INFORMATION TO USERS

NAVIGATION INTEGRITY FOR AIRCRAFT PRECISION LANDING USING THE GLOBAL POSITIONING SYSTEM

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

Boris Steven Pervan

March 1996

UMI Number: 9620529

Copyright 1996 by
Pervan, Boris Steven

UMI
300 North Zeeb Road
Ann Arbor, MI 48103
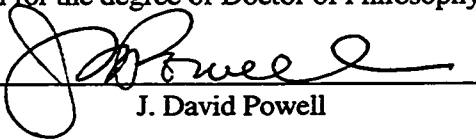
I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Bradford W. Parkinson
(Principal Adviser)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

J. David Powell

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Per Enge

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.
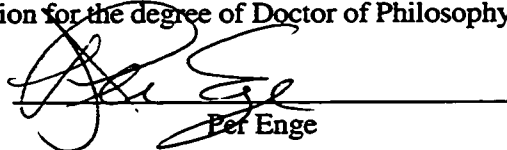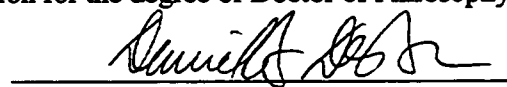
Daniel B. DeBra

Approved for the University Committee on Graduate Studies:

Abstract

The great potential benefit offered to commercial aviation by the Global Positioning System (GPS) lies in the possibility of inexpensive, seamless navigation from takeoff to touchdown. While this goal is highly motivating, significant technical challenges have existed, the most difficult of which have been associated with navigation during zero-visibility (Category III) precision landing. The severe requirements for accuracy, integrity, continuity, and availability have demanded a new level of GPS navigation system performance. For example, the integrity requirement of 'one undetected navigation failure in a billion approaches' has often been perceived as unattainable using GPS. In response, the central focus of this research has been to establish the viability of high-integrity satellite-based navigation for the precision landing of aircraft.

This dissertation demonstrates that highly precise GPS carrier phase measurements from spacecraft and ground-based pseudolites can provide the basis for high navigation integrity. It is shown that the considerable accuracy margin offered by carrier phase provides leverage for autonomous integrity monitoring aboard the aircraft in the sense that extremely tight fault detection thresholds may be set without incurring high false alarm rates. Furthermore, when placed under the approach path, pseudolites provide the means for real-time cycle ambiguity resolution and ensure the availability of redundant measurements for autonomous integrity monitoring.

Prototype algorithms for airborne kinematic carrier phase processing, including high-speed algorithms for the first high-integrity real-time cycle ambiguity resolution,

were developed, implemented and tested. Algorithm performance was verified through an extensive battery of flight tests culminating in 110 successful automatic landings of a United Airlines Boeing 737-300. In addition, the framework of Receiver Autonomous Integrity Monitoring (RAIM) was adapted for application to both cycle ambiguity resolution and kinematic positioning and for the detection of the wide range of navigation system failure scenarios. Navigation integrity and its parametric interrelationship with accuracy, continuity, and availability were quantitatively assessed through analysis, simulation, and flight test.

Acknowledgments

I would like to thank my advisor, Professor Bradford W. Parkinson, for giving me the opportunity to pursue this research. His clear guidance and comprehensive knowledge not only made this thesis possible, but also gave me the confidence and discipline to conduct engineering research in general. I would also like to thank my defense and reading committees, including Professors Daniel B. DeBra and Gene F. Franklin. Special thanks go to Professor Per Enge for his valuable suggestions throughout this research, and to Professor J. David Powell for his continual support since my arrival at Stanford and for providing the aircraft (and pilot!) for many of the flight tests described in this thesis.

To Dr. Clark E. Cohen, who provided me the opportunity to work on the GPS precision landing effort at Stanford, I owe a great debt. Without the foundation provided by his own personal accomplishments, this research would not have been possible. His friendship and daily guidance have been invaluable.

I would like to thank all of the GPS and Gravity Probe B graduate students and research associates for their encouragement and assistance throughout my research. I would especially like to express my gratitude to my fellow Ph.D. candidates David G. Lawrence and H. Stewart Cobb for both their friendship and valuable insights. Without their knowledge and dedication, Stanford's Integrity Beacon Landing System (IBLS) would not exist.

Thanks are due to Trimble Navigation for furnishing GPS receiver hardware used in this research and to the Boeing Commercial Airplane Company for providing me with

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

During the last fifteen years, the Global Positioning System (GPS) has revolutionized the field of navigation. The prospective application of GPS to commercial aviation has been especially alluring in recent years in that seamless satellite-based navigation from takeoff through touchdown may be possible at low cost. While this has been an inspiring goal, serious technical challenges have existed, the most persistent of which have been associated with navigation during zero-visibility (Category III) precision landing, where the stringent requirements for accuracy, integrity, continuity, and availability have demanded a new level of GPS navigation system performance. Specifically, the integrity requirement of 'one undetected navigation failure in a billion approaches' has often been perceived as unattainable using GPS. In response, the central focus of this research has been to establish the viability of high-integrity satellite-based navigation for the precision landing of aircraft.

