 7.99) has a mass of 625 kg and measures $2.7 \times 1.2 \times 1.1$ m^3; it hence belongs to the class of minisatellites. The navigation payload includes 2 rubidium standards and 2 hydrogen masers. Other than GPS, each satellite carries laser reflectors for independent orbit determination. Several deployment strategies are possible, for example up to 8 GALILEO spacecraft simultaneously launched with ARIANE 5, or up to 6 spacecraft simultaneously with the PROTON launcher. The injection is directly into the MEO orbit.

(b) *Ground segment*

The GALILEO ground segment consists of two *GALILEO Control Centers* (GCC). One is responsible for the control of satellites and the generation of navigation and time

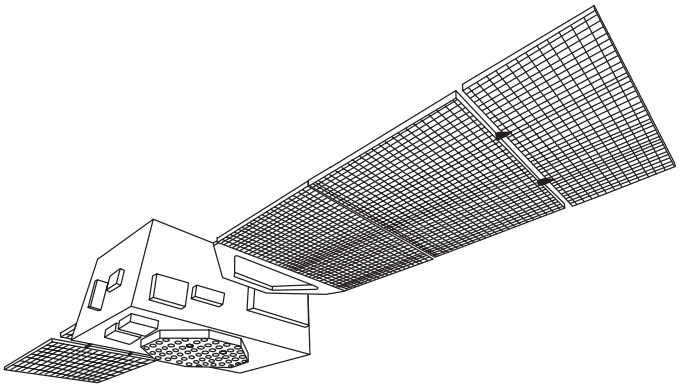

Figure 7.99. GALILEO satellite, possible schematic view

data; the other is responsible for the control of integrity. About 30 globally distributed monitor stations, the *GALILEO Sensor Stations* (GSS), provide data for the GCCs. Data transmission to the satellite is realized via 10 upload stations with S-band and/or C-band antennas.

A further feature is the global *Search and Rescue* (SAR) function. Each satellite is equipped with a transponder which is able to transmit emergency signals to a *rescue center*. A particular link gives a feed back to the user. In Europe the integrity service is closely related to the EGNOS system [7.7.2].

(c) *Services*
Several particular services will be offered within the GALILEO service framework. (Forrest, 2002):

- satellite navigation signals only services,
- combined services with other GNSS or with non GNSS, and
- local services.

Among the satellite only services, besides the "search and rescue service", are four "position, velocity and time services":

- *Open Service*, providing positioning, navigation and time for a mass market, free of charge;
- *Commercial Service*, with added value over the open service, for professional use with service guarantee and user fees;
- *Safety of Life Service*, includes integrity, in particular for landing approach and vehicle guidance;
- *Public Regulated Service*, for applications devoted to European/National security.

For geodesy, surveying and GIS, the open service and the commercial service are of particular interest. The accuracy with a single dual frequency receiver is estimated to be 4 m (horizontal), 8 m (vertical), 50 nsec (time) at the 95% level. The commercial service will have some additional features, such as augmentation with local elements like multiple reference stations [7.5.3].

(d) *User segment*
The receiver architecture will be similar to those used with GPS, but with modern elements in the digital signal processing and reference to the particular GALILEO signal structure. Combined GPS/GALILEO receivers will be designed at least for 4 frequencies. The user market is predicted for 2005 as follows (Eisfeller, 2002):
73% mobile phones,
23% car navigation,
1% aviation,
1% fleet management,
1% leisure, and
1% surveying.

(e) *Signal structure*
The GALILEO signal structure is not yet definitely decided. Probably the carrier frequencies, shown in Table 7.29, will be used in the lower, middle and upper L-band.

