Geomatics Engineering

Department of Geomatics Engineering

Development of a Fast Ambiguity Search Filtering (FASF) Method for GPS Carrier Phase Ambiguity Resolution

By

Dingsheng Chen

December, 1994

UNIVERSITY OF
CALGARY

Calgary, Alberta, Canada

ABSTRACT

A new concept for ambiguity resolution is introduced. Past techniques determined each ambiguity separately regardless of the assumed integers of the other ambiguities. These techniques considered only the partial relationship among ambiguity parameters themselves, and treated these ambiguities as fixed only if their correct values were known. In this study, the search ranges are determined recursively and are related to each other. To determine the uncertainty range of an ambiguity parameter, the effect of an assumed integer on other ambiguities is fully taken into account by constraining the ambiguities into integers. These constrained integers may be correct or incorrect. However, the incorrect integers are rejected later. All observations from the initial to the current epoch are taken into account by a least-squares filter. Furthermore, an index of the possible inability to fix ambiguities is used. Therefore, the full search of all possible integer ambiguities is not required and the computation time is dramatically reduced. Analysis of experimental results shows significant improvements in the time of ambiguity search and the number of epochs required to resolve the ambiguities. The reliability of the ambiguity resolution is also improved.

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOTATIONS

Conventions

$\hat{x}$ the least-squares estimate of the unknown vector x

C^{-1} the inverse of matrix C

δ in δx the correction on

k in x_k at epoch k

$nint(a)$ the nearest integer of a

A^T the matrix transpose of matrix A

$\dfrac{\partial f}{\partial x}$ the partial derivative of the function f with respect to x

$\varDelta$ the single difference between receivers in GPS

∇ the single difference between satellites in GPS

$\nabla \varDelta$ the double difference between satellites and receivers in GPS

Symbols

A the first order design matrix (the partial derivatives of the observation functions with respect to the state vector)

a_{ion} a constant in an ionospheric phase delay model

B the second order design matrix (derivatives of the observation functions with respect to observations)

b the length of the baseline

C the variance-covariance matrix

c the speed of light

$Cc_{k/k}, Ccv_{k/k}, Ccn_{k/k}, Cv_{k/k}, Cvn_{k/k}, Cn_{k/k}$

        the components of the variance-covariance matrix corresponding to the estimated state vector $\hat{x}_{k|k}$

$C\varepsilon c_{k+1/k}$, $C\varepsilon cv_{k+1/k}$, and $C\varepsilon v_{k+1/k}$

        the variance covariance matrices of the prediction noise for the coordinates and the velocity

C_{x_0} the covariance matrix of δx_0

δ the correction on the approximate value of x, i.e., $x = x_0 + \delta$

db the error in baseline b

$d_{ion}(t)$ the ionospheric delay

δ_{k+1} the correction to the predicted value, $\hat{x}_{k+1|k}$

dr the orbital error

dt the time difference between epoch $k+1$ and epoch k

$dT(t)$ the bias of the receiver clock

$dt(t)$ the bias of the satellite clock

d_{trop} the tropospheric delay

$d_{trop}(t)$ the bias of the tropospheric delay at time t

$d_{trop}{}^{\circ}$ the approximate tropospheric delay

$\delta\hat{x}$ and $\delta\hat{y}$ the variation of $\hat{x}$ and $\hat{y}$ from their approximate values, x_0, y_0

δx the correction to the approximated state vector where the observation equation is linearized (to be estimated)

$\delta\hat{x}_C$ the estimated correction of the parameters other than cycle ambiguities

$\delta\hat{x}_N$ the real estimated correction of ambiguity parameters

$\hat{\delta x}_{Nb/N_1\,N_2\,N_{i-1}}$ the real estimated values of $N_i,...,N_n$ under the fixed values of

     $N_1,N_2,...,N_{i-1}$

δx_0 the variation of the *a-priori* state vector from the point where the linearization

     is made

$\Delta\nabla\varepsilon_{ph}$ and $\Delta\nabla\varepsilon_{psr}$

     the measurement noise and other errors remained in the double differenced

     observations

$\Delta\nabla\phi_{AB}^{ij}, \Delta\nabla\rho_{AB}^{ij}, and\ \Delta\nabla N_{AB}^{ij}$ respectively the double differenced phase, geometric

     distance, and ambiguity between satellite i and j, and between stations A and B

$\Delta\nabla l_{ph}$ the double differenced phase observation

$\Delta\nabla l_{psr}$ the double differenced pseudorange observation

$\Delta\nabla\rho$ the double differenced geometric propagation distance from the satellites to the

     receivers

E the eccentric anomaly

e the eccentricity (dimensionless)

ε the unknown observation noise

ε_ϕ the noise in phase observation

$\varepsilon_{k+1,k}$ the uncertainty of the prediction model

ε_P the noise in code observation

$\phi(t)$ the carrier beat phase measurement in cycles

$\phi 1,\ \phi 2$ the phase observations corresponding to GPS *L1* and *L2*

$f1, f2$ the frequencies corresponding to GPS *L1* and *L2* carriers

I the unit matrix

λ the whole wavelength of carrier (for a squared type receiver, it is half of the carrier wavelength)

L1 and *L2*   two GPS carriers

$\lambda 1, \lambda 2$ the carrier wavelengths corresponding to *L1* and *L2*

m the number of constraints

μ the number of unknowns

N the integer carrier beat phase ambiguity

n the number of ambiguity parameters

N1, N2   the integer cycle ambiguities corresponding to *L1* and *L2* carriers

N_F the potential solution of the full ambiguity set

N_I the vector corresponding to the ambiguties constrained to integers

N^n the *n* dimensional integer lattice space

N_T the total electron content along the propagation path in electrons/m^2

p(t)   the carrier beat phase measurement in cycles

$\boldsymbol{P}_{CC}$, $\boldsymbol{P}_{CN}$, $\boldsymbol{P}_{NC}$, and $\boldsymbol{P}_{NN}$

     the sub-matrices of the normal matrix partitioned corresponding to $\boldsymbol{x}_C$ and $\boldsymbol{x}_N$

$\tilde{\boldsymbol{P}}_{Naa}$, $\tilde{\boldsymbol{P}}_{Nab}$, $\tilde{\boldsymbol{P}}_{Nba}$, and $\tilde{\boldsymbol{P}}_{Nbb}$

     the sub-matrices of normal matrix $\tilde{\boldsymbol{P}}_{NN}$, partitioned corresponding to $\hat{\delta \boldsymbol{x}}_{Na}$ and $\hat{\delta \boldsymbol{x}}_{Nb}$,

$\boldsymbol{r}$ the correction to the observations

ρ the satellite to receiver range

ρ(t)   the distance from the satellite to the receiver at time *t*

σ_0 the variance factor

s_{trop} the estimated tropospheric scale parameter

$\boldsymbol{u}_C$ and $\boldsymbol{u}_N$ the constant terms in the normal equations

$\hat{u}_{Na}$ and $\hat{u}_{Nb}$

    the constant terms in the partitioned normal equations corresponding to partitioned ambiguity parameters

Ω the sum of the squared residuals while estimating all ambiguity parameters as real values

$\Omega_{|\boldsymbol{x}_N = N_F}$ the adjusted Ω corresponding to fixed ambiguity values

$\Omega_{|\boldsymbol{y}}$ the weighted sum of residuals while treating $\boldsymbol{y}$ as a known constant

$\boldsymbol{w}$ the misclosure after the linearization

$\hat{\boldsymbol{x}}$ the remaining part in the state vector

ξ the scale from the computed variance to the maximum uncertainty

$\boldsymbol{x}$ the unknown parameter vector to be estimated

$\boldsymbol{x}_{k+1}$ state vector at epoch $k+1$

$\hat{\boldsymbol{x}}_{k+1/k}$ state vector at epoch $k+1$ from all the informaion up to epoch k

$\boldsymbol{x}_{k_o}$ the point of expansion in the Taylor's series of the prediction model; normally it takes the value of $\hat{\boldsymbol{x}}_{k|k}$

$\hat{\boldsymbol{x}}_{k/k}$ the state vector at epoch k from all the informaion up to epoch k

$\hat{\boldsymbol{x}}_N$ the least-squares estimated value of ambiguity

$\boldsymbol{x}_{Na}{}^o$ the approximate value of $\boldsymbol{x}_{Na}$

$\hat{\boldsymbol{x}}_{Na}$ the first part of the partitioned ambiguity parameters

$\hat{\boldsymbol{x}}_{Nb}$ the second part of the partitioned ambiguity parameters

$x_{N_i/N_1,N_2,...,N_{i-1}}{}^{max}$

    the maximum possible value of N_i if the integer values of N_1, N_2, ..., N_{i-1} are correct

$x_{N_i/N_1,N_2,...,N_{i-1}}{}^{min}$

the minimum possible value of N_i if the integer values of N_1, N_2, ..., N_{i-1} are

correct

x_o and y_o the approximated values of x and y

$\hat{x}_{|y}$ the Kalman estimation of x corresponding to the known value of y

$\hat{y}$ the float estimation state vector component whose value could be fixed

y_C the fixed value of $\hat{y}$

Acronyms

AFM	Ambiguity Function Method
AS	Anti-Spoofing
C/A-code	Coarse/Acquisition-code
CACS	Canadian Active Control System
CIGNET	Cooperative International GPS Network
DD	Double Difference
DGPS	Differential GPS
DLLs	Delay Lock Loops
DoD	Department of Defense
DOT	Department of Transportation
FARA	Fast Ambiguity Resolution Approach
FASF	Fast Ambiguity Search Filter
GPS	Global Positioning System
JPO	Joint Program Office
IGS	International GPS Geodynamics Service
NATO	North Atlantic Treaty Organization
OTF	on-the-fly
P-code	the precision code
ppm:	part(s) in per million
PPS	Precise Positioning Service
PRN	Pseudo-Random Noise,
RCSR	Recursive Computation of Search Range,
SA	Selective Availability,
SPS	Standard Positioning Service

CHAPTER 1   INTRODUCTION

1.1   Global Positioning System (GPS) and Ambiguity Resolution

The Navstar Global Positioning System is being developed primarily for the military forces of the North Atlantic Treaty Organization (NATO) for world-wide, real time positioning and continuous navigation. The initial intention of GPS was mainly for navigation of the US military. Due to the tremendous potential of the system and the latest improvements in receiver technology, a growing community is using GPS for a variety of civilian applications (navigation, geodetic positioning, etc.) [Wells et al., 1987].

Many applications require an accuracy no more than 1 m, e.g., navigation in open sea, en-route aircraft navigation, and fleet monitoring [Trimble Navigation, 1989]. However, as will be discussed later, sub-metre and even centimetre-level accuracy is required in many applications.

To achieve centimetre-level accuracy, carrier phase measurements have to be used. The carrier phase measurement is the most precise positioning signal obtainable from GPS. The measurement can be converted into a precise distance between the receiver and the satellite. The position of the receiver can then be computed using the distances from the receiver to different GPS satellites. However, a receiver can measure only the fractional part of the phase and its variation over time. There is a constant unknown, called initial-cycle ambiguity (simply called ambiguity), in every phase measurement. This ambiguity

has to be resolved before the carrier phase measurement becomes an accurate geometric range between the receiver and the satellite. Once all the phase ambiguities are resolved correctly, accurate positioning at the centimetre-level will be readily achievable using at least four satellites.

Double-difference observations between satellites and two receivers are often used. In this case, one receiver acts as a base station (or sometimes called a reference or monitor station). The position of the other receiver (also called remote station or rover) with respect to the base station can then be determined. The corresponding ambiguity parameter is also double-differenced but it is still called the ambiguity.

1.2 Applications of Precise GPS Positioning

Ambiguity resolution is very important for many applications of instantaneous precise positioning. The determination of the ambiguity parameters while the remote receiver is moving is called on-the-fly (OTF) ambiguity resolution. It is also very important for many static surveying projects where only a short period of occupation on the station is allowed. In the following, some existing and potential applications are outlined.

For many geodetic applications, centimetre-level accuracy can be achieved by occupying a site for an extended period of time without integer ambiguity resolution. However, this accuracy is difficult when prolonged occupation is impossible and ambiguity parameters are not resolved.

If ambiguities can be resolved within a few minutes, or even a few seconds, the productivity of surveying can be improved significantly. If on-the-fly ambiguity resolution can be achieved correctly before setting up a GPS antenna, then only one epoch of observations is sufficient to get positioning accuracy to the 2 cm level. This can be

2

especially useful for surveying where fast positioning is very important. Even for prolonged occupation of a site, the resolution of ambiguities will always improve the accuracy of positioning.

As GPS receivers are becoming more affordable, GPS surveying may be extensively used and form the basis for efficient land information systems. GPS will provide many services to the general public. Precise positioning with GPS will be used in precise geodetic surveying, deformation surveys, surveys for detecting and monitoring tectonic movements, geodynamics, etc. [Wells et al., 1987].

Ambiguity resolution is also required in attitude and heading determination using GPS. In these applications, at least two or more antennas must be mounted on a platform (car, ship, plane, etc.). If ambiguities between any two antennas can be resolved, the vectors between antennas can be determined at the centimetre-level. The heading and attitude can then be derived from relative positions of these antennas [Cohen and Parkinson, 1991; Cannon and Haverland 1993; Lu et al., 1994]. At least one baseline is needed to determine heading and at least two baseline vectors are needed to determine the attitude.

GPS antennas and receivers can also be mounted on a spacecraft so that position and attitude information of the spacecraft can be derived [Brock et al., 1994]. Possible applications include altimetry satellites, remote sensing satellites, communication satellites, space ships, etc.

In airborne photogrammetry, if GPS units are mounted at both the ground and on-board the aircraft, precise positions of the camera at each exposure time can be determined if the ambiguities are resolved. Therefore, the number of conventional ground control points can be reduced or even completely eliminated [Hintz and Zhao, 1989; Cannon,

1991]. Similarly, GPS precise positioning can be applied to airborne remote sensing [Schwarz et al., 1994]. Lachapelle et al. [1994a] also experimented with precise aircraft-to-aircraft positioning.

For aircraft navigation, high accuracy is required for landing, especially for automatic landings. Precision GPS may help in reducing risk during the landing[Davis, 1993].

In marine applications, precise GPS can be used to monitor ocean tides, position the points on the sea floor (sea floor geodesy), verify satellite altimetry, determine water levels [Lachapelle et al. 1994c], study sea level variations, and assist with dredging operations [Lachapelle et al. 1993b].

Precise kinematic differential GPS will also be useful in navigating agricultural vehicles. It can play a role in the distribution of work, navigation of the harvesters, and the guidance of tractors, Lachapelle et al. [1994b].

1.3 Previous Studies

Different strategies can be applied to fix ambiguity parameters. A combined strategy is often used. The next sub-section describes techniques in terms of different positioning arrangements. In the second sub-section, ambiguity resolutions in terms of data processing techniques are introduced. In the final sub-section, major developments in the ambiguity search are presented.

1.3.1 Ambiguity Resolution by Arrangement of Occupation

1.3.1.1 Semi-Kinematic Positioning

In semi-kinematic positioning, the remote receiver first rests on a static point. It starts moving once ambiguities are resolved[Cannon, 1990].

One method to resolve ambiguities is prolonged occupation over a static point. This is the oldest method of GPS surveying. After the remote site has been occupied for an prolonged period of time, the satellite geometry changes. The changed geometry leads to precise positioning and accurate estimation of real valued ambiguity parameters (called float ambiguities). Here, the float ambiguity parameters converge and can be rounded off to integers, provided the distance between the reference and the remote station is relatively short. These integers are, therefore, regarded as the correct values.

The disadvantage of this method is its low productivity. Because the geometry of a receiver with respect to the GPS satellites changes slowly, long occupation times are required so that float ambiguities converge to the correct integer values.

A known precise station can also be used to speed up the ambiguity resolution. Here, the remote antenna is first mounted on a point whose relative position with respect to the reference is precisely known. Therefore, the carrier phase ambiguities can almost be computed directly from just a few epochs of observations as long as the effect of errors is well below one cycle of wavelength. Once ambiguities are resolved, the antenna can be transfered to a moving platform or other survey points. The requirement of a precise position is the limitation of this method because obtaining a precise position is often difficult.

All cycle slips must be recovered for the entire kinematic positioning process. To fix the cycle slips of single frequency observations, at least four satellites should be free from cycle slips, otherwise the remote antenna has to go back to a static point. In many cases, it is difficult to start from a fixed point. A short baseline is often used to reduce the effect of un-modeled errors and speed up the ambiguity resolution.

1.3.1.2 Antenna Swapping

This technique was first introduced by Remondi [1986]; see also [Hofmann-Wellenhof and Remondi, 1988]. To describe this technique, two antennas are labeled as 1 and 2, and two stations are labeled as A and B. First, the antenna 1 is mounted at station A and antenna 2 is mounted at station B. Their locations are then exchanged (Figure 1.1). After simultaneously observing at both stations, antenna 1 is moved to station B, and antenna 2 is moved to station A. Two groups of double difference observations are then obtained.
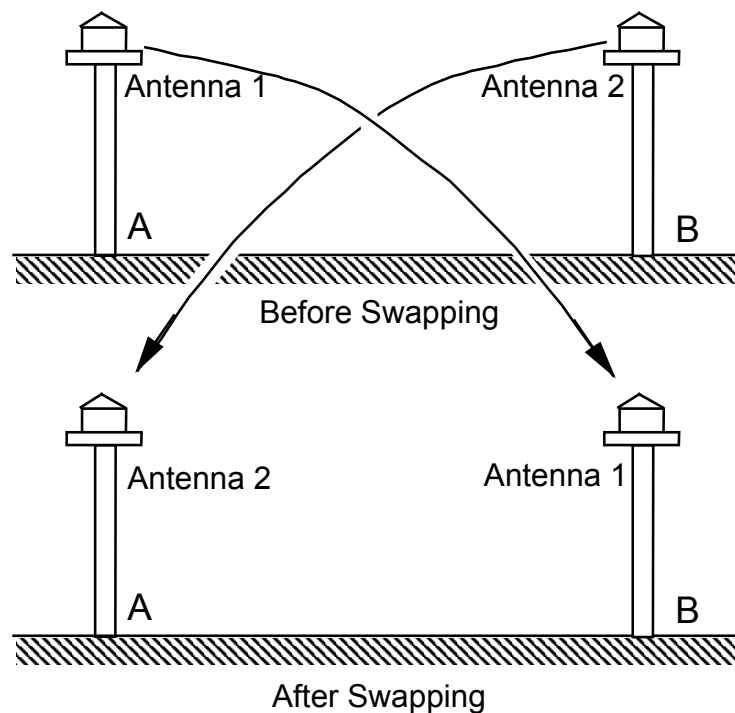
Figure 1.1   Concept of Antenna Swapping

It is assumed that, before antenna swapping, the double difference phase observation is:

$$\Delta\nabla\phi\,^{ij}_{AB} = \Delta\nabla\rho^{ij}_{AB} + \lambda\Delta\nabla N^{ij}_{AB}$$

where,

$\Delta\nabla\phi\,^{ij}_{AB}$, $\Delta\nabla\rho^{ij}_{AB}$, and $\Delta\nabla N^{ij}_{AB}$ are double difference phase, geometric distance, and ambiguity respective between satellites i and j, and between stations A and B. After antenna swapping, the double difference observation can be expressed as:

$$\Delta\nabla\phi'\,^{ij}_{AB} = \Delta\nabla\rho'\,^{ij}_{AB} + \Delta\nabla N'\,^{ij}_{AB}\,\lambda \tag{1.2}$$

If no cycle slips occur during the antenna swapping, the following relationship holds true:

$$\Delta\nabla N'\,^{ij}_{AB} = -\Delta\nabla N^{ij}_{AB} \tag{1.3}$$

Equation (1.2) then becomes

$$\Delta\nabla\phi'\,^{ij}_{AB} = \Delta\nabla\rho'\,^{ij}_{AB} - \Delta\nabla N^{ij}_{AB}\,\lambda\,. \tag{1.4}$$

By adding equations (1.1) and (1.4), the following can be obtained:

$$\Delta\nabla\phi'\,^{ij}_{AB} + \Delta\nabla\phi^{ij}_{AB} = \Delta\nabla\rho'\,^{ij}_{AB} + \Delta\nabla\rho^{ij}_{AB}\,. \tag{1.5}$$

On the other hand, by subtracting equation (1.1) from equation (1.4), the following can be obtained:

$$\Delta\nabla\phi'\,^{ij}_{AB} - \Delta\nabla\phi^{ij}_{AB} = \Delta\nabla\rho'\,^{ij}_{AB} - \Delta\nabla\rho^{ij}_{AB} + 2\,\Delta\nabla N^{ij}_{AB}\,\lambda\,. \tag{1.6}$$

Since equation (1.5) contains no ambiguities, the precise vector between the two stations can be determined precisely by using only a few pairs of observations before and after antenna swapping. Once the position is known, the real valued ambiguities can be computed directly from equation (1.6). Since these values are very precise, rounding to the nearest integers will normally result in the correct integer ambiguities.

1.3.1.3 Kinematic Rapid Positioning Forming a Closed Loop

In this method, the remote unit starts at a point, and then moves to other points of interest. At each occupation, the collection of just one or a few epochs is necessary. After the satellite geometry changes significantly, the receiver returns to the start point. This changed geometry causes a higher constraint in ambiguity resolution and makes the resolution easier. After ambiguity resolution, the precise positioning of other observed points on the closed loop can be determined. A suitable method is still necessary to resolve the ambiguities. The disadvantage of this method is the difficulty in guaranteeing continuous phase lock [Kleusberg, 1990].

All the methods described above suffer from a common problem, the difficulty in maintaining phase lock. Therefore, operational flexibility is restricted. As a result, points would have to be close together throughout the survey mission to guarantee a fast recovery if satellite signals are occasionally lost.

1.3.1.4 Rapid Static GPS Surveying

In this surveying model, each point is occupied for only a short period of time to speed up productivity. Due to the short period of the time, float estimation of the ambiguities may not converge to integer values. Therefore, an ambiguity search method or precise range information should be applied [Kleusberg, 1990].

1.3.1.5 Multiple Occupations of the Same Point

This technique requires re-occupation of points but the continuous tracking during any two consecutive occupations is not required. However, the time gap between two occupations should be large enough to guarantee sufficient geometry in ambiguity

resolution. If all the points can be re-occupied in an orderly manner, for example, if the re-occupation occurs in the same order as the previous one, then this method can improve the productivity of many positioning tasks. This method is also called pseudo-kinematic surveying [Kleusberg, 1990]. The float estimated ambiguities normally do not converge to integers and a proper ambiguity search method is still required in this method.

1.3.2   Ambiguity Resolution Methods in Terms of Data Processing Techniques

1.3.2.1   Classical Convergence of Real Estimated Ambiguity Parameters

One of the earliest ambiguity resolution methods used was to estimate float ambiguities and the float estimations are rounded to the integers if they are close to integers [Langley et al. 1984]. As discussed in Section 1, prolonged occupation on a static point is required so that the changed satellite geometry results in more effective integer ambiguity convergence.

When coordinates are determined very accurately, the real estimated ambiguities may be accurate enough to be rounded to integers. Once the correct ambiguities are obtained, the positioning accuracy can be further improved.

The classical method of rounding to the nearest integers has been improved by many authors. Blewitt [1989] processed undifferenced data, and then formed double-difference ambiguities. His basic assumption was that when some of the ambiguities converge to their correct integers, they can be removed from the estimation; therefore, the geometry of the remaining ambiguities becomes stronger. The variance-covariance matrix was used to select the best determined ambiguity, according to his optimal double-difference transformation. He also used precise code measurements to assist ambiguity resolution. The most likely fixed ambiguity was resolved and removed from the estimation. Then,

the next most likely fixed ambiguity was sought. His criterion of the most likely fixed ambiguity was based on subjective statistics and is only good if some ambiguities converge to the correct integers. Otherwise, an incorrect ambiguity will be selected without trying other alternatives. Dong and Bock [1989] proposed a technique for ambiguity resolution in a network. In this technique, the ambiguities of a few baselines converge at first, and then ambiguities for other baselines converges sequentially. However, this technique offers no help for rapid static and kinematic positioning of a single baseline.

Talbot's approach [1991] is similar to that proposed by Blewitt. The difference is that he tried ambiguity resolution epoch by epoch, sequentially. He also did not use subjective statistics to fix an ambiguity into an integer when it did not converge to an integer with high confidence; he assumed that some of the ambiguity parameters should converge to integers first.

In static applications, it takes a long time before any real estimated ambiguities converge to integers. Therefore, the method is not suitable for detailed precise surveying where many points are to be occupied, or in other areas where only a short period of occupation is allowed. Therefore, an effective ambiguity search algorithm is still very important.

1.3.2.2 *The Combination of Different Observations and the Use of Precise Code*

As will be discussed in Chapter 2, GPS has two primary radio carrier frequencies, namely, L1 and L2. The two fundamental types of GPS observations are carrier phase and code measurements. From these observations, linear combinations between observations of the same type can be formed to resolve ambiguities. These linear combinations are mostly used to reduce, or even eliminate certain biases, or to increase

the wavelength of the combined observation [Wübbena, 1989]. The following criteria are normally considered in forming a combined observation,

- integer ambiguities,

- reasonably large wavelength to help ambiguity fixing,

- low ionosphere influence,

- limited observation noise.

Two important linear combinations useful for ambiguity resolution are the wide and narrow lane combinations [Wübbena, 1989]. The wide lane observation is formed by subtracting the L1 observation from the L2 observation and the narrow lane observation is formed by adding the L1 and L2 observations. Note that the unit of measurement when forming wide and narrow lane observations is cycles. The advantage of the wide lane observable are its long wavelength (≈86 cm, which is 4.5 times larger than L1 wave) which is favorable for ambiguity resolution. The advantage of the narrow lane observable is its low noise level (about half of L1 or L2). However, the wavelength of the narrow lane observable is only about half of that of L1. If potential ambiguities of both wide and narrow lanes can be resolved, the potential ambiguities for individual frequencies can be resolved using an even-odd condition [Seeber, 1993].