1.1 Background

Navigation for civil aircraft precision approach and landing has a five-decade history in the form of the Instrument Landing System (ILS). A brief description of ILS will provide the context for a summary of the navigation requirements for precision

approach and landing. These requirements, or substantial equivalents, must ultimately be met by any proposed GPS-based landing system.

1.1.1 The Instrument Landing System

In 1946 the Instrument Landing System (ILS) was chosen by the International Civil Aeronautics Organization (ICAO) as the international all-weather navigation aid for precision approach [Kayton, b]. During the course of the next fifty years, a large number of runways have been equipped with ILS, and a number of these sites have been upgraded to the point of providing the capability for automatic landing of suitably equipped aircraft. Since its inception, ILS has had a flawless safety record, in that no fatalities have ever been directly attributed to ILS failure.

A typical ILS installation is comprised of a VHF localizer (horizontal) beam transmitter and a UHF glideslope (vertical) beam transmitter and three marker beacons (Figure 1.1) [ICAO]. In case of failure, ground-based monitor systems and "hot-spares" are also present for both the glideslope and localizer transmitters. During final approach, the localizer signal—detected by a VHF receiver onboard the aircraft—is used to determine the aircraft's displacement from the runway centerline. Likewise, the glideslope signal—as measured by a UHF receiver onboard the aircraft—is used to measure the aircraft's deviation from the nominal glidepath (which is usually 3 deg). The three 75 MHz marker beacons, placed under the approach path, provide discrete distance checks to the aircraft during an approach [Kayton, b].

Despite its consistent performance and excellent safety record through the years, there are a number of difficulties associated with ILS which have ultimately motivated the transition to a new navigation system for precision approach and landing—for example:
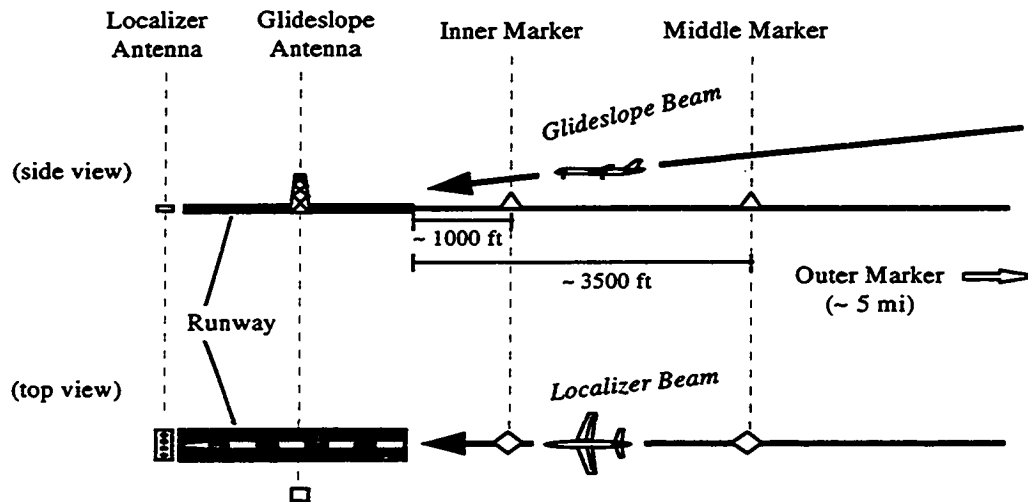
Figure 1.1: The Instrument Landing System (ILS)

1. ILS provides limited capability for curved approaches or parallel instrument approaches [Kayton, *b*].

2. ILS installations are sensitive to local terrain and nearby buildings, and the associated multipath is difficult to eliminate [Skillicorn, Kayton, *a*].

3. Many ILS installations are old, expensive to replace, and spare parts are hard to find [Skillicorn].

4. ILS frequency congestion exists in major urban areas [Skillicorn] and the loss of ILS glideslope frequency allocations in Europe to FM radio is imminent.

1.1.2 Navigation Requirements for Precision Landing

Traditionally, the overall quality and utility of a given navigation and guidance system for precision approach and landing (both ground facilities and aircraft hardware)

has been rated according to a three-tiered structure based on minimum achievable altitude [AC 120-28C]:

Category I. If the horizontal visibility on the runway, known as Runway Visual Range (RVR) is greater than 2400 ft, a *Category I* navigation and guidance system may deliver an aircraft down to a decision height (DH) of 200 ft. If at the DH the pilot is unable to see the runway, a missed approach must be executed.