Table 7.29. GALILEO carrier frequencies, status August 2002

Carrier	Middle frequency (MHz)
E5a (L5)	1176.45
E5b	1207.14
E6	1278.750
E2 - (L1) - E1	1575.42

There are potential interferences with the GPS signals L1 and L5, mitigation of which will require particular modulation techniques. On the other hand, such signal overlap facilitates antenna design for hybrid receivers, and guarantees maximum interoperability. Difficulties also exist with the bandwidths. For E1 and E5 the bandwidth is just 4 MHZ and rather small for a robust signal. As shown, E6 is not a protected band, and hence its use is questionable. In the future at least 5 civil signals will be available to combined GPS/GALILEO receivers, namely code pseudorange and carrier phase:

Modernized GPS: L1, L2, L5
GALILEO: E1-L1-E2, E5a + b.

Simulations show (Eisfeller, 2002) that a combined evaluation of GPS/GALILEO data for geodetic purposes has several advantages:
- increased number of satellites (> 15),
- smaller PDOP (< 1.6),
- increased success rate of the ambiguity fixing, and
- increased positioning accuracy, by a factor of 2 for the horizontal and a factor of 3 for the vertical component.

An interoperable GNSS will enhance the use of satellite-based positioning in difficult environments like mountainous terrain, urban canyons, and around large structures like dams or industrial complexes.

(f) *Applications*
The expected field of possible applications is manyfold, as with GPS [7.6.2]. An overview is given in Table 7.30.

Table 7.30. Possible application markets for GALILEO (Forrest, 2002)

Professional	Mass market	Safety of life
geodesy	personal communication	aviation
precision survey	and navigation	maritime
land survey, GIS	cars	rail
photogrammetry	buses, trucks	police
remote sensing	commercial vehicles	fire
timing	inland waterways	emergency
mining	coastal waters	ambulance
oil and gas	outdoor recreation	search and rescue
environment		personal protection
fleet management		traffic surveillance
precision agriculture		
EEZ delimitation		
fisheries		
vehicle control		
robotics		
construction		
engineering		
meteorology		
space application		

7.8 Services and Organizations Related to GPS

7.8.1 The International GPS Service (IGS)

The IGS was established by the International Association of Geodesy (IAG) and started its activities formally on January 1, 1994, after a pilot phase of about 1 year. The IGS is a member of the Federation of Astronomical and Geophysical Data Analysis Services (FAGS) and it works in close cooperation with the International Earth Rotation Service (IERS). On January 1, 1999, the name of the service was changed from the original "International Global Positioning System (GPS) Service for Geodynamics (IGS)" to *International GPS Service* (IGS). The information in this section is mainly taken from IGS documents, like the "IGS Annual Reports" and the "IGS Directory". These documents are available from the IGS Central Bureau.

Following the *IGS Terms of Reference* (IGS, 2002a), "the primary objective of the IGS is to provide a service to support, through GPS data and data products, geodetic

and geophysical research activities."

The IGS collects, archives and distributes GPS observation data at a number of tracking stations. These data sets are used by the IGS to generate *data products*, namely

- high accuracy GPS satellite ephemerides,
- Earth rotation parameters,
- IGS tracking station coordinates and velocities,
- GPS satellite and tracking station clock information,
- ionospheric information, and
- tropospheric information.

In order to fulfill its tasks, the IGS has a certain structure (Fig. 7.100) with several components:

- Networks of Tracking Stations,
- Data Centers,
- Analysis and Associate Analysis Centers–Analysis Coordinator,
- Working Groups and Pilot Projects,
- Central Bureau, and
- Governing Board.