This property was further exploited by Wübbena [1989] in a technique called extra-widelaning. Besides wide and narrow lanes, Wübbena used a smoothed narrow lane pseudorange and ionospheric delay combination. The phase measurements were also used to derive the narrow lane ambiguities. Obviously, the method can not be applied to single frequency observations.

The measurement of very precise code (cf. Chapter 2) can facilitate ambiguity resolution. If the pseudorange can be measured to a precision of 10 to 40 cm, this will allow

ambiguity determination to occur more rapidly [Euler and Goad, 1991], [Cannon and Lachapelle, 1992], [Lachapelle et al. 1993ᵃ]. However, the availability of P-code receivers is reserved for military users and are not available to the general public. Although the precision C/A receivers are available they are still affected by relatively high noise, such as multipath signals and ionospheric delay. As a result of this noise, the uncertainty ranges for ambiguity parameters are substantially increased.

The fact that many receivers based on single frequency and wide correlator are still being used demands good search method to resolve ambiguities. Even for dual frequency and narrow correlator receivers, float ambiguities may not close to their correct integers and a search scheme is needed to resolve integer ambiguities.

1.3.3   Methods Based on Ambiguity Search

As will be shown in Section 3.1, prolonged occupation over a single point is often difficult, and the float ambiguities can not be rounded to their nearest integer values. Here ambiguity search techniques may be required to get centimetre-level precise positions.

Various methods have been developed for OTF ambiguity resolution. The search process for integer ambiguities is performed by applying certain validation and rejection criteria to the estimated ambiguities or positions. The methods developed for OTF can also be applied for rapid static position, multiple-occupation, and so on.

One of the search criteria for ambiguity-fixing is based on the minimization of the sum of the squared observation residuals (simply called sum of residuals). The ambiguity set that minimizes the sum is fixed as the correct one. Other insurance measures are often taken to make the results more reliable, such as the ratio test [Abidin, 1991; Lu et al. 1994].

Many ambiguity search algorithms are available to minimize the sum of squared residuals. A simple technique is to search through all the possible integer ambiguity sets in the uncertainty ranges. However, such a method consumes enormous computation time, even for a minimal 1 m uncertainty in the position components.

Loomis [1989] used a bank of Kalman filters, each of corresponding to a set of potential ambiguities. Unfortunately, the technique did not work well. The technique did take into account the dependence between ambiguity parameters and was computationally intensive.

Hwang [1989], in a variation of the Loomis approach, recognized that three ambiguities were sufficient to solve the problem. He simply picked the three ambiguities that independently converged the most quickly.

Cohen and Parkinson [1991] and Brown [1992] have also proposed that an OTF ambiguity resolution technique be used specifically for attitude determination of a moving platform using GPS carrier observations. However, the technique cannot be generalized.

Abidin [1991] has proposed an integrated strategy for OTF ambiguity resolution. However, many of the criteria he used almost overlap. As a result, the technique can have only marginal improvement.

Frei and Beutler [1990] proposed the Fast Ambiguity Resolution Approach (FARA). The major characteristic in this approach is that the differences of any two ambiguities in a potential ambiguity set should be included in the corresponding confidence regions of their real estimation. However, the improvement in the computation time is not significant over the full search method, because, when a set of ambiguities is satisfied in

their confidence regions, the difference of any two ambiguities from the set only marginally exceeds the confidence region.

Hatch [1991] proposed an elegant, yet relatively simple approach. Instead of searching through all the possible integer ambiguity sets, he divided the ambiguity parameters into two groups: primary ambiguities (typically three ambiguities); and the secondary ambiguities. Only the primary ambiguities are fully searched. For each set of the primary ambiguities, there is a unique set of secondary ambiguities. Therefore, the search dimension is smaller and the computation time is significantly shorter than the full search approach.

Hatch's method also involves a modified sequential least-squares technique. It searches through all possible integer ambiguity sets at the first epoch. In subsequent epochs, those which do not fit with the data are rejected until only one ambiguity set (or only a few sets) is left. Then, if the ratio test (the second largest sum of squared residuals divided by the minimum sum) exceeds a certain threshold, the ambiguity set corresponding to the minimum is accepted.

Hatch's method is one of the most popular algorithms. However, the ambiguity determination for secondary satellites is not very reliable, and there is the possibility of excluding the correct set of ambiguities at the first epoch. This will be especially true if the geometry of the primarily satellites is poor and non-random errors, such as multipath biases, are large.

Landau and Euler [1992] employ an optimized Cholesky decomposition algorithm (herein, it is simply called Landau and Euler's approach). However, the method is still slow for real time applications, when the uncertainty region is large, or when the data

collection rate is higher than 1 Hz. Furthermore, the time required for ambiguity search is much higher during the initial epochs.

Ober [1993], Teunissen [1993], and Teunissen et al. [1994] applied a lattice transformation [Dieter, 1975, Pohst and Zassenhaus, 1989] on the integer ambiguity parameters. This application is called Z-transformation by the authors. This transformation is advantageous when dual frequency observations are used without explicitly forming widelane observations. By applying the integer transformation on the ambiguity parameters, the correlation between estimated ambiguity parameters is greatly reduced when strong correlation exists between estimated ambiguity parameters. Although the technique will decrease the computational effort in some specific cases, it has no impact on the observation period required to resolve the ambiguities nor on the reliability of the solution.

An ambiguity resolution method similar to the one described by Chen [1993], and further developed in this thesis, was presented by Landau and Vollath [1994]. Although a different formulation is used, the same concepts as proposed by Chen [1993] are utilized.

Another criterion for ambiguity-fixing is the Ambiguity Function Method (AFM) as first described by Counselman and Gourevitch [1981] for static positioning. Mader [1992] applied the technique for OTF ambiguity resolution. However, the AFM is nearly identical to the minimization of the sum, if all the cycle slips are properly handled. The equivalence was discussed by Lachapelle et al. [1992[b]]. Therefore, only methods related to the minimization of the sum of residuals will be compared with the method developed in this thesis.

1.4   The Scope of the Research

As discussed previously, there is an increasing demand for rapid static and kinematic surveying. Even for conventional static surveying, ambiguity resolution will always increase the accuracy of GPS positioning. The key to precise and rapid static surveying is ambiguity resolution.

The exact determination of the ambiguities is a crucial issue for precise GPS positioning and navigation. As will be discussed in the next chapter, the solution of these integer ambiguities is not an easy task. It is limited by various error sources and their patterns, the satellite geometry with respect to the receiver's antenna, and the limitations of ambiguity resolution techniques currently available.

While significant research has been conducted in the area of ambiguity resolution, many problems still remain. Three of the most notable problems are computational speed, observation time required to resolve the ambiguities, and reliability. Unfortunately, techniques for efficient and reliable ambiguity resolution are not yet satisfactory, especially in rapid and kinematic positioning.

The objective of this research is to improve integer ambiguity resolution. The following factors should be considered for an ambiguity resolution algorithm:

- It should be fast enough to be installed in the most common applications.

- It can be applied to both the static and kinematic environments.

- It should be fast enough for both initial search and subsequent searches.

- It should be reliable.

With the above objectives, this dissertation is outlined as follows:

Chapter 2 describes basic GPS observables, error sources, and the effect of the errors on the positioning. Methods on reducing or even eliminating these effects are discussed.

Chapter 3 provides the background of least-squares filtering and develops formulas that can be applied in ambiguity searching.

Chapter 4 presents the method proposed by the author, namely the Fast Ambiguity Search Filter (FASF). The fundamental concepts of FASF are developed in the chapter. To implement these concepts, the least-squares filtering technique is used.

Chapter 5 examines four sets of experimental data which are used to test the effectiveness of FASF and its least-squares implementation. The data sets originate from land, marine, and air experiments of GPS. The computational efficiency, the observation time required to resolve ambiguities, and the repeated ambiguity resolutions are presented. The comparisons with the results from FLYKIN™'93 (a software initially developed according to a modified Hatch's method) are also shown.

The final chapter makes conclusions and suggestions for further developments.

CHAPTER 2   GPS OBSERVABLES AND ERROR SOURCES

Various errors in observations affect position accuracy and also affect the ambiguity resolution. Following the description of GPS observations in Section 2.1, the handling of biases in GPS observations is discussed. The effect of selective availability is briefly discussed in Section 2.3.

2.1   GPS Observations

Three different types of positioning information can be extracted from the signals of a GPS satellite, namely: code (pseudorange), carrier phase, and phase rate (also called Doppler frequency).

2.1.1   Code Measurement (Pseudorange)

A pseudorange is the measurement of the time shift required to align a replica of the GPS PRN code, generated in a GPS receiver, with the code transmitted from a GPS satellite. If the receiver clock is fully synchronized with the GPS time, then the time delay between the transmission and the reception is exactly the travel time of the signal. This delay can be converted into the travel distance of PRN signals between the satellite and the receiver's antenna. However, since satellite and receiver clocks are not synchronized, the range determined in this procedure contains a clock error. The range is, therefore, referred to as the pseudorange. Two types of code are available for GPS signals: P-code (Y-code) and C/A code [Wells et al. 1987].

The pseudorange equation can be written as:

$$p(t) = \rho(t) + d_{orb} + c\{dt(t) - dT(t)\} + d_{trop}(t) + d_{ion}(t) + \varepsilon_P$$

where,

$p(t)$ is the code measurement in metres,

$\rho(t)$ is the distance from the satellite to the receiver in metres,

d_{orb} is the orbital error in metres,

c is the speed of light in metres/second,

$dt(t)$ is the bias of the satellite clock in seconds,

$dT(t)$ is the bias of the receiver clock in seconds,

$d_{trop}(t)$ is the bias of the tropospheric delay in metres,

$d_{ion}(t)$ is the bias of the ionospheric delay in metres, and

ε_P is the measurement noise in metres.

In the ambiguity search procedure, the role of the code measurement is to determine the search space.

2.1.2   Carrier Beat Phase Measurement

The carrier phase can also be measured by beating the received Doppler-shifted satellite carrier with a signal of constant frequency generated in a GPS receiver. The carrier transmitted by a satellite can be extracted either by complete knowledge of the Pseudo-random Codes (C/A-code or P-code), or by codeless signal processing techniques, such as squaring, or cross-correlation. Since a receiver can only measure the fractional part of the beat carrier phase, the integer number of whole wavelengths in every phase measurement is unknown. This integer number is called initial carrier phase ambiguity. If the initial phase ambiguity could be resolved reliably, the phase would be used as the most accurate distance measurement from a GPS satellite.

The carrier phase equation can be written as:

$$\Phi(t) = -\lambda \, \phi(t) = \rho(\tau) + d_{orb} + c \, \{dt(t) - dT(t)\} + d_{trop}(t) - d_{ion}(t) + \lambda \, N + \varepsilon_\phi \, ,$$

where,

$\Phi(t)$ is the carrier phase measurement in metres,

$\phi(t)$ is the carrier phase measurement in cycles,

λ is the carrier wavelength in metres,

N is the integer carrier phase ambiguity, and

ε_ϕ is the measurement noise in metres.

The definitions of the other symbols are the same as in the pseudorange observation.

Double-difference observations are often used. That is, the observations are first differenced between different satellites. Then, these differenced observations are further differenced between the receivers. The advantage of the double difference is that it greatly reduces or eliminates the effects of many errors discussed in Section 2.

2.1.3   Doppler Measurement

A Doppler measurement is the measurement of the instantaneous rate of the GPS carrier phase, i.e., the instantaneous Doppler frequency shift of the incoming carrier. The shift is caused by the relative motion between the receiver and the satellite. The major role of the Doppler measurement is in velocity estimation. In kinematic positioning, it can also be used for roughly detecting and estimating cycle slips (cf. 2.2.9).

2.2   Error Sources in GPS Positioning

Noise and biases in GPS positioning can be grouped into three categories: station dependent biases, observation dependent biases, and satellite dependent biases [Wells et

al. 1987]. Station dependent biases are errors in the base station coordinates and in the receiver clock. Observation dependent biases include the signal propagation delays in the ionosphere and troposphere, carrier phase cycle slips, receiver measurement noise, and multipath. Satellite dependent biases cover errors in the satellite orbit and satellite clock biases.

2.2.1   Measurement Noise

Typical pseudorange measurement resolution is approximately 1 m on a C/A-code and 0.3 m on a P-code. Currently, many receivers can achieve higher accuracy. For example, the NovAtel GPSCard™ use a technique called Narrow Correlator™ in the receiver Delay Lock Loops (DLLs) [Van Dierendonck et al. 1992]. This technique allows GPS receivers to measure the pseudorange at 10 cm noise level on C/A code.

The carrier phase can be measured with millimetre or sub-millimetre precision. The measurement error tends to decrease as the number of observations increases. However, other error sources, such as multipath, are normally much larger than the measurement noise (cf. 2.2.8). In looking at the issue of measurement quality, what is important is the stability of the phase measurements. If a receiver has frequent cycle slips, even worse, half cycle slips and abnormal observations, ambiguity resolution will be very difficult.

2.2.2   Orbital Biases

Orbital errors result from the uncertainties in the orbital information. These uncertainties are due to the accuracy limitations associated with the predicted nature of the broadcast ephemeris and the SA policy instituted by the DoD (see section 2.3). Tests have shown that the orbital error is generally 10 to 50 metres. Under Selective Availability (SA, cf. 2.3), the orbital errors can exceed 100 metres in extreme cases [FRNP, 1990]. Since most

users derive the positions of GPS satellites from the broadcast ephemeris, the uncertainties of the broadcast ephemeris result in errors in positioning.

The orbital error can be greatly reduced in relative positioning by differencing observations between receivers. However, the residual orbital bias increases as the baseline length creases. Therefore, a more effective way to handle orbital bias for a long baseline positioning is to use post-processed precise orbits. Precise orbits require an extensive monitoring network, complicated force modeling and parameter designing [Chen, 1991, Delikaraoglou et al. 1990]. These orbits are currently not available for real time applications.

The effect of the orbital error is as follows, [Vaníče k et al. 1985]:

$$db = dr\, b\, /\, \rho\, ,$$ (2.1)

where,

   db is the error in the baseline,

   b is the length of the baseline,

   dr is the orbital error, and

   ρ is the satellite-receiver range.

The effect of the orbital errors on relative positioning is given in Table 2.1.

Equation (2.1) implies a considerable approximation. A geometry analysis by Chen and Langley [1990[a]] has shown that the popular belief that the above estimation is pessimistic is not always true. In the case of poor geometry in a GPS satellite configuration, the error could be larger than shown in the above equation. This is especially true if only a few epochs of observations are used and the number of tracked satellites is limited. However, if many redundant satellites are tracked or if a prolonged period of observation on a

stationary point is carried out, the orbital errors will tend to reduce by averaging. Only in this case, equation (2.1) can be regarded as a pessimistic estimation. Therefore, equation (2.1) should be used with caution and only for rough estimation. Nevertheless, the equation is convenient.

Table 2.1   Effect of the Orbital Error on Relative
Positioning

Orbital Error (m)	Relative Accuracy (*ppm*) (Assumed ρ=20,000 km)
100	5
20	1
2	0.1

If the orbital error is 20 m and the baseline length is 10 km, the corresponding bias in a double difference phase observation is estimated at one centimetre. That is comparable to multipath effect. Over a longer baseline, the effect of the orbit will increase and the ambiguity resolution will be more difficult. Orbital error is one of the major factors that limit the length of the baseline on which ambiguities can be resolved.

2.2.3   Error in Base Station Coordinates

In differential positioning, the uncertainty of the reference coordinates with respect to the WGS-84 coordinate system will directly propagate to the estimated coordinates of the remote station. The uncertainty has a reduced impact on relative positions (i.e., reduced effect on the coordinate differences between the remote and the base station). The effect is similar to those of orbits, but it is more systematic [Chen and Langley, 1990]. For its effect to be comparable or less than orbital errors, the accuracy of reference coordinates

23

should be better than that of the orbits. For example, if broadcast ephemeris of 20 m are used, the accuracy of the reference station should be better than 20 m.

2.2.4   Satellite Clock Bias

The satellite clock bias is the time offset of the satellite clock with respect to GPS time. For civilian users, the clock effect mostly comes from Selective Availability (SA, cf. 2.3).

Because the difference in the signal emission times from a satellite to two receivers is small, the drift of the satellite clock in the period is negligible. Therefore, the effect of the bias is almost completely removed in double difference observations as stated earlier. The higher order terms of a satellite clock offset are negligible in double difference observations. The residual effect of the satellite clock offsets can be neglected in the ambiguity resolution.

2.2.5   Receiver Clock Bias

The receiver clock bias is the offset of the receiver clock time with respect to GPS time. Although a GPS receiver is supposed to synchronize itself to GPS time at the start of observations, the synchronization is not perfect. Furthermore, the receiver clock will drift after synchronization. However, in double difference observations, just as with satellite clock bias, receiver clock bias can be almost completely removed. The effect of the receiver clock offset and drift on the ambiguity resolution can be ignored for most receivers.

2.2.6 Ionospheric Effect

The ionospheric effect is caused by interference of free electrons when GPS signals pass through the upper layer of the atmosphere. The effect on range may vary from more than 150 m (at midday, during period of maximum sunspot activity, with the satellite near the horizon of the observer) to less than 5 m, (at night, during the period of minimum sunspot activity, with the satellite at the zenith) [Wells et al. 1987]. For GPS carrier frequencies, the ionospheric effect is dispersive, meaning that the amount of ionospheric delay depends on frequency. The dispersive nature of the ionospheric effect can be used as an advantage. For example, for dual frequency GPS phase observations, a linear combination can be formed to eliminate most of the ionospheric effect. However, in this case, the integer ambiguity resolution is difficult to achieve from this ionosphere-free combination alone. The ambiguities are generally estimated as real value.

Ionospheric correction coefficients from the broadcast message can remove only 50% of the ionospheric delay at mid-latitudes [Wells et al. 1987]. Recently, attempts have been made to model the ionospheric effect with polynomials from the difference between the phase and code measurements. Cohen et al. [1992] used a first order spherical harmonics function. Qiu [1993] and Qiu et al. [1994] used third order polynomials. However, these modelling attempts are not accurate enough to be suitable in ambiguity resolution.

Over short baselines below 10 km, the effect of the ionosphere tends to cancel out between the two receivers. Over long baselines, the ionospheric effect increases as the baseline length increases. As a result, the ionospheric effect limits the length of a baseline on which integer ambiguities can be resolved.

2.2.7   Tropospheric Effect

The tropospheric effect is the propagation delay caused by the refraction of a GPS signal in the lower atmosphere and is independent of the GPS carrier frequency. To account for this delay, a tropospheric model has been used. Hopfield's tropospheric model is a popular one [Hopfield, 1971]. However, studies have shown that other models may be better at low elevations [Janes et al. 1990].

The un-modeled tropospheric effect is strongly correlated over a short distance between the reference and the remote if the height difference of the two stations is small. However, when the separation distance or height difference is large, local atmospheric conditions will be more different and the correlation becomes weaker. Consequently, adequate modeling remains difficult, especially for the wet delay component.

Surface meteorological data is not accurate to represent atmospheric conditions along the signal path. To get more accurate data, The water vapor content of the atmosphere along the propagation path can be measured with water vapor radiometers. However, the instruments are very elaborate and expensive and can only be used in major projects.

A parameter estimation approach can be used to model the tropospheric effect. That is to say, a nuisance parameter for each station per observation window is designated for the tropospheric delay [Chen, 1991]. However, in kinematic positioning, estimation of the scale factor will over-parameterize the state model and make the ambiguity resolution more difficult due to the limited number of observations.

As with orbital and ionospheric effects, un-modeled tropospheric effects also limit the length of the baseline on which integer ambiguities can be resolved.

2.2.8   Multipath

Multipath means that reflected signals also reach the antenna in addition to the direct signal. Multipath propagation is almost inevitable for most GPS applications due to all possible reflectors, such as streets, buildings, water, and observing platforms.

Multipath error affects both pseudorange and carrier phase measurements. The amount of multipath for a code observation is much larger than that for a carrier phase. Traditionally, the pseudorange multipath can reach up to one chip length of the PRN code (293 m for the C/A code, and 29.3 m for the P code), while carrier phase multipath is less than 25% for the carrier wavelength [Georgiadou and Kleusberg, 1988]. However, in most cases, observed multipath from C/A code receivers is less than 20m. NovAtel's patented Narrow Correlator™ technique usually reduces the multipath effect on the C/A to submetre levels [Van Dierendonck et al. 1992]. New development on code multipath reduction techniques has been conducted by van Nee [van Nee and Siereveld, 1993].

Multipath is proportional to the ratio of the direct signal power to the reflected signal power. Typically, in static observations, multipath is non-Gaussian in nature and shows sinusoidal oscillations with periods of a few minutes. In kinematic mode, multipath appears more random, due to vehicle movement and environmental change. In a strong multipath environment, the required observation time in the field may have to increase significantly to correctly resolve the satellite carrier phase ambiguities.

For surveying, sites of low multipath can often be chosen. However, for kinematic positioning, the environment may be difficult to control. Choker-rings, absorbing material near the antenna or ground plates may reduce the effect of multipath. Special receiver antenna design and firmware may also help. In principle, another method that

can help reduce the effect of multipath is to model this effect when the relative locations with respect to the reflecting objects and the reflection index are known [Georgiadou and Kleusberg, 1988]. However, it is almost impossible for most users due to complicated nature of the reflecting objects.

For most applications, the multipath effect is independent of the length of the baseline. Multipath is a major error source for differential positioning of short baselines. For static positioning, the effect of the multipath tends to be reduced. When multipath is strong, it will take a substantially longer time to resolve the ambiguities since the multipath effect varies slowly over time. This is especially true when strong multipath is combined with limited number of satellites and unfavorable satellite geometry.

2.2.9   Cycle Slips

Cycle slips occur if the receiver loses phase lock on the satellite signal. This may be caused by external or internal factors.

The examples of external factors include:
- obstructions, e.g., buildings, trees,
- high signal noise, in particular caused by multipath and ionospheric scintillation,
- low satellite elevation, causing low signal strength,
- antenna inclination in kinematic application (airplane, ship).

The examples of internal factors include:
- weak signals, partly caused by signal interference,
- signal processing method used.

In many applications, especially for static positioning, cycle slips pose no problem, since the prediction errors of the observations are normally well below one cycle. The triple difference, i.e., the difference of the double difference observations between two consecutive epochs, is used to detect and eliminate cycle slips. The measurement of the phase rate can also be used to compute the cycle slips if the errors of the integrated values of Doppler are less than half a cycle [Cannon, 1991].

Cycle slips may pose a difficulty for precise kinematic positioning. Neither the prediction method nor triple differencing can be used where the error in the predicted position is larger than half a cycle. The Doppler measurement may detect large cycle slips, but do little to eliminate them since the speed between two consecutive measurements is often not uniform.

If four or more satellites are free of cycle slips, the cycle slips for the remaining satellites can be computed easily. However, it will be difficult to fix cycle slips if the number of satellites without cycle slips is less than four. If cycle slips cannot be fixed, the ambiguities should be resolved again.

The integration of an additional sensor, for example, the integration of an inertial sensor package, can help to bridge gaps caused by cycle slips. The integration can be used to predict GPS measurements. However, the prediction is only good for a very short time (i.e., a few seconds) and most inertial systems will drift more than one cycle after several seconds [Cannon, 1991]. In addition, an inertial system is too expensive for most users.

2.3   Selective Availability and Anti-Spoofing

Since GPS is a military navigation system and is the primary responsibility of the US Department of Defense (DoD), GPS has to meet the national security interests of the

United States. Accordingly, access to the total system accuracy by the national and international civil community is limited.

The service available to the civil community is called Standard Positioning Service (SPS); the service available to authorized users, mainly military, is called the Precise Positioning Service (PPS). Under the current policy of Federal Radio Navigation Plan, the accuracy available to SPS provides 100 m of 2D-RMS; PPS provides 10 to 20 metres in three dimensions [FRNP, 1990].

Anti-Spoofing (AS) entails the encryption of P-code, or in other words, the use of a protected code named Y-code. Only authorized users have the means to access the P-code while AS is activated. Selective Availability (SA) degrades both broadcast ephemeris and satellite clocks, and thus, introduces errors in measurements. As a result, SA is especially bad for single point positioning. However, for double differencing positioning, the effects of SA are almost completely removed since double differencing significantly reduces the effect of SA.

CHAPTER 3   LEAST-SQUARES AND KALMAN FILTERING

This chapter takes a look of the filtering technique from the least-squares estimation. Starting with classical least-squares estimation, the least-squares filtering is derived from the least-square estimation. It is also shown here that least-squares filtering is equivalent to Kalman filtering. Finally, parameter removal is discussed concerning cases in which some estimated parameters become perfectly known. The parameter removal technique discussed is equivalent to the technique that treat the parameters as known values from the beginning.

3.1   The Problem of Least-squares Estimation

3.1.1   Definition of the Problem

Observations are made to derive certain parameters. However, observations often contain biases and errors. To reduce the effect of the errors and assess the accuracy of the solution, redundancy is required. That is, more than the minimum number of observations is required to determine the estimated parameters. These observations must be adjusted so that the solution will be consistent with these adjusted observations. To adjust observations and to obtain the desired parameters, the method of least-squares estimation is often used. In least-squares estimation, parameters and corrected observations are derived by minimizing the weighted sum of the squared residuals. This process is subject to certain constraints among the observations and estimated parameters.

Suppose that the mathematical relation between the observation vector l and the unknown parameter vector x is:

$$l = f(x) + \varepsilon,$$ (3.1)

where,

   ε is the unknown observation noise.

Then, the least-squares estimation of x is to seek a vector $\hat{x}$ ($\hat{x} \in X$, where, X is a real space) such that

$$(l - f(\hat{x}))^T P (l - f(\hat{x})) = \min,$$ (3.2)

where,

   P is the *a-priori* weighting matrix of the observations.

Note that the symbol, $^\wedge$, over a vector is referred to as the corresponding least-squares estimate throughout this thesis.