Category II. If the RVR is greater than 1200 ft, a *Category II* navigation and guidance system may deliver an aircraft down to a DH of 100 ft. Again, if at the DH the pilot is unable to see the runway, a missed approach must be executed.

Category III. While *Category III* ILS ground facilities are generally designed for automatic landing, some variability is allowed, depending on the quality of the specific ground installation and the degree of fault tolerance (via redundant avionics) of the onboard guidance system. Thus, three sub-classes of Category III systems are possible:

- *Cat IIIa.* Ranges from DH > 50 ft and RVR > 700 ft to automatic landing.

- *Cat IIIb.* Automatic landing and rollout.

- *Cat IIIc.* Automatic landing, rollout, and taxi.

Four fundamental parameters provide the basis for allocation of specific requirements for Category I, II, and III navigation systems [Davis, ORD]:

Accuracy. Accuracy is the measure of the navigation output deviation from truth under fault-free conditions—often specified in terms of 95% performance.

Integrity. Integrity is the ability of a system to provide timely warnings to users when the system should not be used for navigation. *Integrity risk* is the

probability of an undetected navigation system error or failure that results in hazardously misleading information onboard the aircraft.

Continuity. Continuity is the likelihood that the navigation signal-in-space supports accuracy and integrity requirements for the duration of intended operation. *Continuity risk* is the probability of a detected but unscheduled navigation function interruption after an approach has been initiated.

Availability. Availability is the fraction of time the navigation function is usable (as determined by its compliance with the accuracy, integrity, and continuity requirements) before the approach is initiated.

The ICAO specifications for vertical accuracy (95% probability) of the ILS signal-in-space plus airborne glideslope receiver and the integrity and continuity specifications for ILS signal-in-space only are given in Table 1.1 [ICAO, Girts]. For a Category III ILS, the vertical accuracy required is roughly two feet at 50 ft altitude. This altitude is generally achieved at the runway threshold. At this point, vertical navigation is obtained primarily

	Vertical Accuracy at DH – 95 % (ft)	Integrity Risk (per approach)	Continuity Risk (per 15 sec)
Category I	14.5	–	–
Category II	6.1	10^{-7}	4×10^{-6}
Category III	2.1*	0.5×10^{-9}	2×10^{-6}

*50 ft DH

Table 1.1: ILS Signal-in-Space Specifications

from a radar altimeter onboard the aircraft. It is also immediately clear from Table 1.1 that for Category II/III navigation, ILS must operate with extremely high integrity and continuity. The *vertical protection limit*–the maximum safe (mean) glideslope deviation– for ILS integrity monitoring is roughly equivalent to 1.1 m at the 50 ft DH. In the event of an anomaly which causes this protection limit to be exceeded, the maximum time allowed to alarm the aircraft of the condition is 2 sec [ICAO]. In addition, while the actual ILS navigation availability at a single-ILS airport is 99.15%, the desired level of availability for a new navigation aid is at least 99.9% [Fernow].

The ILS specifications, of course, are also representative of the required performance for a given GPS-based navigation system that is intended to replace ILS. At the time of this writing, however, debate exists as to the degree of applicability of ILS requirements to GPS-based approach and landing systems. Indeed, a number of alternative requirements definitions have been proposed. The two most prominent examples are the Required Navigation Performance (RNP) concept [Kelly] and, more recently, the FAA Local Area Augmentation System (LAAS) operational requirements [ORD]. Although the allocation of requirements differs somewhat in philosophy among the various sources, Table 1.1 (along with a navigation availability guideline of 99.9%) serves to provide a quantitative guideline for Category III precision landing navigation requirements.

1.1.3 The Global Positioning System

The Global Positioning System (GPS), developed and implemented by the Department of Defense over the course of the past twenty years, is a passive satellite-based ranging system (see Figure 1.2). Timing of signal travel from spacecraft to user provides the basis for range measurement. The typical user receiver obtains ranges from
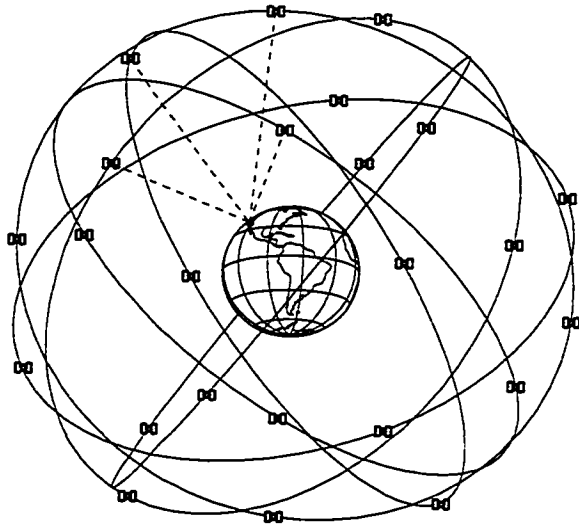
Figure 1.2: The Global Positioning System (GPS)

four (or more) satellites and with knowledge of the spacecraft locations, can solve for the three components of its position and the deviation of its receiver clock from GPS time.