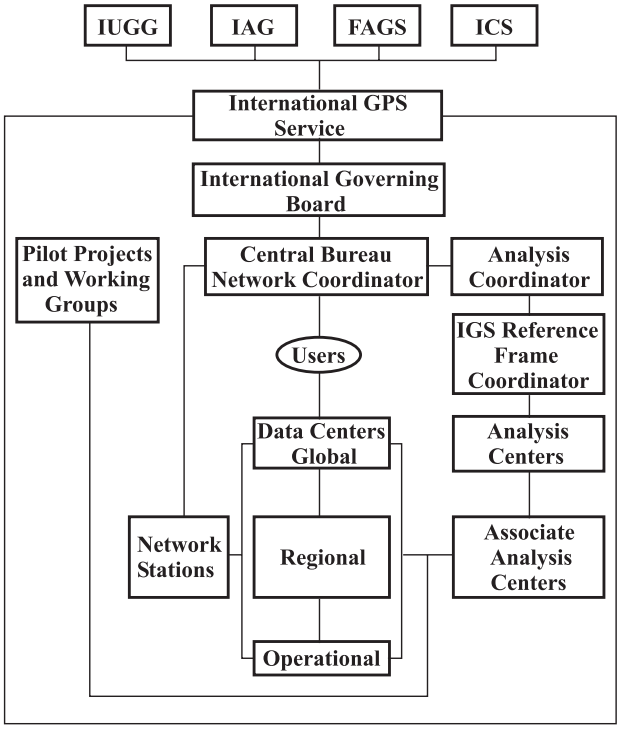

Figure 7.100. Structure of the IGS

The products of the IGS support scientific activities, such as

- improving and extending the International Terrestrial Reference Frame (ITRF),
- monitoring deformations of the solid Earth,
- monitoring Earth rotation,
- monitoring variations in the liquid Earth (sea level, ice sheets),
- determining orbits of scientific satellites,
- monitoring the high atmosphere (ionospheric tomography), and
- climatological research, contributions to weather prediction.

(a) *Network of Tracking Stations*

The global IGS network of permanent dual-frequency GPS tracking stations included more than 300 stations in 2002, representing some 200 agencies around the world. The number is still growing. Fig. 7.101 shows the global station distribution. The stations have to meet certain requirements, in particular they need data transmission facilities for a rapid (at least daily) data transfer to the data centers. The tracking data are analyzed by at least one Analysis Center or Associate Analysis Center. IGS stations which are analyzed by at least three IGS Analysis Centers, for the purpose of orbit generation, are called *IGS Global Stations* (numbering about 120 early in 2002). All IGS stations can be taken as reference stations for regional GPS analyses. The tracking data are available in RINEX format from the Data Centers. Approximately 90 IGS stations are producing hourly 30-seconds RINEX files, and about 35 stations are providing data in near real-time, delivering 15 minute, 1 Hz data files (Schmidt, Moore, 2002). The ensemble of IGS stations form the IGS network or *polyhedron*. In 2002, at about 50 stations GLONASS satellites were also tracked.

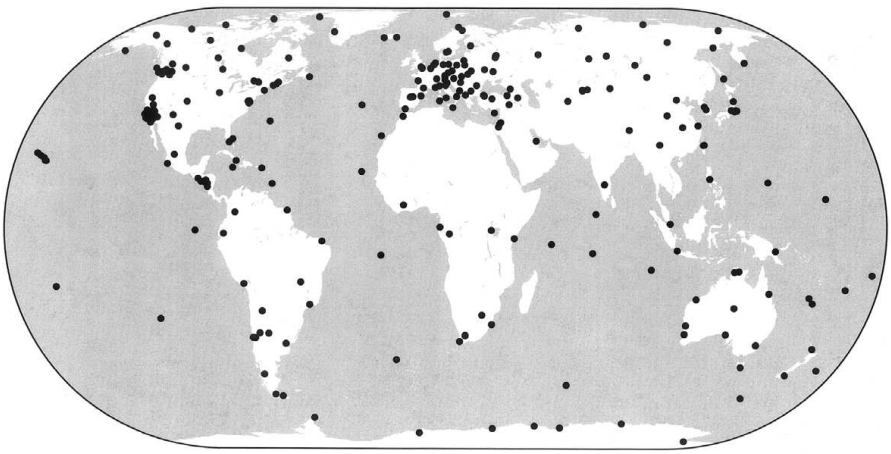

Figure 7.101. IGS Network 2002, source: IGS

(b) *Data Centers*

There are three categories of data centers: Operational, Regional, and Global Data Centers. The *Operational Data Centers* (ODCs), in total 25 at the end of 2000, are in direct contact to the tracking stations. They maintain the data in local archives, validate

and reformat the data (conversion to RINEX and data compression), and transmit them to a regional or global data center. *Regional Data Centers* (RDCs), in total five, collect reformatted tracking data from several data centers, maintain a local archive for users interested in stations of a particular region, and transmit the data to the Global Data Centers.