Here, the *a-priori* information is also regarded as a *quasi*-observation. The constraints in a space could be regarded as a sub-space where the solutions should belong to. The solution space could be discrete (such as cycle ambiguities in GPS carrier phase measurements), or continuous (as for coordinates or velocities), or mixed.

Generally, the solution may not necessarily be unique. However, in many applications, such as, surveying and navigation, a problem is often designed in such a way that only a unique solution can be derived from the least-squares estimation process. The uniqueness is determined by the geometry of observations (including *a-priori* information), the variance-covariance matrix of the observations, and the constraints.

<u>Constrained</u> <u>Adjustment</u> <u>with</u> <u>Unknown</u> <u>Parameters</u>

Assuming the *n*-dimensional observation vector as *l* with variance-covariance matrix C_l, where the observations here also include the *a-priori* information, the constraints are then described by a system of *m* equations such that:

$$f(x, l + r) = 0 ,$$
(3.3a)

where,

x   is the unknown parameter vector to be estimated, and

r   is the correction on observation *l* .

The above equation can be linearized as follows:

$$A \, \delta + B \, r + w = 0 ,$$
(3.3b)

where,

$A \quad = \dfrac{\partial f}{\partial x}\Big|_{x_o,l}$ is the design matrix with respect to the unknown parameter vector,

$B \quad = \dfrac{\partial f}{\partial l}\Big|_{x_o,l}$ is the design matrix with respect to the observation vector,

$w \quad = f(x_o,l)$ is the misclosure vector,

$x_o \quad$ is the vector of approximate value of *x*,

$\delta \quad$ is the correction on the approximate value vector of *x*, i.e., $x = x_o + \delta$, and

r   is the residual vector.

Using the minimum criterion (3.2), the solution of the least-squares estimation for δ and *r* becomes:

$$\hat{\delta} \quad = -[A^T M^{-1} A]^{-1} A^T M^{-1} w ,$$
(3.4a)

$$\hat{r} \quad = -C_l B^T M^{-1} (A \hat{\delta} + w) , \text{ and}$$
(3.4b)

$$M \quad = B \, C_l \, B^T .$$
(3.4c)

The variance-covariance matrices corresponding to the above estimation are:

$$C_{\hat{\delta}} = [A^T M^{-1} A]^{-1} , \text{ and} \qquad (3.5a)$$

$$C_{\hat{r}} = C_l B^T M^{-1} B C_l - C_l B^T M^{-1} A C_{\hat{\delta}} A^T M^{-1} B C_l \qquad (3.5b)$$

[Krakiwsky, 1990]. Note that $\hat{\delta}$ is the variation from approximate state vector which has no weighting in the adjustment. The *a-priori* information is included in observation vector. Therefore,

$$C_{\hat{x}} = C_{\hat{\delta}} . \qquad (3.6)$$

The *a posteriori* variance factor can be computed as:

$$\hat{\sigma}_0^2 = \hat{r}^T C_l^{-1} \hat{r} / (m\text{-}u) . \qquad (3.7)$$

The formula for the computation of $\hat{r}^T C_l^{-1} \hat{r}$ can be derived as:

$$\hat{r}^T C_l^{-1} \hat{r} = w^T M^{-1} w - \hat{\delta}^T C_{\hat{x}}^{-1} \hat{\delta} , \qquad (3.8)$$

where,

     m is the number of constraints, and

     u is the number of unknown parameters.

3.1.3   Parametric Adjustment

When $B = -I$, where, I is the unit matrix, the estimation in Section 3.1.2 becomes a parametric adjustment. The formulas then become:

$$\hat{\delta} = -[A^T C_l^{-1} A]^{-1} A^T C_l^{-1} w , \qquad (3.9a)$$

$$\hat{r} = (A \hat{\delta} + w) , \qquad (3.9b)$$

$$C_{\hat{\delta}} = [A^T C_l^{-1} A]^{-1} , \qquad (3.9c)$$

$$C_{\hat{r}} = C_l - A C_{\hat{\delta}} A^T , \qquad (3.9d)$$

$$= (w^T C_l^{-1} w - \hat{\delta}^T C_{\hat{x}}^{-1} \hat{\delta})/(m\text{-}u) \qquad (3.9e)$$

3.1.4 Conditional Adjustment

When A is a null matrix, the estimation in Section 3.1.2 becomes a conditional adjustment, that is,

$$\hat{r} = - C_l B^T M^{-1} w ,$$ (3.10a)

$$M = B C_l B^T ,$$ (3.10b)

$$C_{\hat{r}} = C_l B^T M^{-1} B C_l ,$$ (3.10c)

$$C_{\hat{l}} = C_l - C_{\hat{r}} , \text{ and}$$ (3.10d)

$$\hat{\sigma}_0^2 = \hat{r}^T C_l^{-1} \hat{r} / m .$$ (3.10e)

The formula for the computation of $\hat{r}^T C_l^{-1} \hat{r}$ can be derived as:

$$\hat{r}^T C_l^{-1} \hat{r} = w^T M^{-1} w .$$ (3.10f)

The above formulas will be used to derive a filter in the following sections.

3.2 From Least-squares to Kalman Filtering

The problem of least-squares filtering is illustrated in Fig. 3.1 for kinematic satellite positioning. The position, velocity, and other parameters (denoted as x_{k+1}) are to be estimated at epoch $k+1$ from the observation vector l_{k+1} and its variance-covariance matrix $C_{l_{k+1}}$. Supposing that r_{k+1} is the correction to l_{k+1}, the relationship between the l_{k+1} and x_{k+1} is defined by the prediction model:

$$l_{k+1} + r_{k+1} = f(x_{k+1}, t_{k+1}) .$$ (3.11)

The relationship between x_{k+1} and x_k is defined by the prediction model, or transition model:

$$x_{k+1} = g(x_k, t_{k+1}, t_k) + \varepsilon_{k+1,k} ,$$ (3.12)

where,

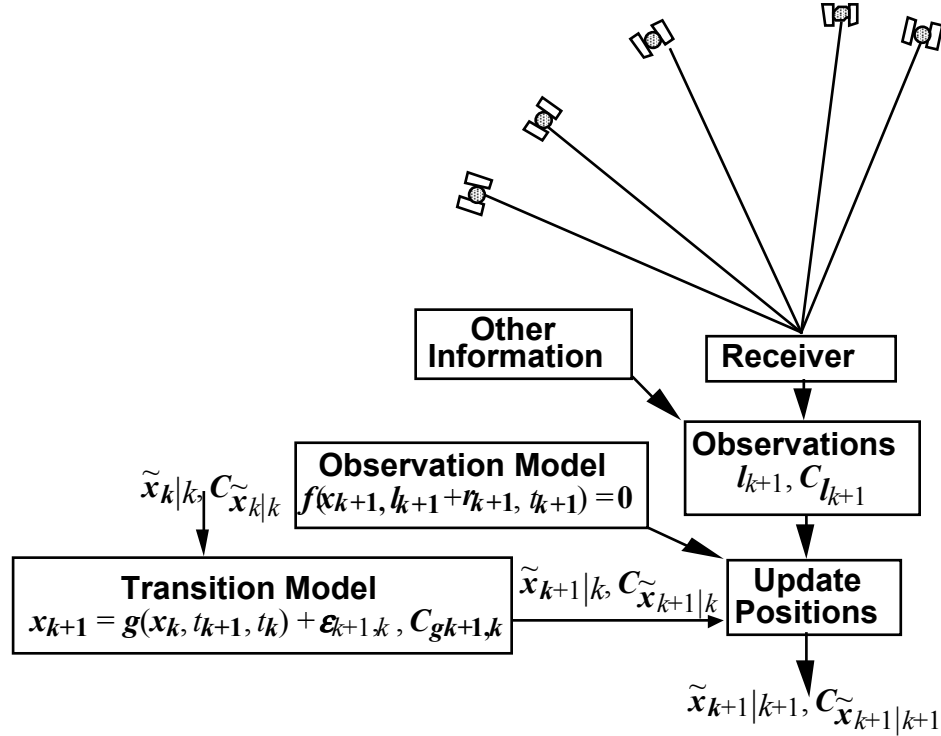$\varepsilon_{k+1,k}$ is the uncertainty of the model, and its variance-covariance matrix is $C\varepsilon_{k+1,k}$.

Figure 3.1   Problem of Kinematic Positioning in GPS

The question now is how to estimate the system parameter vector, x_{k+1}, with all information up to the current epoch, $k+1$. The problem was first solved by Kalman [1960] and is, therefore, called Kalman Filtering. In this section, the least-squares estimation technique described in the previous section is applied to derive the equivalent formulas. The advantage of this approach is that many available formulas in the least-squares adjustments can be used.

The problem is solved in two steps. In the first step, the current state vector is computed using equation (3.12) and its variance-covariance matrix is also computed. In the second step, both the predicted *a-priori* and direct observations are used to derive the state vector. This step is called updating.

36

At the first epoch, the updated value is computed from the *a-priori* and direct observations. All updated values of the state vectors at subsequent epochs are computed from predicted and direct observations.

3.2.1 Prediction of the State Vector

Assuming that $\hat{x}_{k|k}$ is the updated value for state vector x_k, $\hat{x}_{k|k}$ is computed from all the information up to epoch k. The predicted value for the state vector x_{k+1}, namely, $\hat{x}_{k+1|k}$, is computed from $\hat{x}_{k|k}$ using equation (3.12):

$$\hat{x}_{k+1|k} = g(\hat{x}_{k|k}, t_{k+1}, k) + \varepsilon_{k+1,k} , \tag{3.13}$$

where, the expected value of $\varepsilon_{k+1,k}$ is regarded as zero.

In order to obtain the variance-covariance matrix, the kinematic prediction model is linearized as:

$$\hat{x}_{k+1|k} = \mathbf{ø}_{k+1,k}(\hat{x}_{k|k} - x_{k_0}) + g_{k+1,k}(x_{k_0}, t_{k+1}, t_k) + \varepsilon_{k+1,k} , \tag{3.14}$$

where,

$$\mathbf{ø}_{k+1,k} = \frac{\partial\, g_{k+1,k}}{\partial\, x_k}\bigg|x_{k_0} , \text{ and}$$

x_{k_0} is the point of expansion in the Taylor's series and normally takes the value of $\hat{x}_{k|k}$.

By using the law of error propagation, the variance-covariance matrix of the predicted state vector, $C_{\hat{x}_{k+1|k}}$, can be computed as:

$$C_{\hat{x}_{k+1|k}} = \mathbf{ø}_{k+1,k}\, C_{\hat{x}_{k|k}}\, \mathbf{ø}^T_{k+1,k} + C\varepsilon_{k+1,k} . \tag{3.15}$$

3.2.2 Updating Using a Parametric Adjustment

The observation equation can be linearized as:

$$r_{k+1} = A_{k+1}\,\delta_{k+1} + w_{k+1}\,,\tag{3.16}$$

where,

$$A_{k+1} \;=\; \frac{\partial f_{k+1}}{\partial x_{k+1}}\,\big|\hat{x}\,_{k+1|k}\,,$$

$$w_{k+1} \;=\; f_{k+1}(\hat{x}\,_{k+1|k},\, t_{k+1}) - l_{k+1}\,,\ \text{and}$$

$$\delta_{k+1}\quad \text{is the correction to the predicted value, }\hat{x}\,_{k+1|k}\,.$$

The estimated parameters are included in $\delta_{k+1|k+1}$.

In the next two sections, the updating of the state vector will be discussed.

The observation equation corresponding to the *a-priori* state vector is:

$$r_{x_{k+1|k}} = \delta_{k+1}\,.\tag{3.17}$$

By combining equations (3.16) and (3.17) into one, the following equation system can be obtained:

$$r = A\,\delta_{k+1} + w\,,\tag{3.18}$$

with a variance-covariance matrix of

$$C_l = \begin{pmatrix} C\hat{x}_{k+1|k} & 0 \\ 0 & C_{l_{k+1}} \end{pmatrix},\tag{3.19}$$

where,

$$r \;=\; \begin{pmatrix} r_{x_{k+1/k}} \\ r_{k+1} \end{pmatrix}$$

$$A = \begin{pmatrix} I \\ A_{k+1} \end{pmatrix}\ \text{and}$$

$$w \;=\; \begin{pmatrix} 0 \\ w_{k+1} \end{pmatrix}.\tag{3.20}$$

By applying the parametric adjustment formulas from Section 3.1.3, the following estimation can be obtained:

$$\hat{\delta}_{k+1|k+1} = -[A^T C_l^{-1} A]^{-1} A^T C_l^{-1} w \ ,$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \hat{\delta}_{k+1|k+1}$$

$$= \hat{x}_{k+1|k} - [A^T C_l^{-1} A]^{-1} A^T C_l^{-1} w, \text{ and}$$

$$C_{\hat{x}_{k+1|k+1}} = [A^T C_l^{-1} A]^{-1} \ , \tag{3.21}$$

Note that here $|k+1$ means that the values are obtained using the information up to epoch $k+1$.

By placing equations (3.19) and (3.20) into equation (3.21), the following can be derived:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} - [C_{\hat{x}_{k+1|k}}^{-1} + A_{k+1}^T C_{l_{k+1}}^{-1} A_{k+1}]^{-1} A_{k+1}^T C_{l_{k+1}}^{-1} w_{k+1}$$

$$= \hat{x}_{k+1|k} - K w_{k+1} \ , \tag{3.22a}$$

$$C_{\hat{x}_{k+1|k+1}} = [C_{\hat{x}_{k+1|k}}^{-1} + A_{k+1}^T C_{l_{k+1}}^{-1} A_{k+1}]^{-1} \ , \tag{3.22b}$$

where,

$$K = [C_{\hat{x}_{k+1|k}}^{-1} + A_{k+1}^T C_{l_{k+1}}^{-1} A_{k+1}]^{-1} A_{k+1}^T C_{l_{k+1}}^{-1} \ . \tag{3.22c}$$

Comparing with Bayes expression of Kalman filtering in Krakiwsky [1990], the two expressions are equivalent. Here, w_{k+1} is equivalent to $(w_{k+1} + A_{k+1} \hat{\delta}_{k+1|k})$ in Bayes expression. This can be done by the linear expansion of Taylor's series with w_{k+1} at an approximate value of the current state vector x_{k+1}. However, this is not necessary here, since all linearization can be made at the predicted value of the state vector.

3.2.3 Updating Using Conditional Adjustment

The least-squares filtering expression can also be derived from the conditional adjustment. The observations are the same as in the last sub-section. However, the constraint equations can be reformatted as:

$$B\,r + w = 0,$$ (3.23)

where,

$$r \quad = \begin{bmatrix} \delta_{k+1} \\ r_{k+1} \end{bmatrix},$$

$$w \quad = w_{k+1},$$

$$B \quad = (A_{k+1}, \; -I), \text{ and}$$ (3.24)

$$C_l \quad \text{is the same as in equation (3.19).}$$

From the equations of the conditional adjustment, the least-squares estimation is:

$$\hat{r} = -C_l\,B^T\,M^{-1}\,w,$$ (3.25a)

$$M = B\,C_l\,B^T,$$

$$C_{\hat{r}} = C_l\,B^T\,M^{-1}\,B\,C_l, \text{ and}$$ (3.25b)

$$C_{\hat{l}} = C_l - C_{\hat{r}}.$$ (3.25c)

By placing equation (3.24) into equations (3.25a) and (3.25c), the following can be obtained:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \hat{\delta}_{k+1|k+1}$$

$$= \hat{x}_{k+1|k} - C_{\hat{x}_{k+1|k}}\,A_{k+1}^T\,[A_{k+1}\,C_{\hat{x}_{k+1|k}}\,A_{k+1}^T + C_{l_{k+1}}]^{-1}\,w_{k+1}$$

$$= \hat{x}_{k+1|k} - K\,w_{k+1}, \text{ and}$$ (3.26)

$$C_{\hat{x}_{k+1|k+1}} = C_{\hat{x}_{k+1|k}} - K\,A_{k+1}\,C_{\hat{x}_{k+1|k}},$$ (3.27)

where,

$$K = C_{\hat{x}_{k+1|k}}\,A_{k+1}^T\,[A_{k+1}\,C_{\hat{x}_{k+1|k}}^{-1}\,A_{k+1}^T + C_{l_{k+1}}]^{-1}.$$ (3.28)

By comparing the above with the results from Krakiwsky [1990], the above expressions are equivalent to Kalman Filtering except for the term, w_{k+1}. In the formulas given by Krakiwsky, the linearization is at a value other than the predicted one, while here, it is at the predicted value.

3.3 Parameter Removal and Transformation of Normal Equations

During filtering, some estimated constant parameters may become perfectly known without biases and noise such as the ambiguity parameters in GPS. To determine these parameters as if they were perfectly known from the beginning, we want to find a rigorous filter solution. The following discusses a method that transforms estimated constants in filtering to deterministic constants. For convenience, this process is called parameter removal.

To introduce the method for parameter removal, the parametric adjustment method is used. For simplicity in this discussion, all subscripts related to epochs are dropped off. All estimated values are referred to having used all information up to the current epoch.

The state vector is partitioned into two parts: one part corresponds to the constant unknowns and the other part corresponds to remaining unknown. Similarly, the design matrix, normal matrix, and normal equation are all partitioned accordingly.

Theorem. **If** the normal equations of a kinematic system at any epoch are:

$$\boldsymbol{P}_{xx}\, \delta\hat{\boldsymbol{x}} + \boldsymbol{P}_{xy}\ \delta\hat{\boldsymbol{y}} = \boldsymbol{u}_x \,, \tag{3.29a}$$

$$\boldsymbol{P}_{yx}\, \delta\hat{\boldsymbol{x}} + \boldsymbol{P}_{yy}\ \delta\hat{\boldsymbol{y}} = \boldsymbol{u}_y \,, \tag{3.29b}$$

where,

$\hat{y}$ is the constant part of the unknowns in the state vector,

$\hat{x}$ is the remaining part of the state vector,

$\delta\hat{x}$ and $\delta\hat{y}$ are the variation of $\hat{x}$ and $\hat{y}$ from their approximate values, x_o, y_o.

Then, if y becomes a constant known, denoted y_C, the normal equation 3.29 becomes:

$$P_{xx} \, \delta\hat{x}\,|_y = u_{x|y}\,, \qquad\qquad\qquad\qquad (3.30)$$

and the weighted sum of the residual squares becomes:

$$\Omega_{|y} = \Omega + (y_C - \hat{y})^T \tilde{P}_{yy} (y_C - \hat{y})\,, \qquad\qquad\qquad (3.31)$$

where,

$$u_{x|y} \quad = u_x - P_{xy} (y_C - y_o)\,, \qquad\qquad\qquad\qquad (3.32)$$

$$\tilde{P}_{yy} \quad = P_{yy} - P_{yx} P_{xx}^{-1} P_{xy}\,,$$

$\hat{y}$ is the float estimation from equation (3.29),

$\hat{x}\,|_y$ is the estimation of x corresponding to the known value of y,

Ω is the weighted sum of squared residuals while treating y as the estimated parameter vectors, and

$\Omega_{|y}$ is the weighted sum of residuals while treating y as a vector of known constants.

The proof of the same theorem with parametric adjustment is included in the Appendix. Another method to prove the theorem is to use the conditional adjustment method, where, y is constrained to the integer values.

CHAPTER 4   FAST AMBIGUITY SEARCH FILTER: THE CONCEPT AND ITS IMPLEMENTATION

In this section, basic concepts for ambiguity resolution, Fast Ambiguity Search Filter (FASF), are described, followed by a special recursive sequential search algorithm. Then, the least-squares method to implement FASF is discussed. The observation equations are given in Section 4.2 with the prediction model following. Finally, the least-squares implementation of FASF is discussed.

4.1   Development of the Concept of FASF

FASF uses a least-squares filter, a special ambiguity searching process where ambiguities are searched at every epoch until they are fixed, and an index is used to exit the search process without completing full search.

In the current implementation of FASF, the number of potential solutions is used as the index to exit the ambiguity search. An attempt is made to fix the ambiguities if the total number of potential ambiguity sets from the search is less than a certain threshold. If the number is one, the ambiguity set is regarded as the correct one. Otherwise, other tests are carried out, such as the ratio test of the sum of squared residuals: the second minimum divided by the minimum. If the ratio is larger than the specified threshold, the one with the minimum Ω (the weighted sum of the squared residuals) is regarded as the correct

ambiguity set. However, if the radio test fails, the ambiguities are estimated as real values.

Since the full search of potential ambiguities is avoided with FASF, only a relatively small amount of computation time is needed for ambiguity searching. Consequently, the computational efficiency is significantly improved.

The threshold on the number of potential solutions affects computational efficiency. The smaller the threshold, the less the computation is required. However, a small threshold requires a longer observation period before ambiguity resolution is made. From the tests described in Chapter 5, the threshold of ten works reasonably well. However, this value is by no mean an optimal one and should be investigated further.

A least-squares filter, the equivalent of a Kalman filter obtained by applying least-squares theory, is used in FASF and is illustrated in Figure 4.1. In this method, the process noise corresponding to the ambiguity parameters from one epoch to the next is zero. The ambiguities are searched at every epoch, starting from the first epoch, until they can be fixed. Once the ambiguities can be fixed correctly, they are regarded as known integers. The principle of FASF can also be used for static positioning by assigning the noise of predicted positions as zero; in this case, the velocity components are also zero and are not included in the state vector. As will be discussed in section 4.4, the biases in the GPS observations are strongly correlated over the time. Adding processing noise in the estimated ambiguities may partially compensate the impact of the correlation. However, the exact impact should be further investigated.
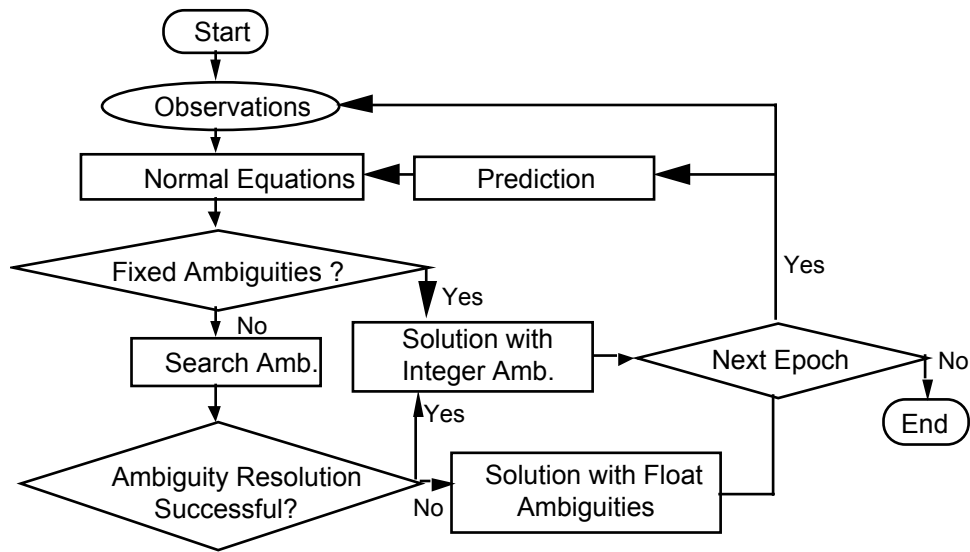
Figure 4.1   Flowchart of Fast Ambiguity Search Filter

In all previous scheme, the value of an ambiguity parameter could be treated as known only if it was correct, or all potential ambiguity sets have to be fully searched. In the paper by Blewitt [1989], ambiguity parameters converge sequentially in a static application. His main assumption is that by arranging the sequence of ambiguity parameters, at least one of the estimated ambiguity parameters will converge to correct integer in terms of a probability level he designed. Once the ambiguities are known values, they can be removed from the estimation. He also used precise code and wide lane ambiguities. The assumption of the convergence may be the case for some static positionings, especially when precise code measurements are also used and the observation period is long enough. However, in many applications, especially in kinematic applications, the convergence can not be guaranteed. The estimation often converges to incorrect values due to various systematic biases.

4.1.1 Search Algorithm

One of the important characteristics of FASF is the ambiguity search procedure. Here, ambiguities are arranged in a series The search range of each ambiguity is determined recursively and sequentially by updating constraints. To compute the search range of an ambiguity parameter in the series, the presumed integer ambiguities on the left of this series (assuming the ambiguities are arranged from the left to the right) are regarded as known. Here, a recursive procedure can be applied and the constraints are updated in computing the uncertainty range for ambiguities from N_1 to N_n. This concept is called the Recursive computation of the Search Range for the ambiguities *(RCSR),* or updating of the constraints.

Assuming the ambiguity series as N_1, N_2, N_3, ..., N_n, the search ranges for the ambiguities are computed from N_1 to N_n, where, n is the number of ambiguity parameters. The search range of possible integers for ambiguity N_i is computed for each specific integer set of ambiguities N_1, N_2, ..., N_{i-1}. In computing the search range, the values of N_1, N_2, ..., N_{i-1} are treated as correct, while N_i, N_{i+1}, N_{i+2}, ..., N_n are treated as estimated parameters.

Note that the search range of N_i should be equal to the maximum uncertainty range of the real estimation of N_i. It can be expressed as:

$$x_{N_i|N_1,\,N_2,\,...,\,N_{i-1}}{}^{min} \leq N_i \leq x_{N_i|N_1,\,N_2,\,...,\,N_{i-1}}{}^{max}\,, \qquad\qquad (4.1)$$

where,

$x_{N_i|N_1,\,N_2,\,...,\,N_{i-1}}{}^{min}$ is the minimum possible value of N_i if the integer values of N_1, N_2, ..., N_{i-1} are correct, and

$x_{N_i|N_1,\,N_2,\,...,\,N_{i-1}}{}^{max}$ is the maximum possible value of N_i if the integer values of N_1, N_2, ..., N_{i-1} are correct.

Of course, to compute the search range of N_i, each potential integer set of ambiguities $N_1, N_2, ..., N_{i-1}$ should also be included in the corresponding ranges derived from the same principle described here. Since N_1 is the first in the series, its maximum uncertainty is derived from the float solution without fixing any ambiguities to possible integer values. All integer sets falling into these ranges should be included.

Not all integer ambiguity sets will be included in these search ranges. However, the correct solution will be included if the search ranges of possible ambiguities are computed properly. Incorrect sets of ambiguities may also be included in the search ranges due to insufficient geometry, the effect of noise, and poor initial *a-priori* information, though the incorrect solutions normally have a higher chance of being rejected in the *RCSR*. However, as observations and geometry accumulate, only the correct solution will continue to satisfy these search ranges. In the case of good geometry, low noise, and good *a-priori* information, it is possible to determine the ambiguities in a few epochs (even one epoch). However, for single frequency observations, it will normally be rare that ambiguities will be resolved in one epoch except in some very favorable conditions, such as a large number of satellites, very small biases and low noise in the phase and pseudorange observations, and special satellite configurations.

4.1.2   Search Algorithm Based on RCSR

To explain the basic concept of FASF in more detail, a search process based on the RCSR can be designed. The search is a procedure of multi-level loops as described below:

Loop 1:   Compute the search range for ambiguity N_1.

1. The search range for N_1 is computed without constraining any ambiguities to integers. The correct value of N_1 is one of the integers in this range if the search range is computed properly.

2. For each possible integer of in the search range of N_1, search for possible integer values of other ambiguities. That is, go to the deeper loops for other ambiguities.

3. Exit the search loop when all the possible integers of N_1 have been searched.

<u>Loop 2</u>:   <u>Compute the search range for N_2</u>

1. Corresponding to each constraint of N_1 to a possible integer, the range of all the possible integers for N_2 is computed as if the constrained value of N_1 was a correct integer value of N_1. When N_1 is a correct value, the correct N_2 will be included in the range and deeper loops can be carried forward.

2. If no integer is in the range of N_2, go back to loop N_1 and search the next possible integer of N_1.

3. Go to upper loop level, loop N_1, if all the possible integers have been searched.

<u>Loop 3</u>: <u>Compute search range for N_3</u>

1. Similarly, for each constraint of N_1 and N_2 to their integers, compute the search range for N_3. The specific integer pair of N_1 and N_2 is treated as correct values in the range computation. When the integer pair of N_1 and N_2 is correct, the correct value of N_3 will be included in the search range.

2. For the integer pair of N_1 and N_2, if no integer is available in the computed range of N_3, continue to search for the next available integer pair of N_1 and N_2 (the next step in loops N_1 and N_2).

3. Go to the upper loop, loop N_2, if all the possible integers of N_3 have been searched.

<u>Loop i</u>:   <u>Compute the search range of N_i</u>

1. Again similarly, compute the search range of all possible integers of N_i corresponding to each set of constrained integer values for $N_1,...,N_{i-1}$. The specific integers of ambiguities $N_1,...,N_{i-1}$ are correct values in the range computation.

2. If the integer set for $N_1,...,N_{i-1}$ is correct and the search range is appropriate, the correct integer of N_i will be included in the search range and the deeper loop can be carried forwards. If no integer is available in the computed range of N_i corresponding to a specific set of $N_1,...,N_{i-1}$, search the next available integer set of $N_1,...,N_{i-1}$ (i.e., the next step in loops $N_1,...,N_{i-1}$).

3. Go to upper level of loop (i.e. N_{i-1}) if all the possible integers of N_i have been searched.

Loop i+1 to Loop n-1: Similar to the above.

Loop n: Compute the search range of for N_n

1. Compute the search range for N_n corresponding to each integer set of $N_1, ..., N_{n-1}$.

2. If no integer is in the search range of N_n, go back to loop N_{n-1} and search the next possible integer of N_{n-1}.

3. Go to the upper loop level, loop N_{n-1}, if all the possible integers of N_n have been searched.

Loop N_n is the deepest loop in the ambiguity search. Therefore, a full set of integer ambiguities is obtained for each possible integer of N_n. That is, a possible ambiguity set can be obtained by putting together the integer ambiguities of different loops since each loop is related to an ambiguity parameter. Therefore, the possible number of integers for ambiguity N_n in this loop is added to the total number of potential solutions.

Whenever the accumulated number is larger than a certain threshold, for example, one, four, ten, etc., depending on the circumstances, it becomes apparent that fixing

ambiguities may not be possible. At this time, the search process stops and the ambiguities are estimated as real numbers. The number of potential ambiguity sets is used as an index of possible inabilities to fix the ambiguities. The use of the index is justified because the greater the number of potential solutions available, the less likely the correct ambiguity resolution will be retained.

4.1.3 Size of Search Windows

The second cause of the reduced computation of FASF is the window size for each ambiguity. As described earlier, in a traditional approach, the window sizes of individual ambiguities are unrelated to one another, except in Hatch's algorithm where the values of the secondary ambiguity parameters are almost uniquely determined by the values of the primary ambiguities. However, by using the *RCSR* concept presented here, the search windows of ambiguity parameters are related to each other. As a result, sequential updating of constraints normally makes subsequent uncertainty ranges smaller and smaller.

The difference in the window sizes is illustrated in Figure 4.2. For convenience, in this figure, the window sizes of different ambiguity parameters in the full search method are assumed to be constant and the ambiguity parameters are arranged in the same order for the different methods.
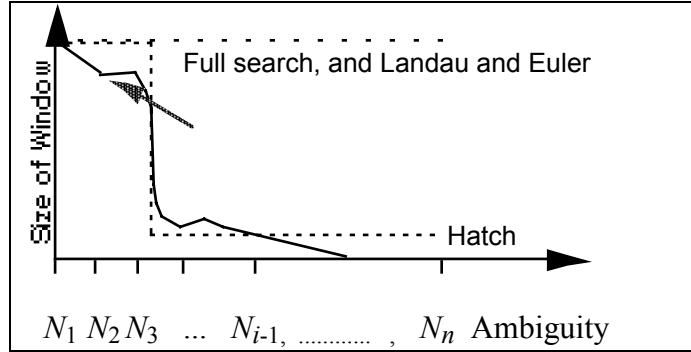
Figure 4.2   Illustration of Window Sizes for Three Different Approaches

Although   the search windows for the first epoch are the same for the full search in Landau and Euler's approach, the latter uses an optimized Cholesky decomposition algorithm to compute Ω.   As demonstrated in Landau and Euler [1992], there is a significant computational difference between the two approaches.

By using the *RCSR* concept, the decrease in window size is dependent on the correlation between the ambiguity parameters.   In applying *RCSR*, the steepness of window size relative to ambiguities will increase as the number of observations increases and geometry changes.

The window sizes for *RCSR* are smaller than for the full search or for Landau and Euler's approach.   The window sizes for the primary satellites in *RCSR* are smaller than for Hatch's method.   Actually, it is not necessary to keep the concept of the primary and the secondary ambiguities in *RCSR*.   The concept is described here only for comparison with Hatch's method.   However, for secondary ambiguities, the window sizes of *RCSR* can be smaller or larger than for Hatch's method.   In Hatch's method, there is a unique set of windows for the secondary ambiguities [Hatch, 1991]; in the *RCSR* concept, there may be none or more than one set of secondary ambiguities for a specific integer set of primary ambiguities.

4.1.4   Validation of the New Concept

With the intent of validating the RCSR concept, the following theorem is first presented:

RCSR Theorem:

If the set of potential ambiguity solutions is:

$$N^n = \{ \ \{N_1,.., \ N_i, \ ...N_n\},$$

while for each N_i *(i=1,.., n),*

$$(x_{N_i}|_{N_1, \ N_2, \ ..., \ N_{i-1}}^{min} \ \leq \ N_i \ \leq x_{N_i}|_{N_1, \ N_2, \ ..., \ N_{i-1}}^{max}) \cap N_i \in I^1\} \ , \qquad (4.2)$$

where,

$x_{N_i}|_{N_1, \ N_2, \ ..., \ N_{i-1}}^{min}$ is the minimum possible value of N_i if the integer values of N_1,

$\qquad N_2, \ ..., \ N_{i-1}$ are correct,

$x_{N_i}|_{N_1, \ N_2, \ ..., \ N_{i-1}}^{max}$ is the maximum possible value of N_i if the integer values of

$\qquad N_1, N_2, \ ..., \ N_{i-1}$ are correct

$I^1 \qquad$ is one dimensional integer space and $\cap$ is logical 'and'.

Note that $x_{N_i}|_{N_1, \ N_2, \ ..., \ N_{i-1}}^{min}$ and $x_{N_i}|_{N_1, \ N_2, \ ..., \ N_{i-1}}^{max}$ should be formulated in a way that when the values of $N_1,.., \ N_{i-1}$ are correct, the correct value of N_i will be included between them;

Then, the correct ambiguity set is included in N^n.

The theorem shows that when the uncertainty ranges are properly computed, the correct values of ambiguities are always included in the potential ambiguity sets by using any recursive method based on the *RCSR*.

The following is the prove of the theorem:

- When $n = 1$, N^n becomes $N^1 = \{ N_1, (x_{N_1}^{min} \leq N_1 \leq x_{N_1}^{max}) \cap N_1 \in I^1 \}$. From the definition, the correct ambiguity solution should be one of the sets from N^1. Therefore, the theorem is true when $n = 1$.

- If we can further prove true for $n = k+1$ under the condition of being true for $n = k$, it becomes true.

   Assuming that the theorem is true for $n = k$, the potential ambiguity set defined in the theorem is denoted as N'^k. According to the assumption, the correct ambiguity set should be inside N'^k. Denoting the correct values of the first k ambiguity parameters as $N'_1, N'_2, ..., N'_k$, from the definition of the maximum uncertainty in the theorem, the correct value of ambiguity N_{k+1} will be included between $x_{N_{k+1}|N'_1, N'_2, ..., N'_k}^{min}$ and $x_{N_{k+1}|N'_1, N'_2, ..., N'_k}^{max}$ corresponding to ambiguity values, $N'_1, N'_2, ..., N'_k$. This correct value is designated as N'_{k+1}.

   Therefore, $N'_1, N'_2, ..., N'_k, N'_{k+1}$ are the correct values of $N_1, N_2.., N_i, ...N_{k+1}$, i.e., the correct values of $N_1, N_2,.., N_i, ..., N_{k+1}$ are included in the theorem for $n = k+1$ if the correct values of $N_1, N_2, ..., N_k$ are included for $n = k$.

- According to the theory of mathematical induction, it can be concluded that the theorem is true.

Although RCSR is rigorous, the difficulty in this method is the determination of the uncertainty range due to the complex nature of noise. On the one hand, the range should be large enough to include the correct solution when other assumed integer ambiguity parameters are held fixed. On the other hand, the range should not be exaggerated. If the range is too large, more incorrect solutions will be included, and the discrimination between the correct solution and incorrect ones will be more difficult. If the uncertainty ranges are not properly computed, the correct solution may not be included in N^n. If that happens, the ambiguity resolution will fail. Many methods are possible to compute the

search ranges. As a first effort, a least-squares method to compute the search range is described in section 4.3.

4.2   Filter Scheme

In this section, the least-squares filter is applied to implement the basic ideas previously described. Following the general double difference observation equations, the prediction of the state vector is described and the adjustment algorithm is applied. The adjustment algorithm applied is equivalent to the standard Kalman filter [Chapter 3].

4.2.1   Observation Equations

The vector form of double difference GPS observations after the various corrections have been applied can be written as:

$$\Delta \nabla l_{ph} = \Delta \nabla \rho + \lambda \, \Delta \nabla N + \Delta \nabla \varepsilon_{ph} \ ,$$

$$\Delta \nabla l_{psr} = \Delta \nabla \rho + \Delta \nabla \varepsilon_{psr} \ , \tag{4.3}$$

where,

$\Delta \nabla l_{ph}$ is the vector of double difference phase observations in metres,

$\Delta \nabla l_{psr}$ is the vector of the double difference pseudorange observations in metres,

$\Delta \nabla \rho$ is the vector of double differenced geometric propagation distances from

the   satellites to the receivers, in metres,

λ is the whole wavelength of the carrier, L1, L2, or the wavelength of a

linear

        combination of dual frequency observations. However, for squaring

    type receivers [Wells et al. 1987], it is the half of the carrier wavelength, in

    metres, and

$\Delta \nabla \varepsilon_{ph}$ and $\Delta \nabla \varepsilon_{psr}$ are residual vectors in metres.

After linearization, equation (4.3) can be written as:

$$r = A \, \delta x + w,\tag{4.4}$$

where,

$r = -\begin{bmatrix} \Delta\nabla\varepsilon_{ph} \\ \Delta\nabla\varepsilon_{psr} \end{bmatrix}$ is the residual vector. The variance-covariance matrix of the

observations is denoted as C_l,

δx is the correction to the approximated state vector,

A is the partial derivatives of the observations with respect to the state vector

(called   the first order design matrix), and

w is the misclosure after linearization.

To obtain the solution of δx, the weighted parametric least-squares adjustment can be applied. The normal equation is:

$$P \, \hat{\delta x} = u,\tag{4.5}$$

where,

$\hat{\delta x}$ is the least-squares estimation using all information up to the current epoch, and

$$u = A^T C_l^{-1} w + C_x^{-1} \delta x_0,$$
$$P = A^T C_l^{-1} A + C_{x,}^{-1},\tag{4.6}$$

where,

δx_0 is the variation of the *a-priori* state vector from the point where the

linearization is   made, and

$C_{x,}$ is the variance-covariance matrix of x_0.

For the combined solution of the phase and code,

$$A^T C_l^{-1} A = A_{ph}^T C_{ph}^{-1} A_{ph} + A_{psr}^T C_{psr}^{-1} A_{psr},$$
$$A^T C_l^{-1} w = A_{ph}^T C_{ph}^{-1} w_{ph} + A_{psr}^T C_{psr}^{-1} w_{psr},\tag{4.7}$$

where,

A_{ph}, C_{ph}, and w_{ph} correspond to the phase observable, and

A_{psr}, C_{psr}, and w_{psr} correspond to the pseudorange observable.

For the code only solution,

$$A^T C_l^{-1} A = A_{psr}^T C_{psr}^{-1} A_{psr} \text{ , and}$$
$$A^T C_l^{-1} w = A_{psr}^T C_{psr}^{-1} w_{psr} \text{ .} \tag{4.8}$$

For the phase only solution,

$$A^T C_l^1 A = A_{psr}^T C_{psr}^{-1} A_{psr} \text{ , and}$$
$$A^T C_l^{-1} w = A_{ph}^T C_{ph}^{-1} w_{ph} \text{ .} \tag{4.9}$$

If there is no *a-priori* information, equation (4.6) becomes:

$$u = A^T C_l^{-1} w \text{ , and}$$
$$P = A^T C_l^{-1} A \text{ .} \tag{4.10}$$

In a kinematic application, the observation equation can be linearized around the *a-priori* state vector. Then, equation (4.6) is reduced to:

$$u = A^T C_l^{-1} w \text{ , and}$$
$$P = A^T C_l^{-1} A + C_x^{-1} \text{ .} \tag{4.11}$$

The state vector normally includes position, velocity, and ambiguity parameters. The ambiguity parameters are not included in the state vector if no phase observations are used or if ambiguity parameters are fixed.

Since only the difference between ambiguity parameters will affect the observation equation (4.3), the ambiguity parameters are defined as the differences between a specific satellite and all other satellites. The differences are defined in a way such that each double difference ambiguity in equation (4.3) can be uniquely and linearly combined by these parameters.

The components of the design matrix are listed in Table 4.1, assuming that the double difference observations are made with respect to the first satellite.

The variance-covariance matrix for the double difference phase or the code observations is:

$$C = 2\,\sigma_0 \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 2 \end{bmatrix},$$

(4.12)

where,

σ_0 is the *a-priori* variance for undifferenced observations, and

C can be inverted by the general method. However, the inversion can be computed simply using the following formula:

$$C^{-1} = 1/2\,\sigma_0^{-1} \begin{bmatrix} n/(n+1) & -1/(n+1) & \cdots & -1/(n+1) \\ -1/(n+1) & n/(n+1) & \cdots & -1/(n+1) \\ \cdots & \cdots & \cdots & \cdots \\ -1/(n+1) & -1/(n+1) & \cdots & n/(n+1) \end{bmatrix}.$$

(4.13)

In kinematic positioning, the *a-priori* state vector is normally computed from the estimated state vector of the previous epoch by using a prediction model (see next sub-section).

Table 4.1   Partial Derivatives with Respect to the Components of the State Vector.
(Assuming the double difference are made with respect to 1'st satellite)

Observable	Pseudorange	Carrier phase
With respect to coordinates	where,	$$\begin{bmatrix} a^{1T} \\ a^{2T} \\ ... \\ a^{nT} \end{bmatrix}$$ where, $a^i = u^i{}_R - u^0{}_R$, $i = 1,2,..,n$. (Assumed $n+1$ satellites available). $u^i{}_R$ and $u^0{}_R$ are unit vectors from satellite i and 0 to the remote station.
With respect to velocity		~ 0
With respect to ambiguities (Only if ambiguities are estimated)	0	A special case if the first satellite is the reference satellite and the order of the satellites is the same as the order of the ambiguity parameters: $$\lambda \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}.$$ If not, each element at row i and column j is: $$\lambda \, (\delta^i{}_j - \delta^0{}_j),$$ where, λ is the wavelength, $$\delta^i{}_j = \begin{cases} 1, \text{if } i \text{ and } j \text{ are referred to the same satellites} \\ 0, \text{otherwise} \end{cases}$$
n is the number of the double difference observations		

4.2.2   Prediction of the State Vector

Assuming that the state vector contains the coordinates, velocities, and ambiguity parameters, the predicted state vector is then:

58

$$\hat{x}_{k+1|k} = \begin{bmatrix} \hat{x}c_{k+1|k} \\ \hat{x}v_{k+1|k} \\ \hat{x}n_{k+1|k} \end{bmatrix} = \begin{bmatrix} \hat{x}c_{k|k} + \hat{x}v_{k|k}\,(t_{k+1} - t_k) \\ \hat{x}v_{k|k} \\ \hat{x}n_{k|k} \end{bmatrix} + \begin{bmatrix} \varepsilon c_{k+1|k} \\ \varepsilon v_{k+1|k} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} I & (dt\,I) & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}c_{k|k} \\ \hat{x}v_{k|k} \\ \hat{x}n_{k|k} \end{bmatrix} + \begin{bmatrix} \varepsilon c_{k+1|k} \\ \varepsilon v_{k+1|k} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} I & (dt\,I) & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \hat{x}_{k|k} + \begin{bmatrix} \varepsilon c_{k+1|k} \\ \varepsilon v_{k+1|k} \\ 0 \end{bmatrix} , \tag{4.14}$$

where,

$\hat{x}_{k+1|k}$ is the state vector at epoch $k+1$, estimated from all the information up to

epoch k (*i.e.*, the predicted state vector at epoch $k+1$),

$\hat{x}c_{k+1|k}, \hat{x}v_{k+1|k},$ and $\hat{x}n_{k+1|k}$ are the components of $\hat{x}_{k+1|k}$ for the coordinates,

velocity, and the ambiguity parameters correspondingly,

$\hat{x}c_{k|k}, \hat{x}v_{k|k},$ and $\hat{x}n_{k|k}$ are the state vectors at epoch k estimated from the

information up to epoch k, and

$\varepsilon c_{k+1|k}$ and $\varepsilon v_{k+1|k}$ are the noise of the prediction model on the coordinate and

velocity vectors.

From the law of error propagation, the variance-covariance of $\hat{x}_{k+1|k}$ is:

$$C\hat{x}_{k+1|k} = \begin{bmatrix} I & (dt\,I) & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} C\hat{x}_{k|k} \begin{bmatrix} I & 0 & 0 \\ (dt\,I) & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} C\varepsilon c_{k+1|k} & C\varepsilon cv_{k+1|k} & 0 \\ C\varepsilon cv_{k+1|k}^T & C\varepsilon v_{k+1|k} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} I & (dt\,I) & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} Cc_{k|k} & Ccv_{k|k} & Ccn_{k|k} \\ Ccv_{k|k}^T & Cv_{k|k} & Cvn_{k|k} \\ Ccn_{k|k}^T & Cvn_{k|k}^T & Cn_{k|k} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ (dt\,I) & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$+ \begin{bmatrix} C\varepsilon c_{k+1|k} & C\varepsilon v_{k+1|k} & 0 \\ C\varepsilon c v_{k+1|k}^{T} & C\varepsilon v_{k+1|k} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} Cc_{k|k}+dt\,(Ccv_{k|k}^{T}+Ccv_{k|k})+dt^{2}Cv_{k|k} & Ccv_{k|k}+dt\,Cv_{k|k} & Ccn_{k|k}+dtCvn_{k|k} \\ (Ccv_{k|k}+dt\,Cv_{k|k})^{T} & Cv_{k|k} & Cvn_{k|k} \\ Ccn_{k|k}^{T}+dt\,Cvn_{k|k}^{T} & Cvn_{k|k}^{T} & Cn_{k|k} \end{bmatrix}$$

$$+ \begin{bmatrix} C\varepsilon c_{k+1|k} & C\varepsilon v_{k+1|k} & 0 \\ C\varepsilon c v_{k+1|k}^{T} & C\varepsilon v_{k+1|k} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{4.15}$$

where,

I is the unit matrix,

dt is the time difference between epochs $k+1$ and k,

$Cc_{k|k}$, $Ccv_{k|k}$, $Ccn_{k|k}$, $Cv_{k|k}$, $Cvn_{k|k}$, $Cn_{k|k}$ are the components of the variance-covariance matrix corresponding to the estimated state vector $\hat{x}_{k|k}$, and

$C\varepsilon c_{k+1|k}$, $C\varepsilon cv_{k+1|k}$, and $C\varepsilon v_{k+1|k}$ are the variance-covariance matrices of the prediction noise for coordinates and velocity.

If the velocity is not included in the state vector, all components corresponding to the velocity become null. The same happens if the ambiguity parameters are not included.

4.3 Computation of the Uncertainty Range Using Least-Squares

In the previous sections, a specific method to compute the search range was not mentioned. In this section, a least-squares filter is applied to compute the search range. The predicted state vector and the direct GPS observations are all regarded as observables in the estimation process.

When the cycle ambiguity parameters are estimated as real values, the system of normal equations can then be derived using the parametric adjustment method as follows (Section 4.2.1):

$$P_{CC} \, \hat{\delta x}_C + P_{CN} \, \hat{\delta x}_N = u_C \text{, and}$$

$$P_{NC} \, \hat{\delta x}_C + P_{NN} \, \hat{\delta x}_N = u_N \text{,} \tag{4.16}$$

where,

$\hat{\delta x}_N$ is the estimated correction of the ambiguity parameters to be estimated as the real values,

$\hat{\delta x}_C$ is the estimated correction of the other parameters,

P_{CC}, P_{CN}, P_{NC}, and P_{NN} are the sub-matrices of the normal matrix partitioned corresponding to x_C and x_N, and

u_C and u_N are the constant terms in the normal equations.

By eliminating the parameter $\hat{\delta x}_C$ from the second part of equation (4.16), the following equation can be obtained:

$$\tilde{P}_{NN} \, \hat{\delta x}_N = \tilde{u} \text{,} \tag{4.17}$$

where,

$$\tilde{P}_{NN} = P_{NN} - P_{NC} \, P_{CC}^{-1} \, P_{CN} \text{, and}$$

$$\hat{u}_N = u_N - P_{NC} \, P_{CC}^{-1} \, u_C \text{.} \tag{4.18}$$

Assuming that ambiguities $N_1, N_2.., N_{i-1}$ have been assigned integer values, denoted as N_I, a least-squares solution is now required for $N_i, ..., N_n$ while treating $N_1, N_2.., N_{i-1}$ as known values.

A tree of filters can be used. Each filter corresponds to each different set of ambiguity values. However, such an approach will complicate programming and computation. A

simpler solution is to convert the solution of equation (4.17) to an equivalent solution of partially fixed ambiguities using the parameter removal theorem discussed in the last chapter.

To get the filter solution when ambiguities $N_1, N_2, ..., N_{i-1}$ are treated as known integers, the normal equations are partitioned before ambiguity fixing. That is, the full float estimated ambiguity parameters are partitioned into two parts, one corresponding to $N_1, N_2.., N_{i-1}$, denoted as $\delta\hat{x}_{Na}$ (the vector form of $\delta\hat{x}_{N1}, ..., \delta\hat{x}_{Ni-1}$); the other corresponding to $N_i, ..., N_n$, denoted as $\delta\hat{x}_{Nb}$ (the vector form of $\delta\hat{x}_{Ni}, ..., \delta\hat{x}_{Nn}$). Then, equation (4.17) can be partitioned like the partition in equation (4.16):

$$\tilde{P}_{Naa}\,\delta\hat{x}_{Na} + \tilde{P}_{Nab}\,\delta\hat{x}_{Nb} = \hat{u}_{Na}\text{, and}$$
$$\tilde{P}_{Nba}\,\delta\hat{x}_{Na} + \tilde{P}_{Nbb}\,\delta\hat{x}_{Nb} = \tilde{u}_{Nb}\text{,} \qquad (4.19)$$

where,

$\tilde{P}_{Naa}, \tilde{P}_{Nab}, \tilde{P}_{Nba}$, and $\tilde{P}_{Nbb}$ are the sub-matrices of the normal matrix ($\tilde{P}_{NN}$) partitioned corresponding to $\delta\hat{x}_{Na}$ and $\delta\hat{x}_{Nb}$, and

$\hat{u}_{Na}$ and $\hat{u}_{Nb}$ are the constant terms in the partitioned normal equations.

Note: In the above equations, all ambiguity parameters are estimated as real.

When x_{Na} becomes known, denoted as N_I, values of $N_i, ..., N_n$ are denoted as $\delta\hat{x}_{Nb|N1\,N2\,Ni-1}$. According to the theorem in Chapter 3, when x_{Na} is treated as the fixed value vector, the corresponding normal equation after eliminating the non-ambiguity parameters becomes:

$$\tilde{P}_{Nbb},\,\delta\hat{x}_{Nb|N1,\,N2,\,...,\,Ni-1} = \tilde{u}_{Nb} - \tilde{P}_{Nba}\,(N_I - x_{Na}{}^o)\text{,} \qquad (4.20)$$

where

$x_{Na}{}^o$ is the approximate value that δx_{Na} is referred to.