The *Global Data Centers* (GDCs) are the main interfaces to the Analysis Centers and to the general user community. Among other tasks they archive and provide on-line access to tracking data and IGS products. GDCs provide an on-line archive of at least 100 days of GPS data in the RINEX format, including the data from all global IGS sites. The GDCs also provide an on-line archive of derived products, generated by the IGS analysis or associate analysis centers. There is also on-line access to IGS products generated since the start of the IGS test campaign in 1992. The three Global Data Centers are

CDDIS	Crustal Dynamics Data Information System, NASA Goddard Space Flight Center, USA
IGN	Institut Géographique National, France, and
SIO	Scripps Institution of Oceanography, USA

(c) *Analysis Centers*

There are two categories of analysis centers: Analysis Centers (ACs) and Associate Analysis Centers (AACs). The *Analysis Centers* receive and process tracking data from one or more data centers and generate IGS products, as a minimum ephemerides, Earth rotation parameters, station coordinates, and clock information as well as other recommended products. The products are delivered to the Global Data Centers and to the IERS, using designated standards. The Analysis Centers are

COD	Center for Orbit Determination in Europe, University of Berne, Switzerland,
EMR	Geodetic Resources Division, Natural Resources Canada, Ottawa, Canada,
ESA	European Space Operations Center, European Space Agency, Darmstadt, Germany,
GFZ	GeoForschungsZentrum, Potsdam, Germany,
JPL	Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA,
NGS	National Oceanic and Atmospheric Administration/National Geodetic Survey, Silver Spring, Maryland, USA, and
SIO	Scripps Institution of Oceanography, University of California, San Diego, California, USA.

The *Associate Analysis Centers* produce specialized products, for example ionospheric maps or station coordinates and velocities for a global or regional sub-network. Of particular importance are the "Global or Regional Network Associate Analysis Centers" (GNAACs or RNAACs), producing weekly solutions of the global polyhedron or regional subsets thereof. Examples for RNAACs are the EUREF, NAREF or SIRGAS networks for Europe, North- and South America (Blewitt, 1998).

The *Analysis Coordinator* monitors the activities of the Analysis Centers. This person is also responsible for the appropriate combination of the Analysis Center products into a single set of products, in particular a single IGS ephemeris for each GPS satellite.

(d) *Working Groups and Pilot Projects*
A *Working Group* works on a particular topic. A *Pilot Project* has the objective to develop a particular IGS product or service relying on the IGS infrastructure. Active Working Groups and Pilot Projects by the end of 2002 were, for example,

- Reference Frame Densification Working Group,
- IGS/BIPM Time and Frequency Pilot Project,
- Ionosphere Working Group,
- Troposphere Working Group,
- International GLONASS Service Pilot Project,
- Low Earth Orbiter Pilot Project
- Real-time Working Group,
- Tide Gauge Pilot Project, and
- African Reference System (AFREF) Pilot Project

(e) *Central Bureau and Governing Board*
The *Central Bureau* (CB) is responsible for the general management of IGS and coordinates all IGS activities. The *Governing Board* (GB) sets the IGS policy and exercises a broad oversight of all IGS functions and components. The Central Bureau, for the time being, is located at the Jet Propulsion Laboratory in Pasadena, California and maintains an *IGS Information System* (CBIS), accessible at http://igscb.jpl.nasa.gov (World Wide Web). The CBIS contains information on the availability of and access to tracking data and IGS products, IGS orbits, Earth rotation parameters and other data. CBIS also gives access to IGS publications like annual reports and workshop proceedings. An overview of current IGS products is given in Table 7.31.

7.8.2 Other Services

A large number of international and national services provide information on GPS and other GNSS systems. Here only some indications are given. Note that web-addresses may change. For updated information see the regular "Almanac" pages in the August and December editions of *GPS World*.