Therefore, the float estimation of remaining ambiguity parameters is:

$$\hat{\delta x}_{Nb|N1, N2, ..., Ni-1} = \tilde{P}_{Nbb}^{-1} \left(\hat{u}_{Nb} - \tilde{P}_{Nba} (N_I - x_{Na}{}^o) \right) . \tag{4.21}$$

The corresponding variance-covariance matrix is $\tilde{P}_{Nbb}{}^{-1}$.

From equation (4.19), the float estimation of $\hat{\delta x}_{Nb}$ without fixing any ambiguity, is:

$$\hat{\delta x}_{Nb} = \tilde{P}_{Nbb}^{-1} \left(\hat{u}_{Nb} - \tilde{P}_{Nba} \hat{\delta x}_{Na} \right) . \tag{4.22}$$

By differencing between equations (4.21) and (4.22), the following can be obtained:

$$\hat{\delta x}_{Nb|N1, N2, ..., Ni-1} - \hat{\delta x}_{Nb} = \tilde{P}_{Nbb}^{-1} \tilde{P}_{Nba} \left(\hat{\delta x}_{Na} - (N_I - x_{Na}{}^o) \right) . \tag{4.23}$$

Since $\hat{x}_{Na} = \hat{\delta x}_{Na} + x_{Na}{}^o$, the above equation becomes:

$$\hat{\delta x}_{Nb|N1, N2, ..., Ni-1} - \hat{\delta x}_{Nb} = \tilde{P}_{Nbb}^{-1} \tilde{P}_{Nba} \left(\hat{x}_{Na} - N_I \right). \tag{4.24}$$

That is,

$$\hat{x}_{Nb|N1, N2, ..., Ni-1} - \hat{x}_{Nb} = \tilde{P}_{Nbb}^{-1} \tilde{P}_{Nba} \left(\hat{x}_{Na} - N_I \right) . \tag{4.25}$$

Also since only the float estimation of N_i is of concern, only the first element of $\hat{\delta x}_{Nb|N1 N2 Ni-1}$ is to be computed. That is,

$$\hat{x}_{Ni|N1, N2, ..., Ni-1} = \hat{x}_{Ni} + a_i{}^T \left(\hat{x}_{Na} - N_I \right) , \tag{4.26}$$

where,

$$a_i{}^T \quad = 1\text{'st row of } (\tilde{P}_{Nbb}^{-1} \tilde{P}_{Nba})$$
$$= 1\text{'st row of } (\tilde{P}_{Nbb}^{-1}) \ \tilde{P}_{Nba} .$$

The variance of the estimated $\hat{\delta x}_{Nb|N1 N2 Ni-1}$ corresponds to the elements of the first row and the first column of $\tilde{P}_{Nbb}{}^{-1}$ multiplied by the variance factor (σ_o). That is,

$$\sigma_{x_{Ni}} = \sigma_o \, [\, \tilde{P}_{Nbb}^{-1}]_{11} , \tag{4.27}$$

As seen, only the first row of $\tilde{P}_{Nbb}{}^{-1}$ needs to be computed.

After the float solution and its variance is obtained, the maximum uncertainty range for possible integer values of x_{N_i} can be computed as:

$$x_{N_i|N_1, N_2, ..., N_{i-1}} - \xi\, \sigma_{x_{N_i|N_1, N_2, ..., N_{i-1}}} \leq N_i \leq$$

$$x_{N_i|N_1, N_2, ..., N_{i-1}} + \xi\, \sigma_{x_{N_i|N_1, N_2, ..., N_{i-1}}} \,, \tag{4.28}$$

where,

ξ is the expansion factor from the standard deviation to the maximum possible uncertainty.

Since both $\boldsymbol{a}_i$ and $\xi\, \sigma_{x_{N_i|N_1, N_2, ..., N_{i-1}}}$ are independent on the values of $N_1, N_2, ..., N_{i-1}$, both can be computed outside the search loops. Theoretically, ξ can be related to a probability level if the noises in the GPS observations are uncorrelated in the time domain. However, this is not always true. The noise of a GPS observable is correlated from epoch to epoch and the form of the correlation is unknown. Therefore, the probability level based on the assumption of independent observations is not suitable here. For the first ambiguity, N_1, x_{N1} and σ_{xN_1} are computed without fixing any other ambiguities.

As the uncertainty range is linearly proportional to both the *a-priori* standard deviation of the observation and the expansion factor, the increase or decrease of the expansion factor can also be equivalently achieved by the same amount of the increase or decrease in the expansion factor. However, this increase or decrease will make the *a-priori* standard deviation deviate from the actual level of the noise and the biases. Therefore, it is better to keep *a-priori* standard deviation close to the actual value and adjust the expansion factor so that the correct ambiguities will be included in all range checks described in equation (4.28).

According to the parameter removal theorem in Chapter 3, the weighted sum of the squared residuals can be computed as:

$$\Omega|_{\mathbf{x}_N=N_F} = \Omega + (N_F - \hat{\mathbf{x}}_N)^T \tilde{\mathbf{P}}_{NN} (N_F - \hat{\mathbf{x}}_N) , \qquad (4.29)$$

where,

N_F is the potential solution of the full ambiguity set,

$\Omega|_{\mathbf{x}_N=N_F}$ is the adjusted Ω corresponding to fixed ambiguity values,

$\hat{\mathbf{x}}_N$ is the float ambiguity vector from the least-squares estimation, and

Ω is the sum of the squared residuals while all ambiguity parameters are estimated as real values.

4.4 Expansion Factor

The presumption for the standard deviation of the float ambiguities is that the noise in the observation should be uncorrelated from epoch to epoch. In this ideal case, equation (4.28) can be explained by a statistical confidence interval [Vaníček and Krakiwsky, 1987]. Assuming that the observation error is Gaussian distributed and uncorrelated over the time, in order to achieve a probability level of $1-\alpha$ so that a correct solution is included as a potential solution satisfying equation (4.28), the expansion factor in the equation can be expressed as:

$$\xi = \xi_{1-\alpha/(2n)} , \qquad (4.30)$$

where,

$\xi_{1-\alpha/(2n)}$ is the expansion factor of individual two-tailed statistical test with the significant level of α/n, and

n is the number of ambiguities.

With six ambiguity parameters, and the confidence level of 0.99, the expansion factor will be 3.765. Similarly, the expansion factors 3 and 15 correspond to confidence levels of 0.984 and .99999999 for six ambiguities.

In many cases, there is a strong correlation between the errors from different epochs. The standard deviations derived from least-squares may not be able to represent the uncertainties correctly in many cases. This correlation makes the estimated standard deviation too optimistic. Typically, a value between 3 and 15 is used for the least-squares FASF. The exact value is dependent on the correlation over time. If the correlation is large and only a small value of the expansion factor is used, the correct ambiguities are most likely to be excluded from the search ranges. Therefore, if the correlation level is unknown, a conservative value of 15 is recommended. However, if the biases in the observations are small and a large expansion factor is used, it will take longer.

The following two simplified examples may help to understand the problem better. The correct values in the examples are zeros and are supposed to be unknown; the measured values are 0.00 + noise; there are 401 measurements. The purpose in presenting these problems is to derive the correct integer value from equation (4.28).

Example 1: Pure white noise. The *a-priori* standard deviation of the measurement error is σ=1.00 (Figure 4.3). The value from the least-squares estimation is -0.025 with the standard deviation of $\sigma/20$=0.05 (from the *a-priori* standard deviation). With a confidence level of .99, the expansion factor is 3.765. By applying equation (3.32), the search range is from -0.214 to 0.164. The only integer value included is zero. In this case, the correct value is obtained.
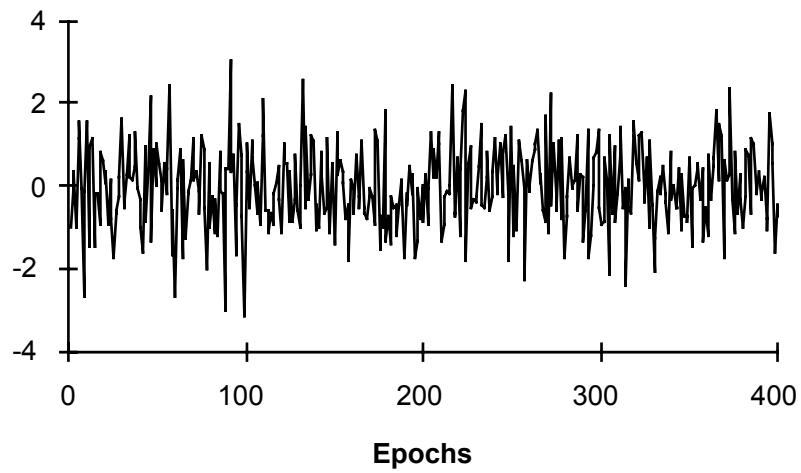
Figure 4.3   White noise

The adjusted value is 1.0. With a confidence level of .99, the search range is from 0.811 to 1.189. Only the incorrect solution is included.

Example 2: Constant noise with value of one (see Figure 4.4). The *a-priori* standard deviation is $\sigma=1$.
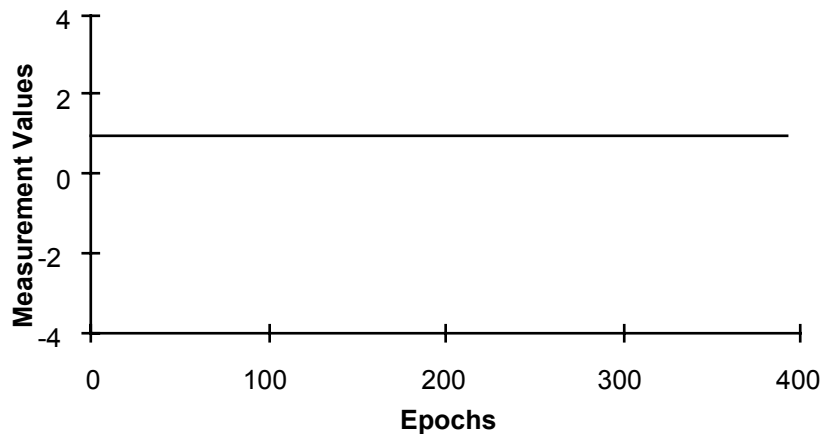
Figure 4.4   Constant Noise

For the strongly correlated observations from epoch to epoch, the standard deviation from the least-squares estimation for the partially assumed ambiguities is too optimistic. This problem can be partially considered by the proper selection of the expansion factor in

converting the standard deviation to the maximum possible error for the float estimation. However, the difficulty here is that this expansion factor is highly dependent on the correlation of errors in the observations. The exact relationship between the correlation and the expansion factor is also unknown. Furthermore, no current models exist on the relationship between the correlation and the environment. The selection of a universal value is still a problem to be solved. The correlation coupled with the geometry makes the estimated standard deviation even more complicated. The relationship between the geometry, error pattern, success rate, and time required to resolve ambiguities also needs to be further investigated. Under strongly correlated observation noise, the longer the period of the observations for the float solution, the more serious the distortion.

The least-squares FASF will work fine where only a few epochs of observations are needed to resolve ambiguities since the distortion of standard deviation will be serious only when extensive observations are accumulated. For example, in the case of favorable geometry, the FASF should be good even if there is strong correlation in the observations from epoch to epoch. Also, for dual frequency receivers, where the wide lane ambiguities can be formed either explicitly or implicitly, FASF will also work properly because of the relatively small size of the search range in terms of cycles.

CHAPTER 5   TESTING AND ANALYSIS OF FASF

The theoretical advantage of FASF is tested with four data sets. One of the data sets is static, but processed as if it were kinematic. The other three sets are kinematic data from land, air, and sea experiments. Some of the kinematic data have previously been used for ambiguity resolutions based on a modified Hatch's algorithm. All these data sets are analyzed using a program developed by the author based on the FASF concept and its least-squares implementation.   The results of these analyses are presented here in the comparison with previous results.

The static data set is analyzed in Section 5.1; the land kinematic experiment is analyzed in Section 5.2; the air experiment is presented in Section 5.3; the marine data is examined in Section 5.4. The threshold of maximum number of potential solutions (Chapter 4) is taken as ten in all the analysis.

5.1   Kinematic Analysis of Static Data

5.1.1   Description of Test Data

The experiment was conducted on February 12, 1993, on the Springbank Test Range, using NovAtel L1 GPSCard™ (model 951R) units. In this analysis, a static data over a short baseline was used.

Although the code observations of the NovAtel GPSCard™ are very precise (Cannon and Lachapelle, 1992), this analysis is intended to demonstrate only the computational

advantage of FASF and thus, the code observations are not used. The inclusion of code measurements will be discussed in later sections.

5.1.2   Data Analysis

A 486/50 PC was used for the data reduction. The characteristics of the data processing are listed in Table 5.1. The standard deviation σ_{ph} was assigned arbitrarily.

Table 5.1   The specifications of Data Processing Characteristics

σ of initial state	σ of prediction for x	σ of the prediction for v	σ_{ph}	Mask angle
33 m	3 m	3 m/s	1.5 cm	10 degrees

The computation time versus the number of epochs is shown in Figure 5.1. The short vertical lines represent the time required to search the integer ambiguities. The time needed for ambiguity searching is either zero or 0.06 second. The zero time is caused by the resolution of the PC system time.
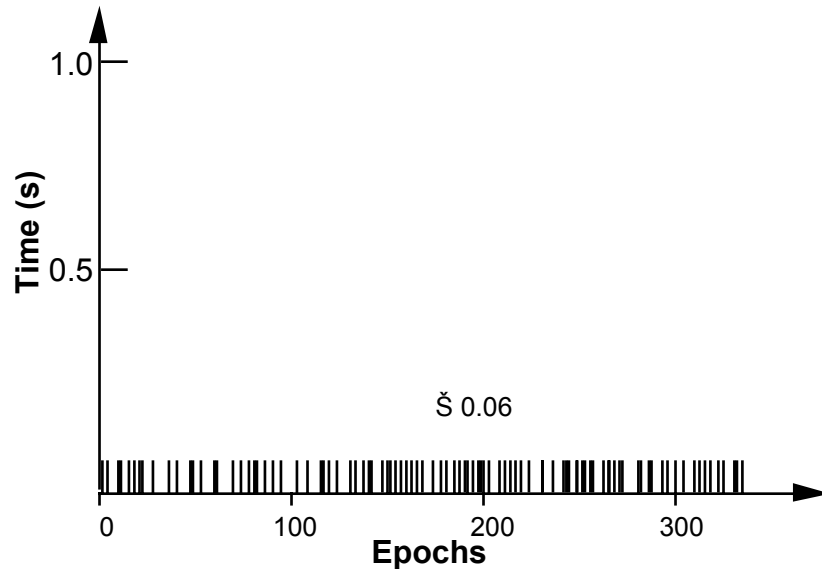
Figure 5.1   Computation Time for Epoch by Epoch Ambiguity Search

70

A comparison of the computation time with other methods is shown in Table 5.2. As shown, FASF has a clear advantage in terms of computational time. The time required for Landau and Euler's method is an optimistic estimation. At the initial search epochs, the time required in Landau and Euler's method is much longer than the values presented here.

Table 5.2   Computation Time: FASF *vs.* Others
(six Satellites, 486/50 Computer)

range in cycles / methods	±10	±50	±500
FASF	0.06s	0.06s	0.06s
Landau & Euler	0.1*	1.5*	71*
Hatch	0.2**	3x10**	3x10^{4}**
Full search	9x10***	3x10^{5}***	3x10^{10}***

* Estimated from Landau and Euler's paper, [1992]. The time required for the window of ±500 in their method was projected from the time required for the windows of ±10 and ±50. The computation formula is $1.5*(W/50)^{1.68}$, where, W is the size of the window.

** The time for the window of ±50 was experienced by Lu [1993], using the 1993 version of FLYKIN^{TM}[Lachapelle et al. 1993]. The other two values were computed by $30*(W/50)^{3}$.

*** The time for the full search method was computed by the time required for Hatch's method multiplied by $(2*W)^{n-3}$, where, n was the number of the ambiguity parameters.

The variations from the FASF solution in height are depicted in Figure 5.2. As seen, before the ambiguities are fixed, the float solution deviates by as much as 10 m. However, after correctly fixing the ambiguities, the deviation becomes much smaller. The convergence time, the time required to resolve ambiguities, is about 1100 seconds. As stated previously, no code measurements were used in this analysis. The standard variation of the *a-priori* coordinates is 33 metres. That could be improved if code information was used.

The variations of the updated heights using the fixed ambiguity parameters from the OTF solution are shown in Figure 5.3. No systematic errors larger than 1.5 cm are evident.

The small deviations are mostly caused by carrier phase multipath. A constant value of 1184.39 m has been subtracted from the height in Figures 5.2 and 5.3. The plots in these figures are thinned out by a factor of five. The same was also applied to the carrier phase residuals in Figure 5.4. Similarly in Figure 5.4, there is no noticeable long term systematic effect. The small periodic variations are mostly caused by multipath effects.
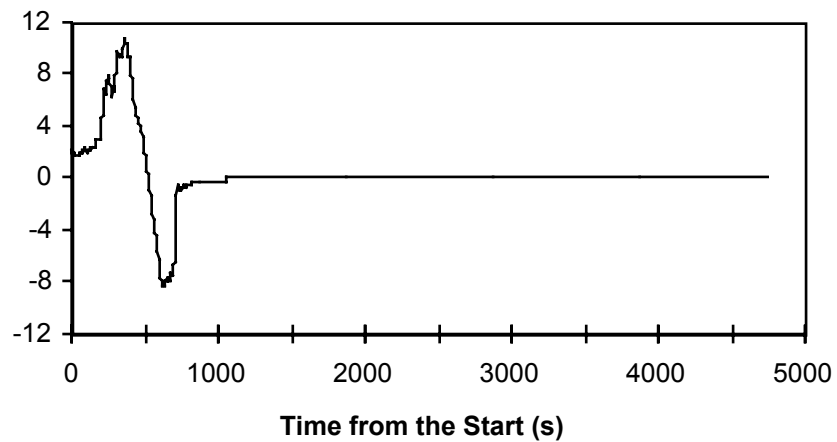
Figure 5.2   Height Variations before and after Correct Ambiguity Fixing
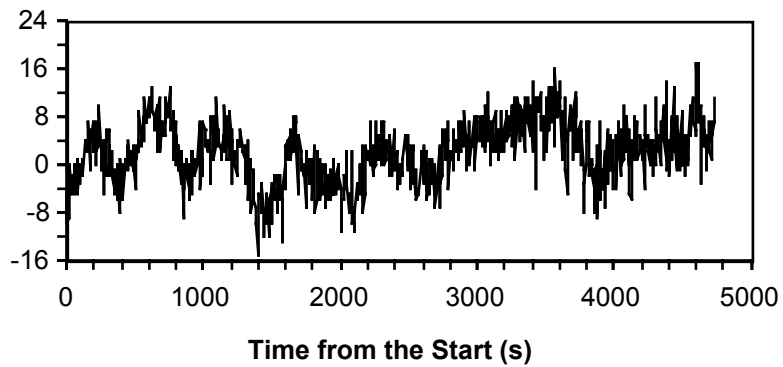
Figure 5.3   Height Variations Using the Fixed Ambiguities

72
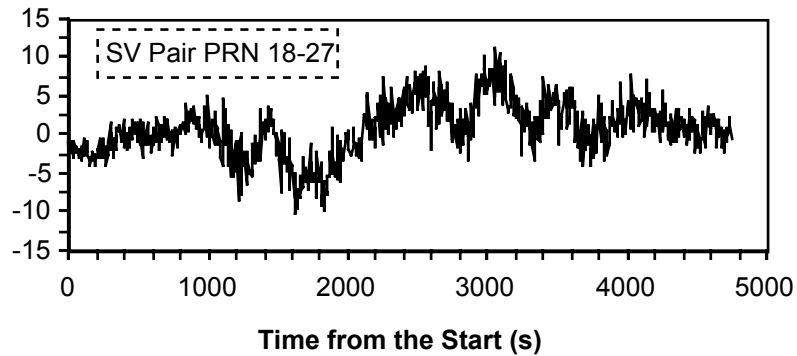
Figure 5.4   Residuals of Phase Observations using the Fixed Ambiguities

5.2   Kinematic Land Testing

5.2.1   Description of the Experimental Data

The data was obtained on Aug. 25, 1992, on the Springbank Test Range, located 20 km west of Calgary.   The experiment was originally intended to compare P-code and C/A-code L1 Narrow Correlator™ spacing receiver technologies to resolve ambiguities using Hatch's method [Lachapelle et al. 1993a].   In this analysis, only a portion of the L1 data between GPS times of 260218 and 261193 seconds is used.   The receivers were all NovAtel GPSCards™ (Model 951R) and antennas were geodetic type (Model 501) with choker-ring groundplanes.

The remote unit was mounted on a vehicle travelling at speeds of up to 70 km h^{-1}.   The distance between the reference station and the vehicle did not exceed a few kilometres (see Figure 5.5). A static initialization at the control points was performed at the start and end of the 15-minute trial to assess independently the correctness of the solutions.   The return to the original point was also used as a check on the ambiguity solution.
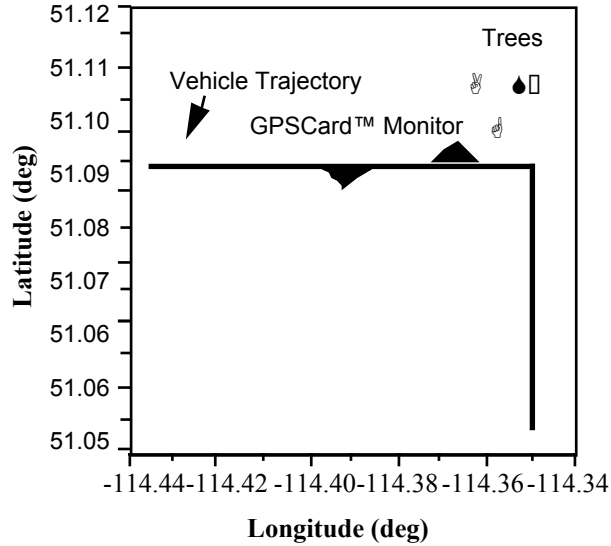
Figure 5.5 Trajectory of the Moving Unit (after Lachapelle et al. [1993a])

As indicated in the figure, trees located near the corner of the L-shaped trajectory created a multipath signal on the nearby highway. The period of the most multipath effect was between the GPS time of 260528 and 260758 seconds (see Lachapelle et al. [1993a] for the details). Seven satellites were available during the experiment and the PDOP was less than three.

5.2.2   Analysis of Results

The accuracy of the initial coordinates of the moving unit was set at 3m (1σ). The process noise of the moving vehicle, error of the predicted position and velocity, was set at 3 m (1σ) for the positioning components, and 3 ms^{-1} for the velocity components. Observation noise was set at 1.8 cm (1σ), the value previously used by Lachapelle et al. [1993a], and an expansion factor of three was used. The masking angle was set at 10° and the original data was collected at the rate of 1 Hz.

The observation time required for FASF ambiguity resolution OTF was investigated by conducting numerous computational trials on the kinematic portion of the data, each one

74

shifted in time by 10 s. The number of epochs required for each trial is shown in Figure 5.6. In the period corresponding to the strong multipath interference, significantly more epochs are required to resolve the cycle ambiguities. The statistics of such trials are summarized in Table 5.3. A comparison with a previous analysis using a modified Hatch's method by Lachapelle et al. [1993ᵃ] is summarized in Table 5.4. Two groups are listed, one was under the effect of the trees (i.e., between GPS times of 260528 and 260758 seconds); the other was clear from the trees. The number of trials, number of identical solutions, success rate (number of identical trials divided by total trials), and average epochs required for the success trials are listed. As seen, a substantial improvement is achieved with FASF. The impact of carrier phase multipath caused by trees on convergence time is still significant but much less than in the case of the least-squares search technique.
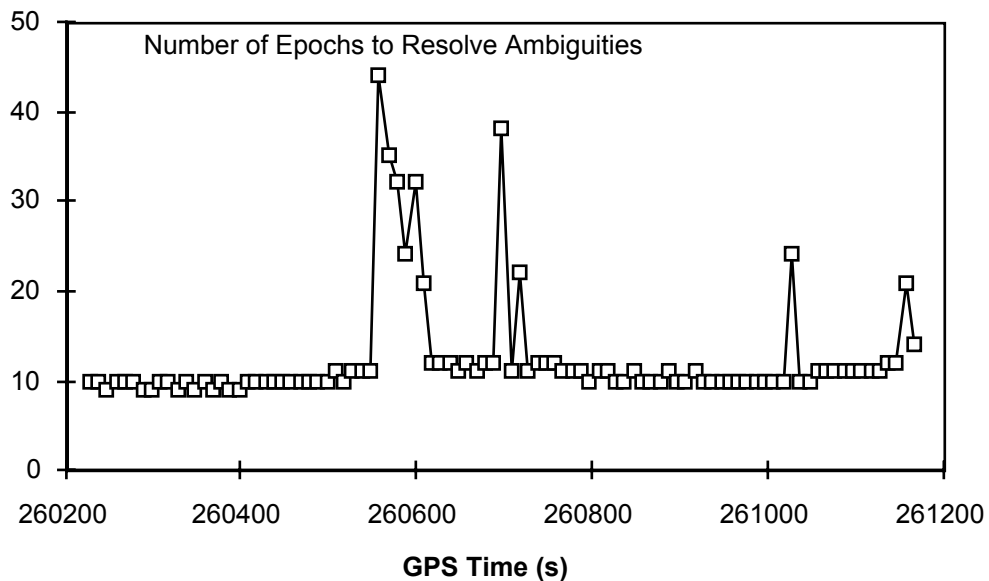
Figure 5.6   Epochs Required for Ambiguity Resolution Using Different Start Times

75

Table 5.3   Summary of the Ambiguity Resolution Using the FASF Method

Multipath	$\delta\Delta\nabla$	Number of Trials	Average Epochs Required	Success Rate
clear	1.8 cm	72	11	100%
trees	1.8 cm	24	18	100%

Table 5.4   Statistics of the Repeated Ambiguity Solutions Based on Hatch's Method [Lachapelle et al., 1993[a]].

Multipath	$\delta\Delta\nabla$	Number of Trials	Average Epochs Required	Success Rate
clear	1.8 cm	43	106	100%
tree	1.8 cm	30	500	100%

It should be pointed out that substantially more trials have been obtained in this analysis than those in Lachapelle et al. [1993[a]]. However, the same data was used. In the analysis by Lachapelle et al.[1993[a]], ambiguity solutions could not be obtained close to the end of the session when Hatch's least-squares method was used. These trials were not counted in Table 5.3. However, the ambiguity resolutions were still obtained at the very end of the session by using FASF. Another difference in the analysis is that the number of trials under the effect of trees is less than those in Lachapelle et al. [1993[a]].

Static ambiguity resolution was also made using the data at the beginning of the session. As expected, the same ambiguity resolution was obtained as those from kinematic solutions. The misclosure between the positions of the start and the end epochs was also compared.   The misclosure was at millimetre level in both height and horizontal components.
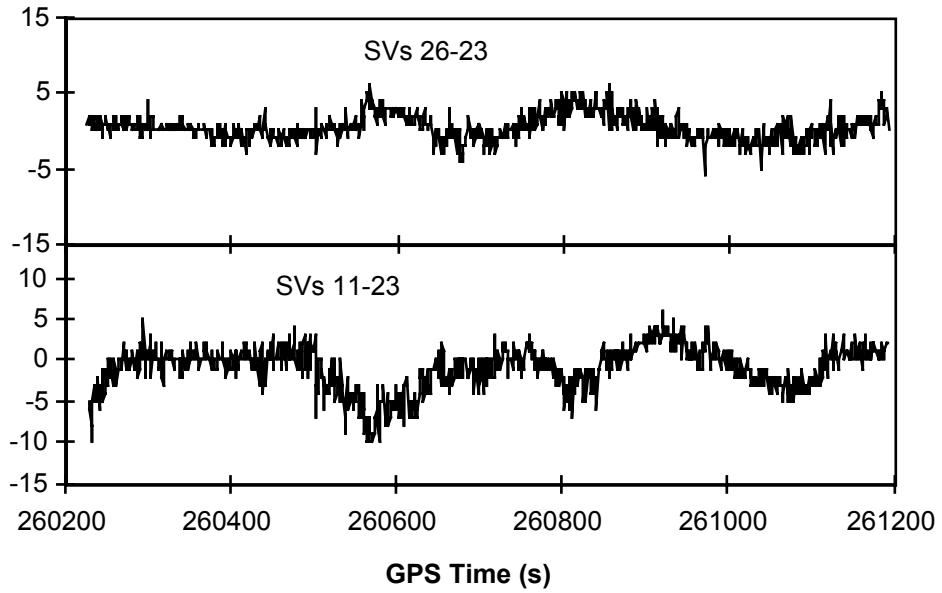
Figure 5.7   Residuals of Double Difference Phase Observations

Using the fixed ambiguities from OTF solution, the whole data set was reprocessed. The residuals of double difference phase measurements are shown in Figure 5.7.   The residuals obtained from the estimation process grow to about 10 mm during the trajectory segment affected by trees.   The fact that no long term trend is affecting the residuals shows fairly reliably that the correct ambiguity solution has been obtained. The periodic variations were mostly caused by multipath interference. The high frequency variation at the millimetre level was caused by the observation noise. The residuals of the C/A code measurements are shown in Figure 5.8. Most of the residuals are clearly below 1 m.
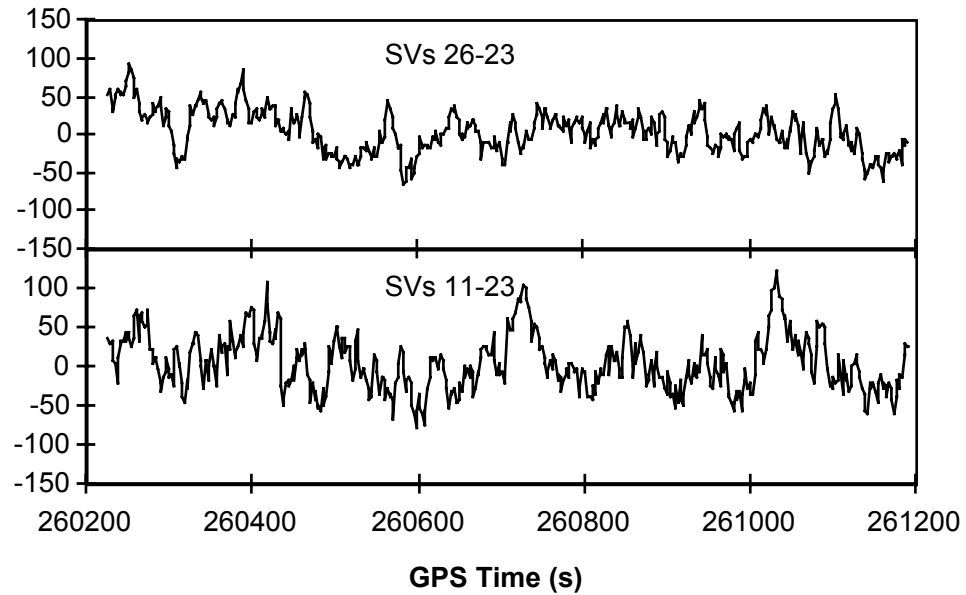
Figure 5.8   Residuals of Double Difference C/A Code Observations Using the Fixed
Ambiguities from OTF Solution

5.3   Airborne Testing

Data obtained with a pair of dual frequency Trimble 4000SSE receivers operating in

P-code mode was used in this case. The data was provided by GeoSURV Inc., of Ottawa,

and was collected in the Muskoka area of Ontario in Spring of 1993. The station, located

at an airport, was used as the base station for data analysis here. At the start, the distance

from the base station was up to 66 km and the height was up to 2.5 km (Figures 5.9). The

plane traveled at a speed of 80 m/s with occasional rates of up to 104 m/s (Figure 5.10).

The data between GPS time of 60914 and 65000 seconds, at GPS week 696, are analyzed

here.

The number of the visible satellites was usually seven (Figure 5.11). Towards the end of

the period, the number dropped to six and five. The number was less than five at some

periods and thus the ambiguity resolutions were not possible. The infinite PDOPs (Figure

5.12) [Wells et al., 1987] correspond to the time that less than four satellites were

available. The elevations of the observed satellites during this period are plotted in Figure 5.13.
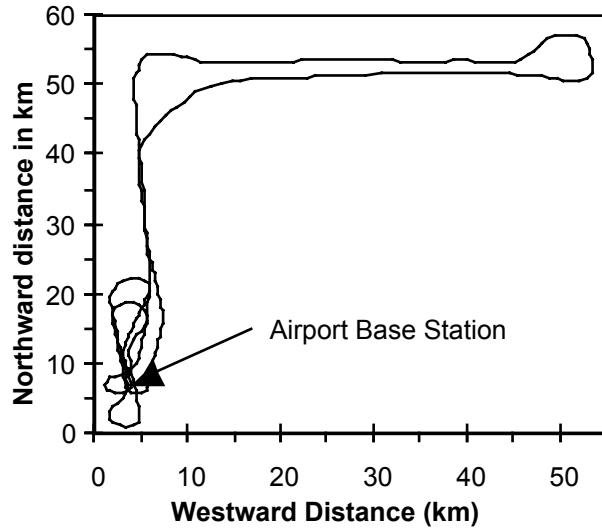
Figure 5.9   The Horizontal Trajectory of the Aircraft
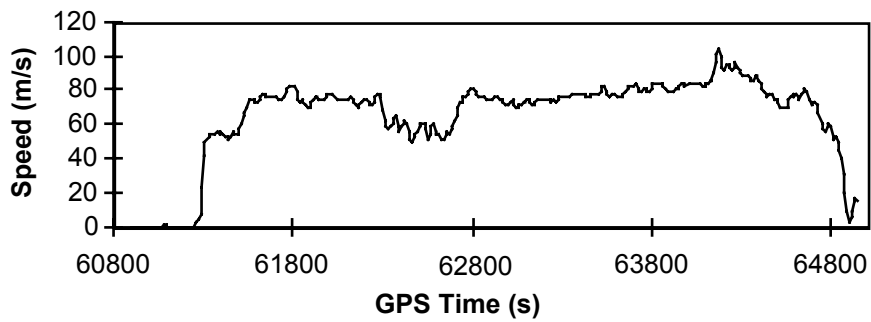
Figure 5.10   The Speed of the Aircraft.
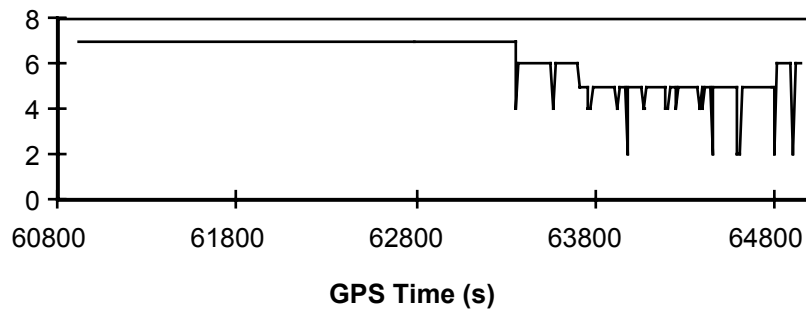
Figure 5.11   The Number of the Satellites Viewed by both Receivers
with Mask Angle of 10 Degrees

79
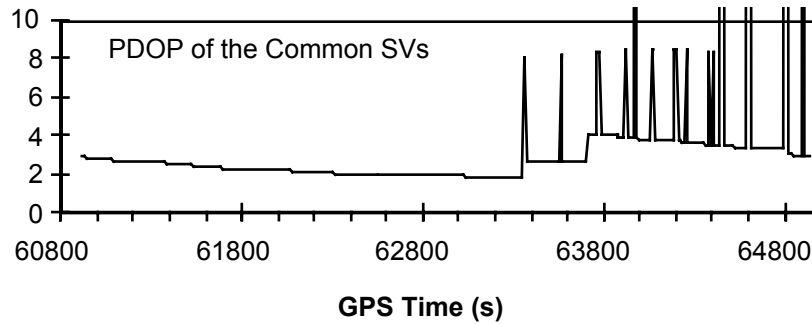
Figure 5.12   PDOPs from Observed Satellites with Mask Angles of 10 Degrees
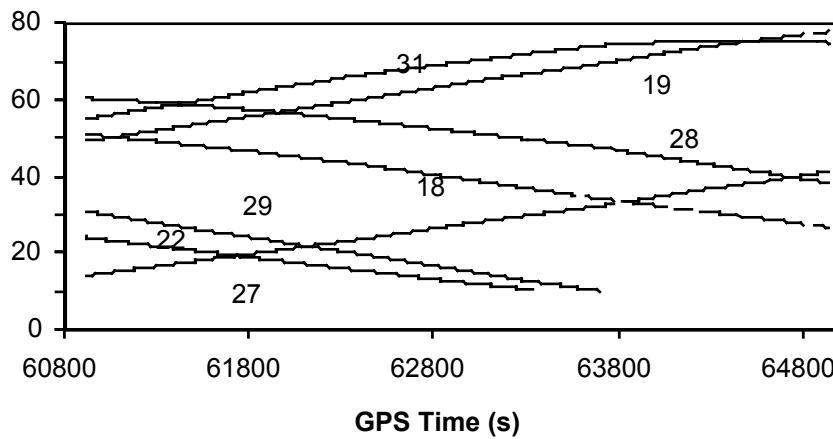
Figure 5.13   The Elevation of Observed Satellites with a Mask Angle of
10 Degrees

The wide lane combination of phase observations (L1-L2) was used. The mask angle was
ten degrees. The noise level was set at 2.8 cm for the double difference widelane phase
observable (1σ) and the expansion factor of three was used.

As in the case of the land test, the observation time required for ambiguity resolution
OTF was investigated by conducting numerous computational trials, each one shifted in
time by 10 s. The statistics of such trials are summarized in Table 5.5. The number of
epochs required for resolution during each trial is shown in Figure 5.15.   The wrong and
correct solutions are shown separately. Overall, the success rate was about 79%. The
same data was processed using FLYKIN™ based on modified Hatch's least-squares

method [Sun, 1994]. The results of this analysis are listed in Table 5.6. As seen, significantly more epochs were required to resolve ambiguities and furthermore, the success rate was lower.

Table 5.5   Statistics of the Ambiguity Resolution Using FASF

	$\delta_{\Delta\nabla}$	Number of Trails	Percentage	Number of Epochs Required
Success Trials	2.8 cm	274	79%	2.7
Wrong Trials	2.8 cm	77	21%	7.7

Table 5.6   Repeated Ambiguity Search Using Hatch's Method

Type of Solution	$\delta_{\Delta\nabla}$	Number of Trails	Percentage	Number of Epochs Required
Success Trials	2.8 cm	183	53.8%	61.377
Wrong Trials	2.8 cm	157	45.2%	258.433

The distance of the aircraft from the base station and its height are superimposed on the number of epochs (Figure 5.14) thereby showing the correlation between the number of epochs required for ambiguity resolution with the distance and height of the aircraft. As seen, most failures were at high altitude and further away from the base station. At high altitude, the effect of the un-modeled troposphere can not be reduced since the troposphere effects between the aircraft and the base station are different. Similarly, the ionospheric effect will increase when the distance and the height increase. The relatively lower number of the satellites during the later part also contributes to the wrong ambiguity resolution. Some of the failures were also caused by abnormal data such as half cycle slips.

The double difference phase residuals are shown in Figure 5.15. Most of the residuals are below 5 cm. The spikes at some epochs are caused by unresolved cycle slips. When

cycle slips occurred and could not be fixed, the data from that epoch was not used and the ambiguity search was re-initialized. Nonetheless, the residuals are still shown. Then. Residuals as large as 9 cm between the GPS time of 62311 and 62412 seconds were caused by the erroneous fixing of cycle slips. The residual variations of the C/A code measurements are shown in Figure 5.16. Most of the residuals are below 1.5 m.
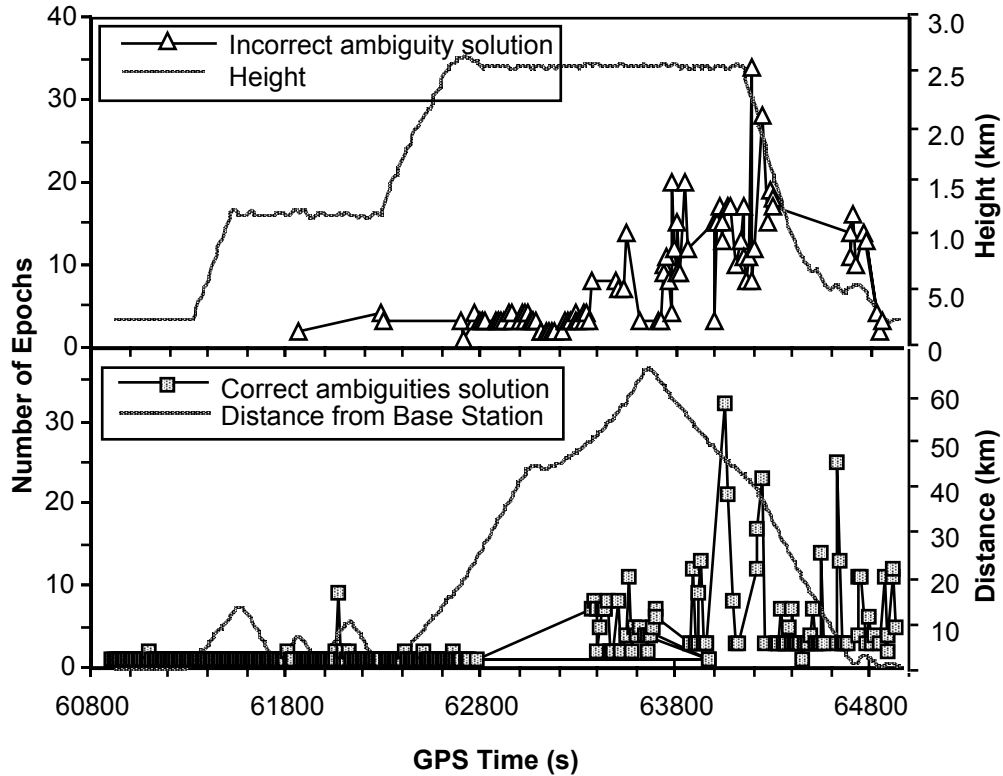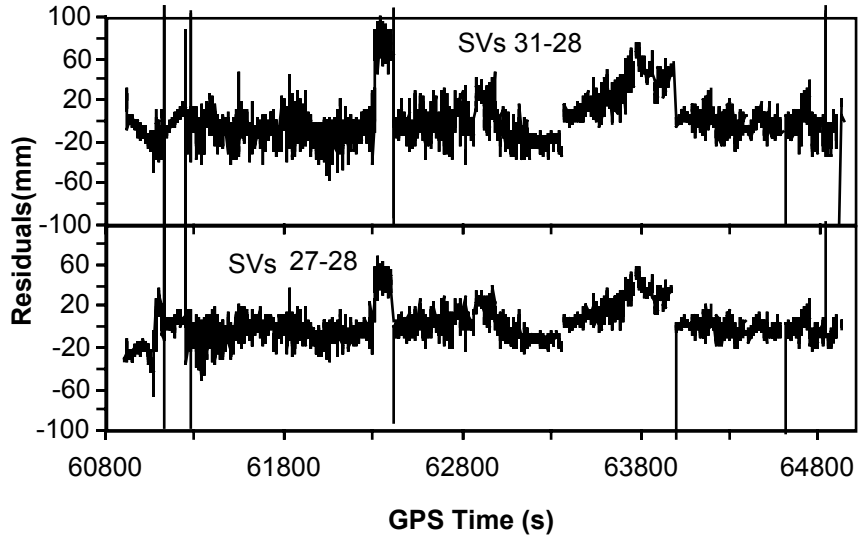

Figure 5.14 Number of Epochs Required to Fix Ambiguities Using FASF

Figure 5.15   Double Difference Phase Residuals
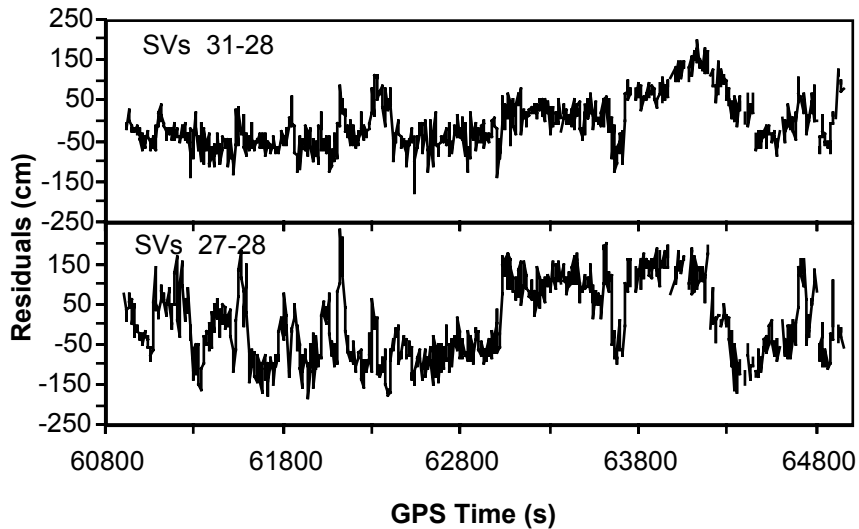
Figure 5.16   Double Difference Code Residuals

The same data was also processed with the *a-priori* standard deviation at 3 cm. In this case, the success rate was improved from 79% to 85%. This indicates that the standard *a-priori* weighting used in this processing was not optimal.

5.4   Marine Experiment

A marine experiment was conducted by The University of Calgary and the Canadian Hydrographic Service (Pacific Region) in early September 1992 in the Sidney, B.C., area

(see Figure 5.17) using a 12-m launch (Figure 5.18) [Lachapelle et al., 1993[b]]. Six satellites were available, and the PDOP varied between 1.9 and 2.6. The distance between the shore unit and the launch ranged from 10 to 24 km, and the speed ranged between 18 to 27 km h[-1] during the trial. The roll and pitch angles did not exceed five degrees.

Three GPSCard™ sensors were mounted on the launch. Both code and carrier phase data were recorded at a data rate of 2 Hz using PC laptops. The distances among the three GPSCard™ units were measured with an accuracy of about 1 cm as shown in Figure 5.18. These distances will be used later to independently check the double difference carrier phase ambiguities estimated between the shore antenna and each one of the three launch-based GPSCard™ antennas.
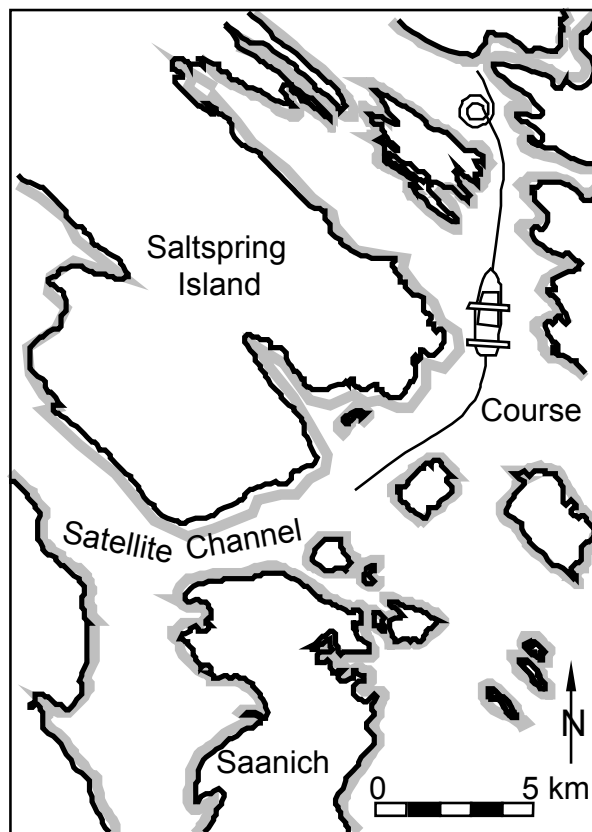
Figure 5.17 Launch Track Observed in the Marine Experiment in Sidney, B.C. Area
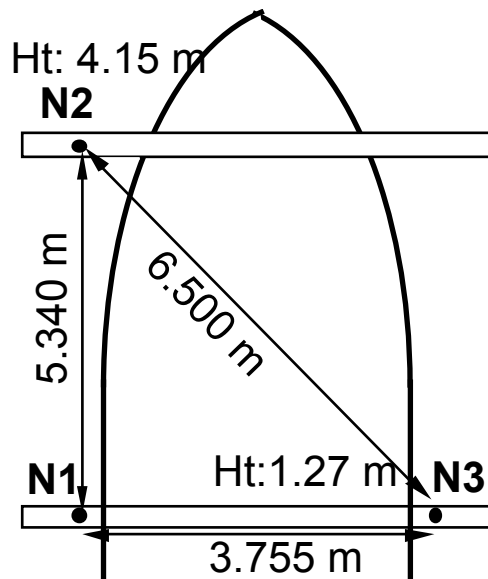
Figure 5.18   GPS Antenna Configuration on Launch

All the GPSCard™ antennas were equipped with choker-ring groundplanes except for one antenna on the survey launch. The use of such groundplanes has proven effective in minimizing multipath effects during previous experiments [Cannon and Lachapelle, 1992]. In this case, however, their use appears to make little difference on code multipath as shown by Lachapelle et al., 1993b.

A 40-minute data set previously used by Lachapelle et al. [1993b] is analyzed here. The analysis includes three parts: (1) the ambiguity search statistics; (2) the residuals; and (3) the distances between antennas. The data rate is 2 Hz. Both code and phase observables are used for the processing of this data. The mask angle is set at 10 degrees.

As in the previous analysis, the data was repeatedly processed. Several quasi-independent solutions were thus obtained. The results are summarized in Table 5.7. The previous analysis by Lachapelle et al. [1993b] is summarized in Table 5.8. As seen, the improvement is not as good as in the previous two cases. The results for antennas 2 and 3 are better when FASF is used, but the results for antenna 1 are about the same.

Table 5.7   Repeated Ambiguity Resolution (FASF)

Receivers	$\sigma_{\Delta\nabla}$ (Max)	Number of Trials	Success Rate	Average Period Required
GPSCard™ No. 1 (choker-ring)	1.8 cm	25	100%	1044
GPSCard™ No. 2 (no choker-ring)	1.8 cm	17	100%	836
GPSCard™ No. 3 (choker-ring)	1.8 cm	31	100%	700

The residuals from Figures 5.19 to 5.24 (thinned out by a factor of ten) were computed using fixed ambiguities. As seen, no phase residuals were larger than 1.5 cm. Most of residuals were caused by the effect of multipath. Random variations at the millimetre level were caused by receiver measurement noises. It is not evident that antenna 2, without choker-rings has a larger multipath effect, because antenna 2 was at the highest point on the launch. Antennas 1 and 3 are at lower positions, but they were equipped with choker-rings, and these choker-rings balanced out the effect of the relatively unfavorable multipath conditions. The differences in the number of epochs required to resolve ambiguities could also be caused by small differences in the error patterns among the receivers.