Canadian Space Geodesy Forum, `http://gauss.gge.unb.ca/CANSPACE.html`
A service of the University of New Bruswick, Canada. Presents daily GPS constellation status reports and ionospheric disturbance warnings. News and discussion about GPS and other space-based positioning systems.

U.S. Coast Guard Navigation Center, `http://www.navcen.uscg.gov/gps`
GPS constellation status, almanac data, information on DGPS, Loran C.

U.S. National Geodetic Survey, `http://www.ngs.noaa.gov/GPS/GPS.html`
Precise and rapid orbits. General information on GPS.

Table 7.31. IGS products, status August 2002, source: IGS Central Bureau

	Accuracy	Latency	Updates	Sample Interval
GPS Orbits & Clocks				
Broadcast *)	∼260 cm / ∼7 ns	real time	–	daily
Ultra-Rapid	∼25 cm / ∼5 ns	real time	twice daily	$15^m/15^m$
Rapid	5 cm / 0.2 ns	17 hours	daily	$15^m/5^m$
Final	<5 cm / 0.1 ns	∼ 13 days	weekly	$15^m/5^m$
GLONASS Orbits				
Final	30 cm	∼ 4 weeks	weekly	15^m
Geocentric Coordinates IGS Tracking Stations				
Final horiz./vert. position	3 mm / 6mm	12 days	weekly	weekly
Final horiz./vert. velocity	2 mm / 3 mm per yr	12 days	weekly	weekly
Earth Rotation				
Rapid polar motion/ polar motion rates/ LOD	0.2 mas / 0.4 mas/d 0.029 ms	17 hours	daily	daily
Final polar motion/ polar motion rates/ LOD	0.1 mas / 0.2 mas/d 0.020 ms	∼ 13 days	weekly	daily
Atmospheric Parameters				
Final tropospheric	4 mm zenith path delay	< 4 weeks	weekly	2 hours
Ionospheric TEC grid	under development			
*) for comparison				

Geodetic Survey Division, Nat.Res. Canada; `http://www.geod.nrcan.gc.ca`
Provides access to data of the Canadian Spatial Reference System (CSRS) and Canadian Active Control System (CACS).

Scripps Orbit and Permanent Array Center (SOPAC); `http://sopac.ucsd.edu`
A service of the Scripps Institution of Oceanography, University of California. SOPAC provides precise, rapid, ultra rapid, hourly and predicted orbits for the IGS in SP3 format. Further SOPAC archives daily RINEX data from about 800 continuous GPS sites from various networks and arrays (IGS, SCIGN, CORS, EUREF and others). SOPAC is also the operational Center for the California Spatial Reference Center.

National Imagery and Mapping Agency (NIMA); `http://164.214.2.59/GandG/sathtml/`
Offers precise GPS orbit information based on tracking data from U.S. Air Force, NIMA, and IGS sites. Daily precise ephemerides and satellite clock estimates in SP3 format. Earth Orientation Parameter predictions.

National Mapping Division, Australia (AUSLIG); `http://www.auslig.gov.au/`
Comprehensive www-site with information on various GPS-related topics. Data from

the Australian Regional GPS Network. On-line GPS processing service for RINEX data.

Federal Agency for Cartography and Geodesy (BKG), Germany; `http://gibs.leipzig.ifag.de`

The information site "GIBS" is a service of the "Bundesamt für Kartographie und Geodäsie", Fankfurt/Leipzig. GIBS (GPS Informations- und Beobachtungssystem) provides almanac data, information on GPS status, datum transformations, satellite visibility, precise ephemerides, DGPS, GLONASS status and almanac data. The site also contains comprehensive information material on GPS and services like SAPOS.

For information on Galileo see

Galileo Homepage; `http://www.galileo-pgm.org`

Genesis Office; `http://www.genesis-office.org/`.

Genesis is a project providing support to the European Commission on its GALILEO activities. GENESIS communicates and disseminates information related to GALILEO and also distributes a "Galileo Newsletter".