Table 5.8 Repeated Ambiguity Resolution (Hatch's Least-squares Method)

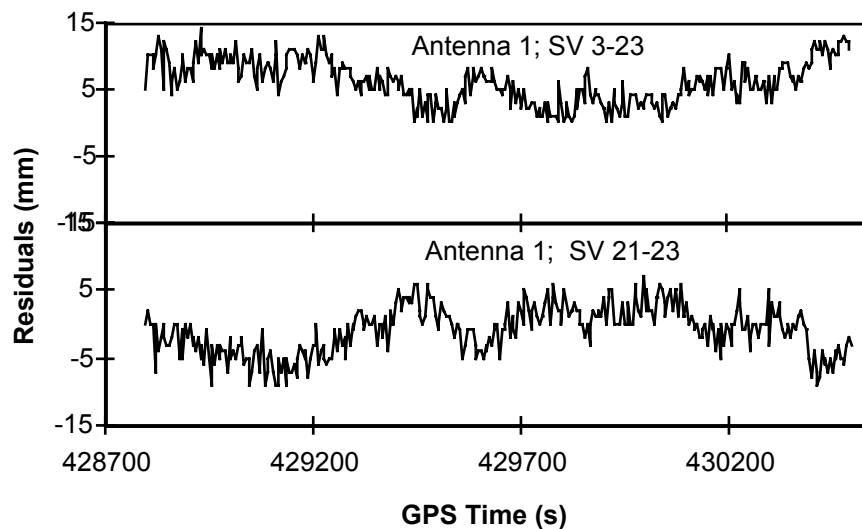| Receivers	$\sigma_{\Delta\nabla}$ (max)	Number of Trials	Success rate	Average Period Required
GPSCard™ No. 1 (choker-ring)	1.8 cm	9	100%	1032 s
GPSCard™ No. 2 (no choker-ring)	1.8 cm	7	100%	1825 s
GPSCard™ No. 3 (choker-ring)	1.8 cm	7	100%	1146 s

Figure 5.19   Double Difference Carrier Phase Residuals for Antenna 1 (Choker-ring Ground Planes at both the Reference and Launch).

Comparing with the performance on the land and air, the improvement under the water environment is limited The following factors contribute to this limit in improvement. (1) The multipath effect. Multipath under sea environment seems to have a longer period, and it cannot be filtered out by the method developed here. (2) Fewer satellites. In the previous case, nominally seven satellites were observed; while in this case, only six satellites were observed; (3) Other biases, such as, orbital errors, ionosphere, and troposphere. As compared with the previous two sections, there are larger residuals for

both carrier phase and code measurements because of a greater distance relative to the wavelength.
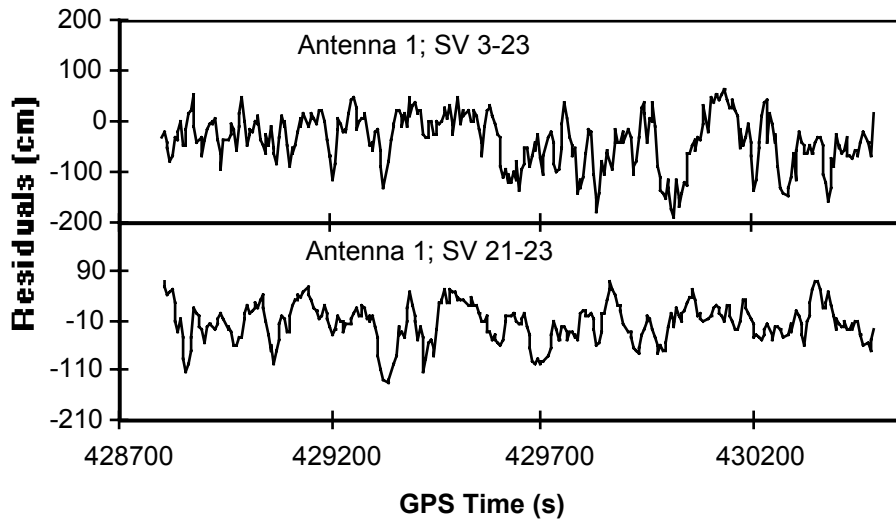
Figure 5.20 Double Difference C/A Code Residuals for Antenna 1 (Choker-ring Ground Planes at both the Reference and Launch).
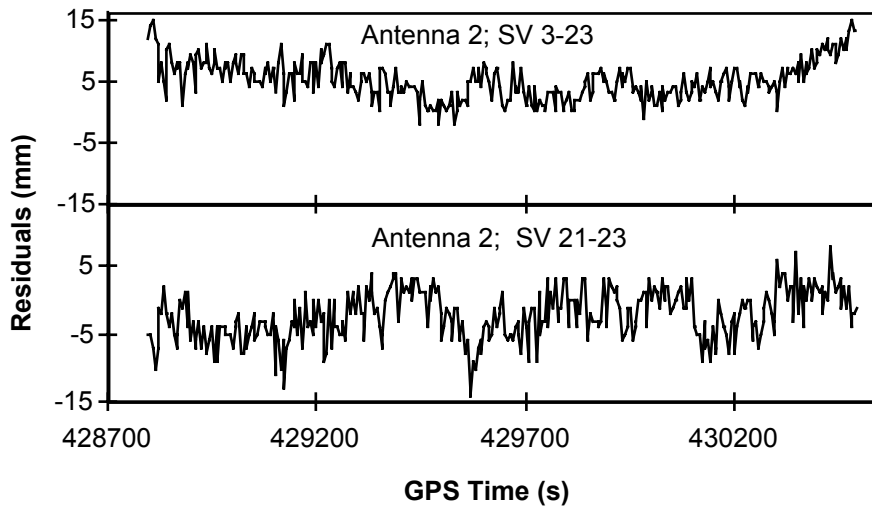

Figure 5.21 Double Difference Phase Residuals for Antenna 2 (No Choker-ring Ground Planes at Antenna 2).
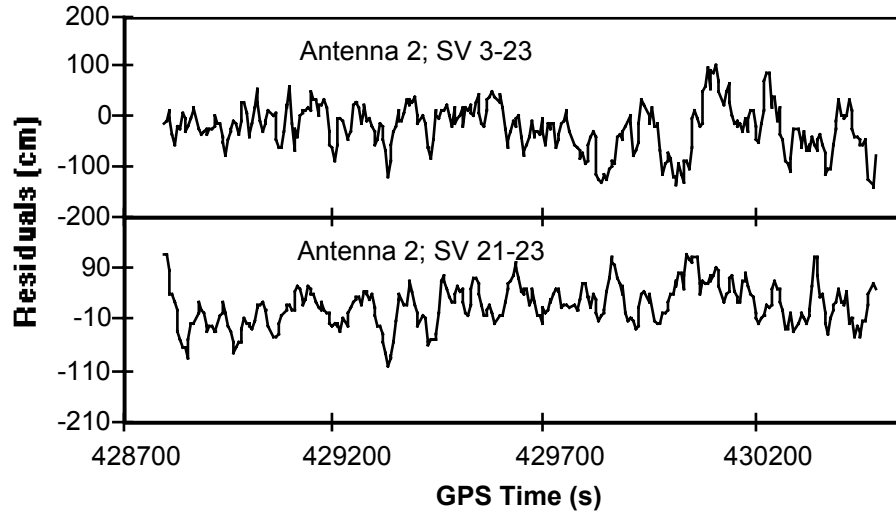
Figure 5.22   Double Difference C/A Code Residuals for Antenna 2 (No
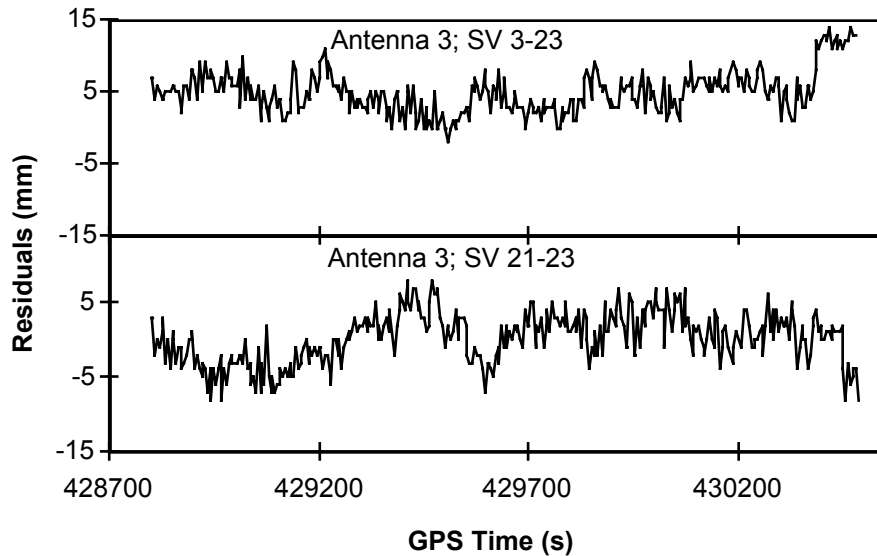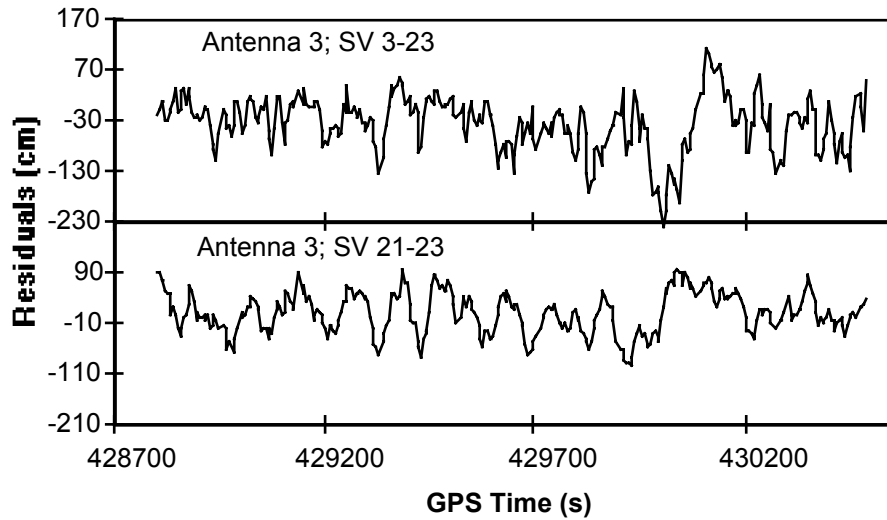Choker-ring Ground Planes at Antenna 2).

Figure 5.23   Double Difference Carrier Phase Residuals for Antenna 3 (Choker-ring
Ground Planes at both the Reference and Launch).

The independent distance measurements between the antennas on board the vessel were
also used to check ambiguity resolution. The differences between the computed distances
from the GPS and the direct measurements are shown in Figure 5.24. Apart from a small
constant of about 1 cm, there is no systematic deviation. The small constant values are
caused by differences in measurement centers between the two types of distances. The

89

distance measurements are referred to the base of the antennas, while the distances from GPS are between the phase centers. Other causes of the deviations were the multipath interference and the measurement errors in the distances.

Figure 5.24 Double Difference C/A Code Residuals for Antenna 3 (Choker-ring Ground Planes at both the Reference and Launch).
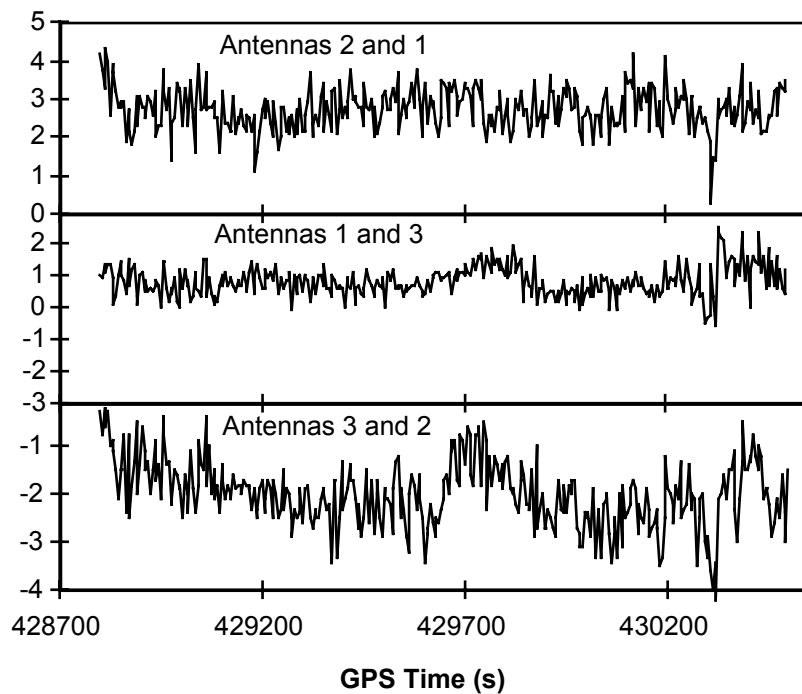

Figure 5.25 Differences Between Calculated and Measured Distances Using Fixed Ambiguity Solutions

5.5   Correlation Analysis of Land and Marine Tests

As discussed previously, the standard deviations of float ambiguities from least-squares may not reflect the actual uncertainties due to the correlation of the noise under some environmental effects. These correlations were partly responsible for the long period of observations required to resolve ambiguities in the marine test.   As a result of the correlation, the standard deviation from the least-squares was too optimistic.   To compensate this effect, a much larger expansion factor was used for the marine data analysis. That is, the expansion factor was three in the land test, while it was fifteen in the marine test. As a result, it took significantly more time to resolve ambiguities in the marine case than in the land case.

In this section, the numerical analysis of the correlations for the land and marine tests will be presented. The Fourier analysis of the correlation will also be shown.

Figure 5.26 shows the correlation of the double difference residuals for satellite pair 26 and 23 in the land test and Figures 5. 27 to 5.29 are the correlations for satellite pair 23 and 3 in the marine test. As seen, in the land case, the correlation quickly decreased and fluctuated between -0.4 and 0.4 after 25 seconds.   The effect of mixed positive and
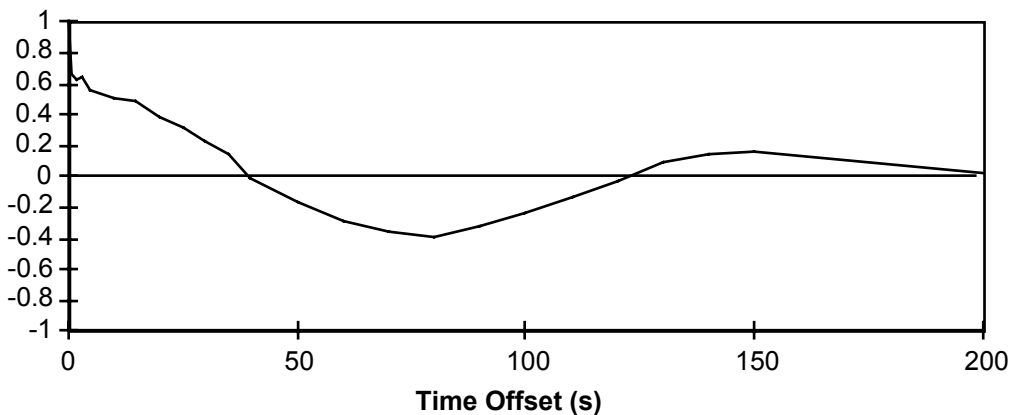

Figure 5.26   Correlation of the Phase Residuals for SV 26-23 for Land Test

negative correlations will mostly be canceled out over the time. However, in the marine case, the correlation decreased much more slowly. The effect of this correlation can not be canceled out over a short period.
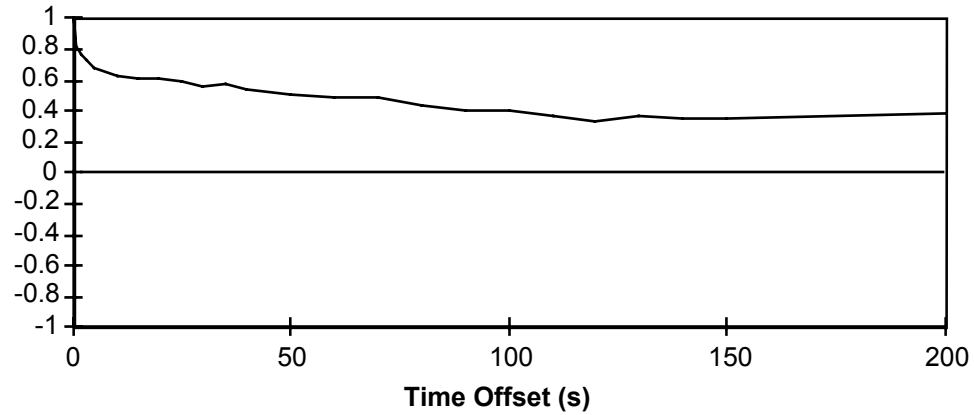

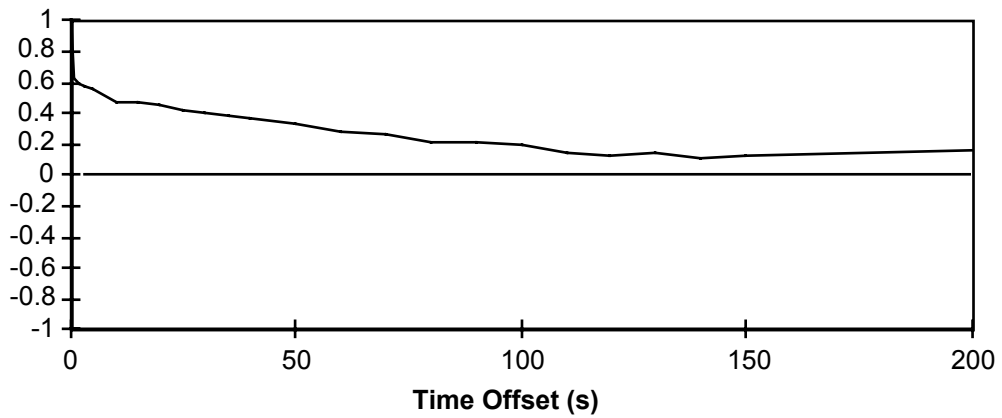Figure 5.27   Correlation of the Phase Residuals for SV 23-3 for Marine Test (Antenna 1)

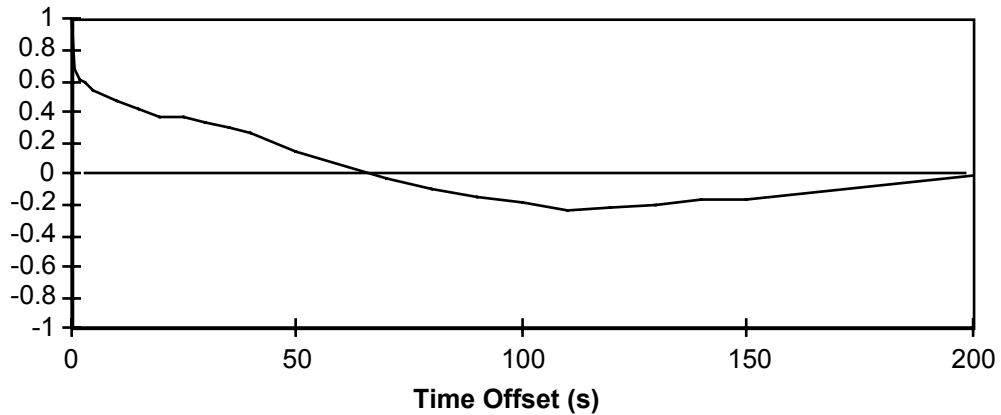Figure 5.28   Correlation of the Phase Residuals for SV 23-3 for Marine Test (Antenna 2)

Figure 5.29   Correlation of the Phase Residuals for SV 23-3 for Marine Test (Antenna 3)

The Fourier analysis of the phase residuals in the land test is shown in Figure 5.30 and the Fourier analysis for the marine test is shown in Figures 5.31 to 5.33. It should be noted that to show the Fourier spectrums clearly at the same scale for both the land and marine tests, the spectrums for f=0 (the constant component) and f=0.001 Hz are not shown for the marine tests. Instead, only the numerical values for the Fourier spectrums are shown in the Figures. In the marine case, there are large spectrums for frequencies between 0 and 0.001 Hz caused by the orbital errors. They cannot be averaged out by the least-squares method within short periods. In the land case, all the spectrums are smaller than 8 Hz.

It is these low frequency biases in the marine tests that caused the strong correlation demonstrated in Figures 5.27 to 5.29. As a result, the search ranges became larger and more observations were required to resolve ambiguities. Besides the low frequency multipath effect, these low frequencies were mainly caused by the orbital errors which could be in the period of hours due to the relatively larger separation between the rover and the monitor. The spectrums for three antennas in the marine show only small

93

differences. The differences in ambiguity resolution time were likely caused by the random nature of the noise.

Both land and marine tests show that there are spectrums below 1.7 Hz at all frequencies. These spectrums were caused by the white noise and can be easily represented in the standard deviation from the least-squares estimation. The spectrums at 0.16 to 0.17 Hz for the land test are short terms with periods between five and seven seconds) which can also be averaged out by the least-squares estimation. These spectrums were responsible for the fluctuated correlation in Figure 5.26. They are much smaller as compared with the spectrums at low frequencies in the marine test.
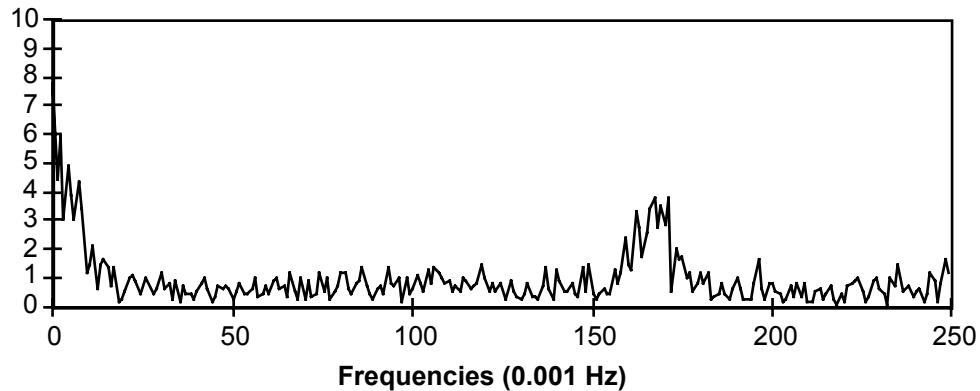

Figure 5.30   Fourier Spectrums of the Phase Residuals for SV 26-23 for Land Test
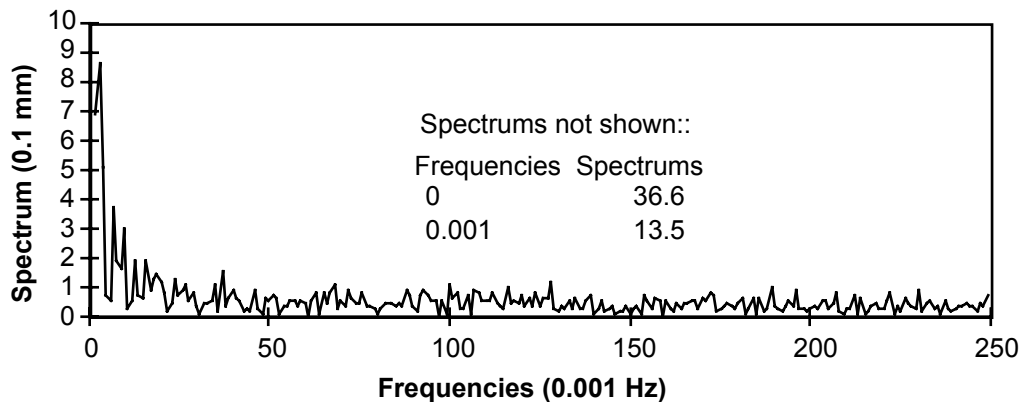

Figure 5.31   Fourier Spectrums of the Phase Residuals for SV 23-3 for Marine Test
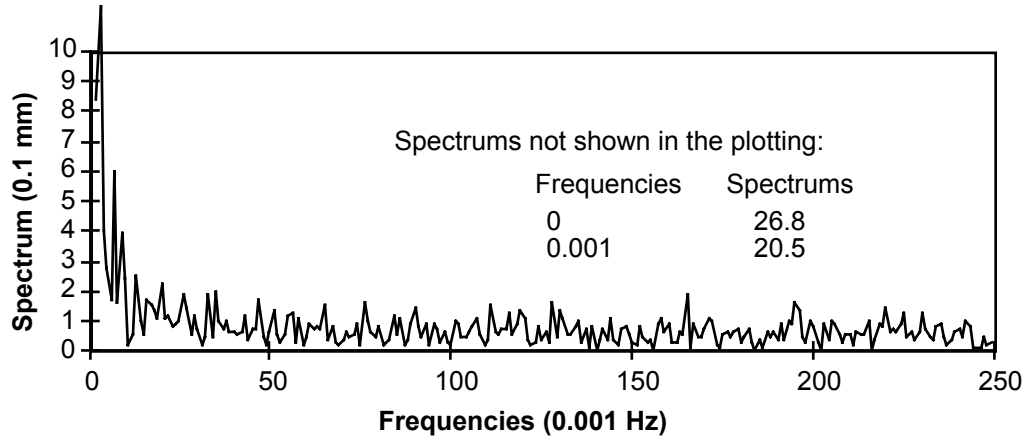(Antenna 1)

Figure 5.32   Fourier Spectrums of the Phase Residuals for SV 23-3 for Marine Test (Antenna 2)
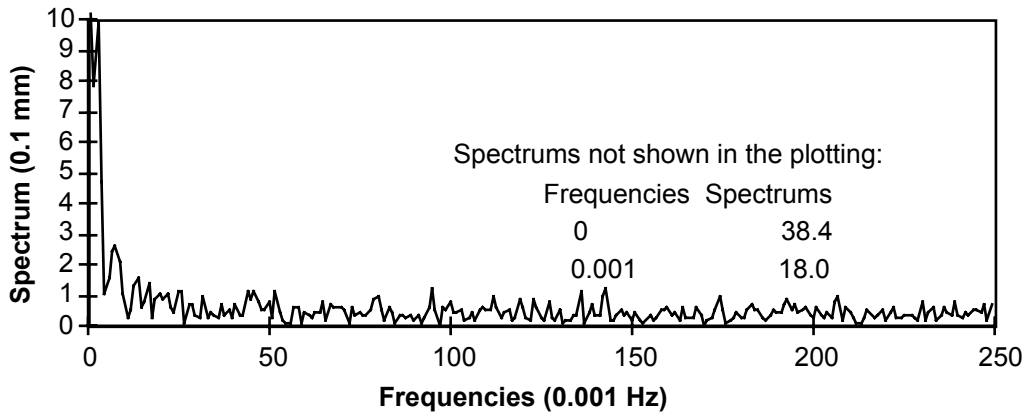

Figure 5.33   Fourier Spectrums of the Phase Residuals for SV 23-3 for Marine Test (Antenna 3)

5.6   Summary

The results presented in this chapter have shown that FASF reduces both computation and observation times required for ambiguity resolution OTF as compared with the least-squares search method.   Both land and airborne testings have shown significant improvement.   However, marine testing did not show much improvement, primarily because of the relatively stronger systematic errors.

CHAPTER 6   CONCLUSIONS AND RECOMMENDATIONS

A new concept for ambiguity resolution is developed. In the past, the search range of each ambiguity was determined separately, regardless of the assumed integers of the other ambiguities. The relationship between potentially fixed ambiguities was not fully considered. In FASF, the search ranges are determined recursively and are related to each other. To determine the uncertainty range of an ambiguity parameter, the effect of an assumed integer on others is fully taken into account. The geometry information is exploited more effectively by constraining ambiguities to the possible integer values in the range computation. Therefore, less observation time is required to resolve ambiguities.

A threshold to exit ambiguity search is used. Ambiguity resolution is not made until there is a higher chance of ambiguity fixing. Therefore, unnecessary search is avoided and the computation can be dramatically reduced. Currently, the maximum number of the potential solutions is used as the threshold. This is justified because the greater the number of potential ambiguity sets, the more likely the ambiguity resolution fails.

The computational improvement comes from two aspects; one is the recursive computation of the search range (RCSR); the other is the threshold to exit the search process as stated previously. RCSR makes the search range sequentially smaller as more ambiguities are treated as fixed in the ambiguity series.

Another characteristic is that all observations, from the initial to the current epoch, are taken into account by a least-squares filter (or Kalman Filter). Another advantage of the

least-squares filtering is the tolerance to large isolated errors. A method based on the epoch by epoch rejection of incorrect solutions, such as Hatch's least-squares method, will be very sensitive to any unexpected large error at any single epoch. That is, any error larger than the pre-specified threshold in any single epoch will risk rejecting the correct solution. While this method tends to smooth out random errors, it can not reduce any systematic biases. That is especially true if it combines with poor satellite geometry and small number of satellites.

Since the method is computationally fast, it can be applied in real time. For example, in rapid static surveying, ambiguities can be searched in real time. Therefore, GPS data can be collected until ambiguities are fixed. This will be useful in many kinematic environments where a high data collection rate is required. In attitude determination systems where ambiguity search is a serious burden because multi-baselines have to be searched, the fast ambiguity resolution method will have a clear computational advantage.

As a first attempt to implement FASF, variances from the least-squares estimation were used to compute the search ranges. However, other problems remain to be resolved, as will be discussed in the following.

The effect of the systematic biases on ambiguity resolution should be investigated further, including multipath, troposphere, ionosphere, and orbits. While in some cases, the method works well, in other cases, it does not. Mostly, it is caused by the computation of the standard deviation for the partially fixed ambiguities from least-squares estimation. The presumption is that the noise in the observation should be uncorrelated from epoch to epoch. However, in many cases, there is a strong correlation between the errors from different epochs. This correlation makes the estimated standard deviation too optimistic. It can be partially considered by proper selection of the expansion factor that converts the

97

standard deviation to the maximum possible error in the float estimation. However, the difficulty is that this expansion factor is highly dependent on correlation. The exact relationship between the correlation and the expansion factor and the relationship between the correlation and the environments should be studied. The relationship between the geometry, error pattern, success rate, and time required to resolve ambiguity also needs to be further investigated.

Most of the results are compared with Hatch's least-squares methods. Comparison with other methods should be carried out. Also more research should be carried out to test the effectiveness of the method under different environments.

The least-squares implementation of FASF will work fine where only a few epochs of observations are needed to resolve ambiguities since the distortion of standard deviation will be serious only when extensive observations are accumulated. For example, in the case of favorable geometry, the least-squares implementation of FASF should be good even if there is strong correlation in the observations from epoch to epoch. Also, for dual frequency receivers, where the wide lane ambiguities can be formed either explicitly or implicitly, FASF will also work properly because of the relatively small size of the search range in terms of the number of cycles.

For post mission processing, precise orbits can be used. That will be very useful for a longer baseline. A better tropospheric model should be used to take into account the height effect. An ionospheric model should be considered when the distance or height between the reference and remote is large.

Cycle slips also need to be properly handled. When the cycle slips can not be fixed exactly, new ambiguity parameters have to be introduced.

REFERENCES

Abidin, H.Z. (1991). "New Strategy for On-The-Fly Ambiguity Resolution." *Proc. of ION GPS-91*, The Institute of Navigation, Alexandria, VA, pp. 875-886.

Blewitt, G. (1989). "Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km." J. Geophysics. Res., Vol 94, No. B8, pp. 10.187-10.203.

Brock, K., R. Fuller, Hur-Diaz S., and J. Rodden (1994). "GPS attitude and orbit determination for JAWSAT." *Proc. of ION GPS-94*, The Institute of Navigation, Alexandria, VA, pp. 1251-1261.

Brown, R.A. (1992). "Instantaneous GPS Attitude Determination." *Proc. of PLANS'92, IEEE Position Location and Navigation Sym.*, Held at Monterey, California, 24-27 Mar., pp. 113-120.

Cannon, M.E., G. Lachapelle, H. Ayers, and K.P. Schwarz. (1990). "A Comparison of SEMIKIN and KINSRVY for Kinematic Applications." *Proc. of ION GPS-90*, The Institute of Navigation, Alexandria, VA, pp.80-84.

Cannon, M.E. (1991). "Airborne GPS/INS with an Application to Aerotriangulation." UCSE Report Number 20040, Department of Geomatics Engineering, The University of Calgary, Calgary, Canada.

Cannon, M.E. and G. Lachapelle (1992). "Analysis of a High Performance C/A Code GPS Receiver in Kinematic Mode." Navigation, Vol. 39, No. 3, The Institute of Navigation, Alexandria, VA, pp. 285-299.

Cannon, M.E. and M. Haverland (1993). "Experiences of GPS Attitude Determination within a Helicopter Pod." *Proc. of ION GPS-93*, The Institute of Navigation, Alexandria, VA, pp. 633-640.

Chen, D. (1991). A *Long Arc Approach to GPS Satellite Orbit Improvement.* M.S.E. Thesis, The Department of Surveying Engineering, The University of New Brunswick.

Chen, D.S., R.B. Langley (1990). "A Geometrical Analysis of the Effect of Satellite Orbit Error on GPS Relative Positioning." *Proc. of GPS'90*, Canadian Institute of Geomatics, Ottawa, pp. 757-771.

Chen, D. (1993). "FAST Ambiguity Search Filter (FASF): A Novel Concept for GPS Ambiguity Resolution." *Proc. of ION GPS-93*, The Institute of Navigation, Alexandria, VA, pp. 781-787.

Cohen, C., B. Pervan, B. Parkinson (1992). "Estimation of Absolute Ionospheric Delay Exclusively through Single-Frequency GPS Measurements." *Proc. of ION GPS-92*, The Institute of Navigation, Alexandria, VA, pp. 325-330.

Cohen, C.E. and B.W. Parkinson (1991). "Expanding the Performance Envelope of GPS Based Attitude." *Proc. of ION GPS-91,* The Institute of Navigation, Alexandria, pp. 1001-1011.

Counselman, C. C., and S. A. Gourevitch (1981). "Miniature Interferometer Terminals for Earth Surveying: Ambiguity and Multipath with Global Positioning System." IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-19, No. 4, pp. 244-252.

Davis, J.M. (1993). "RNP Tunnel Concept for Precision Approach with GNSS Application." *Proc. of 49 ION National Technical Meeting*, The Institute of Navigation, Alexandria, VA, pp. 135-154.

Delikaraoglou, D., H. Dragert, J. Kouba, K. Lockhead, J. Popelar (1990). "The Development of a Canadian GPS Active Control System: Status of the Current Array." *Proc. of GPS'90*, Canadian Institute of Geomatics, pp. 190-202.

Dieter, U. (1975) "How to Calculate Shortest Vectors in Lattice." <u>Mathematics of Computations,</u> Vol. 29, No. 131, pp. 827-833.

Dong, D., and Y. Bock (1989). "GPS Network Analysis with Phase Ambiguity Resolution Applied to Crustal Deformation Studies in California." <u>J. Geophysics. Res.</u>, 94, (B4), 3949-3966.

Euler, H.-J., C.C. Goad (1991). "On Optimal Filtering of Dual-Frequency Observations without using Orbit Information." Bulletin Géodésique, Springer Verlag, New York, Vol. 65, *pp*. 130-143.

Frei, E., G. Beutler (1990). "Rapid Static Positioning based on the Fast Ambiguity Resolution Approach: The Alternative to Kinematic Positioning." *Proc. of the Second International Symposium on Precise Positioning with the Global Positioning System,* "GPS 90", Canadian Institute of Geomatics, Ottawa.

FRNP (1990) *1990 Federal Radio Navigation Plan.* Rep. No. DOD-4650.4 and DOT-VNTSC-RSPA-90-3.

Georgiadou, Y., and A. Kleusberg (1988). "On Carrier Signal Multipath Effects in Relative GPS Positioning." <u>Manuscripta Geodaetica</u>, Springer Verlag, Vol. 13, pp. 172-179.

Hatch, R. (1991). "Instantaneous Ambiguity Resolution." Kinematic Systems in Geodesy, Surveying and Remote Sensing, International Association of Geodesy Symposia 107, Springer Verlag, New York, pp. 299-308.

Hintz, R.J. and M.Z. Zhao (1989). "Considerations in the Implementation of Airotriangulation with GPS derived Exposure Station Positions." Photogrammetric Engineering and Remote Sensing. Vol. 55, No. 12, December, pp. 1731-1735.

Hofmann-Wellenhof, B., and B.W. Remondi (1988). "The Antenna Exchange: One Aspect of High-precision GPS Kinematic Surveying." Presented at the International GPS Workshop, GPS Techniques Applied to Geodesy and Surveying, Darmstadt, German, April 10-13.

Hopfield, H.S. (1971). "Tropospheric Effect on Electromagnetically Measured Range: Prediction from Surface Weather Data." Radio Sci., Vol. 6, pp. 357-367.

Hwang, P.Y.C. (1990). "Kinematic GPS: Resolving Integer Ambiguities On-The-Fly." *Proc. of Plans'90, IEEE Position Location and Navigation Sym.,* Las Vegas, Nev., pp. 579-586.

Janes, H.W., R.B. Langley, S.P. Newby (1990). "Analysis of Tropospheric Delay Prediction Models: Comparisons with Ray-Tracking and Implications for GPS Relative Positioning (a Summary)." *Proc. of GPS'90*, Canadian Institute of Geomatics, Ottawa, Canada, pp. 444-450.

Kalman, R.E. (1960). "A New Approach to Linear Filtering and Prediction." *Journal of Basic Engineering, ASME*, 82D.

Kleusberg, A. (1990). "A Review of Kinematic and Static GPS Surveying Procedures." *Proc. of GPS'90*, Canadian Institute of Geomatics, Ottawa, pp. 1102-1113.

Krakiwsky, E. J. (1990). "The Method of Least-squares: a Synthesis of Advances." Department of Surveying Engineering, The University of Calgary.

Lachapelle. G., M.E Cannon, and G. Lu (1992[a]). "High Precision GPS Navigation with Emphasis on Carrier Phase Ambiguity Resolution." Marine Geodesy, Vol 15, 4 , pp. 253-269.

Lachapelle. G., M.E Cannon, and G. Lu (1992[b]). "Ambiguity Resolution On-The-Fly - A Comparison of P Code and High Performance C/A Code Receiver Technologies." *Proc. of ION GPS-92*, The Institute of Navigation, Alexandria, VA , pp. 1025-1032.

Lachapelle, G., M.E. Cannon, and G. Lu (1993[a]). "A Comparison of P code and High Performance C/A GPS Receivers for On-The-Fly Ambiguity Resolution." Bulletin Géodésique, Springer Verlag, New York , Vol. 67, No. 3, pp. 185-192.

Lachapelle, G., C. Liu, G. Lu, B. Townsend, M.E Cannon, and R. Hare. (1993[b]). "Precise Marine DGPS Positioning Using P code and High Performance C/A Code Technologies." Geomatica, Canadian Institute of Geomatics, Ottawa, Vol 47, No. 2, pp. 117-128.

Lachapelle, G., H. Sun, M.E. Cannon, and G. Lu (1994[a]). "Precise Aircraft-to-Aircraft Positioning Using a Multiple Receiver Configuration." Canadian Aeronautics and Space Journal, Canadian Aeronautics and Space Institute, Vol. 40, No. 2, pp. 74-78.

Lachapelle, G., M.E. Cannon, H. Gehue, T. Goddard, and D. Penney (1994[b]). "GPS System Integration and Field Approaches in Precision Farming." Navigation, Vol. 41, No. 3, The Institute of Navigation, Alexandria, pp. 271-278.

Lachapelle, G., C. Liu, G. Lu, Q. Weigen, and R. Hare (1994[c]). "Water-Borne Leveling with GPS." Marine Geodesy, Vol. 17, No. 4, pp. 271-278.

Landau, H., and U. Vollath (1994) "Differential GPS - New Developments on High Precision Positioning." *Proc. of 3'rd International Conf. on Differential Satellite Navigation Systems (DSNS)*, The Royal Institute of Navigation, London, Paper No. 23.

Landau, H., H.-J. Euler (1992). "On-the-Fly Ambiguity Resolution for Precise Differential Positioning." *Proc. of ION-92*, The Institute of Navigation, Alexandria, VA., pp. 607-613.

Langley, R.B., G. Beutler, D. Delikaraoglou, B.G. Nickerson, R. Santerre, P. Vanicek, and D.E. Wells (1984). "Studies in the Application of the Global Positioning System to Differential Positioning." Department of Surveying Engineering Technical Report No. 108, University of New Brunswick, Fredericton, N.B.

Loomis, P. (1989). "A Kinematic GPS Double-Differencing Algorithm." *Proc. of Fifth International Geodetic Symposium on Satellite Positioning*, DMA, U.S., DoD, NGS, NOAA, Vol. 2, Held at Las Cruces, N. Mex., pp. 611-620.

Lu, G. (1993). Personal communication. Department of Geomatics Engineering, The University of Calgary.

Mader, G.L. (1992). "Rapid static and kinematic Global Positioning System Solutions Using the Ambiguity Function Technique." J. of Geophysical Res., Vol. 97, No. B3, pp. 3271-3283.

Ober, P.B. (1993) "Split-second Ambiguity Resolution for GPS Using the Basis-Reduction Method." Report A-565, Telecommunications and Traffic-control Systems Group, Faculty of Electrical Engineering, Delft U. of Tech.

Pohst, M. and H. Zassenhaus (1989) Algorithmic Algebraic Number Theory. Cambridge U. Press.

Qiu, W. (1993). "An Analysis of Some Critical Error Sources in Static GPS Surveying." Technical Report 20054, Ms Thesis, Department of Geomatics Engineering, The University of Calgary.

Qiu, W., G. Lachapelle, and M.E. Cannon (1995). "Ionospheric Effect Modeling for Single Frequency GPS Users." <u>Manuscripta Geodetica</u>, Springer Verlag, New York, 20, pp. 96-109.

Remondi, B.W. (1986). "Performing Centimetre Accuracy Relative Surveys in Seconds Using GPS Carrier Phase: Initial Results." <u>Navigation</u>, Journal of The Institute of Navigation, Vol. 32, No. 4, 1985, pp. 386-400.

Schwarz, K.P., M.A. Chapman, M.E. Cannon, P. Gong, and D. Cosandier (1994). "A Precise Positioning/Attitude System in Support of Airborne Remote Sensing." The Symposium of ISPRS Commission II, pp. 191-202.

Seeber, G. (1993). <u>Satellite Geodesy: Foundations, Methods, and Applications</u>. Walter de Gruyter, Berlin, New York.

Sun, H. (1993). Personal communication. Department of Geomatics Engineering, The University of Calgary.

Talbot, N.C. (1991). "Sequential Phase Ambiguity Resolution for Real-Time Static Differential Positioning." <u>Manuscripta Geodetica</u>, Springer Verlag, 16, 274-282.

Teunissen, P.J.G. (1994) "A New Method for Fast Carrier Phase Ambiguity Estimation." *Proc. of PLANS94,* IEEE, New York, pp. 562-573.

Trimble Navigation (1989). *GPS A Guide to the Next Utility.* Sunnyvale, California.

Van Dierendonck, A.J., P. Fenton, and T. Ford (1992). "Theory and Performance of Narrow Correlator Spacing in A GPS Receiver." <u>Navigation</u>, Journal of The Institute of Navigation, Alexandria, VA, Vol. 39, No. 3, pp. 265-283.

van Nee, R. D. J. and J. Siereveld (1993). "The Multipath Estimating Delay Lock Loop - Approaching Theoretical Accuracy Limits." *Proc. of ION GPS-93*, The Institute of Navigation, Alexandria, VA, pp. 921-922.

Vaníčěk, P., G. Beutler, A. Kleusberg, R. B. Langley, R. Santerre, and D. E. Wells (1985). "DIPOP: DIfferential POsitioning Program Package for the Global Positioning System." Technical Report No. 115, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B.

Wells, D.E., N. Beck, D. Delikaraoglou, A. Kleusberg, E.J. Krakiwsky, G. Lachapelle, R.B. Langley, M. Nakiboglu, K.P. Schwarz, J.M. Tranquilla, and P. Vanicek (1987). *Guide to GPS Positioning*. 2'nd printing with corrections, Canadian GPS Associates, Fredericton, N.B.

Wübbena, G. (1989). "The GPS Adjustment Software Package GEONAP, Concepts and Models." *Proc. of Fifth International Geodetic Symposium on Satellite Positioning*, DMA, U.S., DoD, NGS, NOAA, held at Las Cruzes, N. Mex., Vol. 1, pp. 452-461.

APPENDIX   PROOF OF THE PARAMETER
REMOVAL THEOREM

The thereom of parameter removal (Chapter 3) can be proven using the parametric adjustment theorem as follows:

<u>Proof of the Normal Equations</u>

The observation equations are described by

$$l + r = f(x, y) ,$$ (1)

with the weight of the observation as P_l. If the constant parameter y is treated as unknown, the linearized observation equations can be described as:

$$r = [A_x, A_y]\begin{bmatrix} \delta x \\ \delta y \end{bmatrix} + w ,$$ (2)

where,

$$A_x = \frac{\partial f}{\partial x}\bigg|_{x=x_o} ,$$

$$A_y = \frac{\partial f}{\partial y}\bigg|_{y=y_o} ,$$

$$w = f(x_o, y_o) ,$$

where, x_o and y_o are the approximated values of x and y,

and the normal equations are:

$$\begin{bmatrix} P_{xx} & P_{xy} \\ P_{yx} & P_{yy} \end{bmatrix}\begin{bmatrix} \hat{\delta x} \\ \hat{\delta y} \end{bmatrix} = \begin{bmatrix} u_x \\ u_y \end{bmatrix} ,$$ (3a)

where,

$$\begin{bmatrix} P_{xx} & P_{xy} \\ P_{yx} & P_{yy} \end{bmatrix} = \begin{bmatrix} A_x^T P_l A_x & A_x^T P_l A_y \\ A_y^T P_l A_x & A_y^T P_l A_y \end{bmatrix},$$

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} A_x^T P_l w \\ A_y^T P_l w \end{bmatrix}, \text{ and} \tag{3b}$$

$$\begin{bmatrix} \hat{\delta x} \\ \hat{\delta y} \end{bmatrix} \text{ is the least-squared estimation of } \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}.$$

If y becomes the known constants, then the linearized observation equations become

$$r|_y = A_x \, \delta x + w|_y, \tag{4}$$

where,

$$w|_y = f(x_o, y_o) + A_y (y_C - y_o)$$
$$= w + A_y (y_C - y_o). \tag{4b}$$

Here, equation (2) and 4a are assumed to be linearized at the same values. Then, the corresponding normal equations are in the form of

$$\tilde{P}_{xx} \, \hat{\delta x}|_y = u_x|_y, \tag{5}$$

where,

$$\tilde{P}_{xx} = A_x^T P_l A_x, \text{ and}$$
$$u_x|_y = A_x^T P_l w|_y. \tag{6}$$

Comparing with equation (3b), we know that $\tilde{P}_{xx} = P_{xx}$. Inserting equations (4b) into (6), the following can be obtained:

$$u_x|_y = -A_x^T P_l [w + A_y (y_C - y_o)],$$
$$= u_x - A_x^T P_l A_y (y_C - y_o),$$
$$= u_x - P_{xy} (y_C - y_o). \tag{7}$$

Therefore, equations (3.30) and (3.32) are true.

Proof of Ω

$$\hat{r}\,|_y = w|_y + A_x\,\delta\hat{x}\,|_y$$

$$= w + A_y\,(y_C - y_o) + A_x\,\delta\hat{x}\,|_y$$

$$= w + A_y\,(y_C - \hat{y} - y_o + \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x} + \delta\hat{x})$$

$$= w + A_y\,(y_C - \hat{y} + \delta\hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x} + \delta\hat{x})$$

$$= (w + A_x\,\delta\hat{x} + A_y\,\delta\hat{y}) + [A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]$$

$$= \hat{r} + [A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]$$

$$\Omega|_y = \hat{r}\,|_y{}^T\,P_l\,\hat{r}\,|_y$$

$$= \{\hat{r} + [A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]\}^T\,P_l\,\{\hat{r} + [A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]\}$$

$$= \hat{r}\,^T\,P_l\,\hat{r} + 2[A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]^T\,P_l\,\hat{r}$$

$$+ [A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]^T\,P_l\,[A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]. \qquad (8)$$

In the above formula,

$$[A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})\,]^T\,P_l\,\hat{r}$$

$$= [(\delta\hat{x}\,|_y - \delta\hat{x})^T\,,\,(y_C - \hat{y})^T]\,[A_x\,A_y]^T\,P_l\,\hat{r}$$

Inserting $\hat{r} = w + A_x\,\delta\hat{x} + A_y\,\delta\hat{y}$ into the above, we obtain:

$$[A_x\,A_y]^T\,P_l\,\hat{r}$$

$$= [A_x\,A_y]^T\,P_l\,w + [A_x\,A_y]^T\,P_l\,[A_x\ \ A_y]\begin{bmatrix} \delta\hat{x} \\ \delta\hat{y} \end{bmatrix}$$

$$= -\begin{bmatrix} u_x \\ u_y \end{bmatrix} + \begin{bmatrix} P_{xx} & P_{xy} \\ P_{yx} & P_{yy} \end{bmatrix}\begin{bmatrix} \delta\hat{x} \\ \delta\hat{y} \end{bmatrix} = 0. \qquad (9)$$

Therefore,

$$\Omega|_y = \hat{r}\,^T\,P_l\,\hat{r}$$

$$+ [A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]^T\,P_l\,[A_y\,(y_C - \hat{y}) + A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})]$$

$$= \hat{r}\,^T\,P_l\,\hat{r} + (y_C - \hat{y})^T\,A_y{}^T\,P_l\,A_y\,(y_C - \hat{y})$$

$$+ (y_C - \hat{y})^T\,A_y{}^T\,P_l\,A_x\,(\delta\hat{x}\,|_y - \delta\hat{x}) + (\delta\hat{x}\,|_y - \delta\hat{x})^T\,A_x{}^T\,P_l\,A_y\,(y_C - \hat{y})$$

$$+ (\delta\hat{x}\,|_y - \delta\hat{x})^T\,A_x{}^T\,P_l\,A_x\,(\delta\hat{x}\,|_y - \delta\hat{x})$$

109

$$= \hat{r}^T P_l \hat{r} + (y_C - \hat{y})^T P_{yy} (y_C - \hat{y}) + (y_C - \hat{y})^T P_{yx} (\hat{\delta x}|_y - \hat{\delta x})$$
$$+ (\hat{x}|_y - \hat{\delta x})^T P_{xy} (y_C - \hat{y}) + (\hat{x}|_y - \hat{\delta x})^T P_{xx} (\hat{x}|_y - \hat{\delta x}), \tag{10}$$

However, from equations (3.30) and (3.32), $\hat{\delta x}|_y = P_{xx}^{-1} [u_{xx} - P_{xy} (y_C - y_o)]$; and from equation (3.29), $\hat{\delta x} = P_{xx}^{-1} [u_x - P_{xy} \hat{\delta y}]$. Therefore,

$$\hat{\delta x}|_y - \hat{\delta x} = -P_{xx}^{-1} P_{xy} (y_C - y_o - \hat{\delta y})$$
$$= -P_{xx}^{-1} P_{xy} (y_C - \hat{y}). \tag{11}$$

Inserting the above equation into equation (10), the following can be obtained:

$$\Omega|y = \hat{r}^T P_l \hat{r} + (y_C - \hat{y})^T P_{yy} (y_C - \hat{y})$$
$$- (y_C - \hat{y})^T P_{yx} P_{xx}^{-1} P_{xy} (y_C - \hat{y})$$
$$= \hat{r}^T P_l \hat{r} + (y_C - \hat{y})^T [P_{yy} - P_{yx} P_{xx}^{-1} P_{xy}] (y_C - \hat{y})$$
$$= \hat{r}^T P_l \hat{r} + (y_C - \hat{y})^T \tilde{P}_{yy} (y_C - \hat{y}) \tag{12}$$

Therefore the theorem is true.