GPS provides global coverage with a 24 space vehicle (SV) constellation, in which each SV is continuously broadcasting an L-band signal. The portion of the transmission normally accessible to the civil user is comprised of an L-band carrier modulated with a pseudorandom noise (PRN) code and a data stream. A unique PRN code is used for each spacecraft. For a given SV, ranging is based on measuring the offset between the received PRN code phase and an identical code generated internally in the receiver. The received data stream provides the receiver with the necessary information on spacecraft location [Spilker, b].

Standard civil GPS positioning accuracy is limited to roughly 100 meters (95% probability) by a number of error sources including an intentional degradation of the GPS signal by the DoD known as Selective Availability (S/A) [SPS]. Furthermore, although the GPS Operational Control Segment (OCS) does monitor the end-to-end performance of the navigation system, it does not have the means to provide timely alarms to the user in the event of a satellite or OCS ground segment failure.

Clearly, the application of GPS to civil aviation applications–particularly for precision approach and landing–is contingent upon significant improvements in both navigation accuracy and integrity.

1.2 Previous Work

A significant amount of research effort has been devoted to improving GPS navigation for high precision and high integrity applications. Great improvements in positioning accuracy and integrity have been made through the application of the *differential* GPS (DGPS) concept. Further progress in integrity has been achieved with Receiver Autonomous Integrity Monitoring (RAIM). In addition, a number of efforts at directly applying GPS to precision approach and landing navigation have been made.

1.2.1 Differential Positioning

The concept of Differential GPS–using code measurements–was introduced by Teasley, Hoover, and Johnson [Teasley] in 1980 and further detailed by Beser and Parkinson [Beser] in 1982. With DGPS, a reference GPS receiver is placed at a precisely known location in the near-vicinity of the user to calibrate those errors that are highly spatially correlated (including S/A). Positioning accuracy can thus be improved to roughly three meters (95% probability). In addition, the application of DGPS improves integrity in comparison to traditional 'stand-alone' positioning since, many types of spacecraft signal anomalies are mitigated through differencing.

In 1979, Counselman and Shapiro [Counselman] and MacDoran [MacDoran] independently extended the existing techniques of very long baseline interferometry (VLBI) to GPS, effectively making the first use of highly-precise GPS *carrier phase* measurements for static survey applications. Remondi (1986) expanded the concept for application to kinematic survey.

Although differential carrier phase positioning—often called *kinematic* positioning—can provide centimeter-level accuracy, the resolution of the *integer cycle ambiguity* for at least four satellites in view is mandatory. It is an unfortunate fact, however, that high-integrity, real-time cycle ambiguity resolution has been an elusive goal.

1.2.2 Integrity Monitoring

In the general areas of fault detection and isolation (FDI), an impressive body of literature has been accumulated during the last twenty-five years [Frank, Willsky, a]. Much of this work, in fact, has been targeted specifically toward the use of redundant sensor measurements in aerospace navigation systems (see for example [Daly]).

Within the context of satellite-based navigation, the concept of receiver autonomous integrity monitoring (RAIM) has been founded on the use of redundant satellite measurements. A number of important papers have provided the structure for the growing field of RAIM-based FDI. Brown and McBurney (1988) proposed the use of the maximum separation between satellite subset solutions for the detection of satellite failure [Brown, a]. Parkinson and Axelrad (1988) introduced the least-squares pseudorange residual as a test statistic for FDI [Parkinson, a]. Sturza (1988) formalized the residual-based approach using parity-space methodology [Sturza].

The limitations encountered in RAIM-based FDI performance for high-integrity/high-accuracy applications can be attributed directly to:

1. The achievable pseudorange measurement precision.

2. The availability of redundant measurements and adequate satellite geometry.

1.2.3 Application to Precision Approach and Landing

In the last ten years there have been a number of approaches directed toward the application of GPS to precision approach and landing. Klein and Parkinson (1985) [Klein] and Parkinson and Fitzgibbon (1989) [Parkinson, b] suggested the use of terminal area ground-based GPS pseudo-satellites ('pseudolites') to improve ranging geometry—and therefore, position accuracy—during precision approach and landing. Paielli, et al. (1993) at NASA Ames [Paielli], Van Graas, et al. (1993) of Ohio University [Van Graas, a], and Romrell, et al. (1995) at E-Systems [Romrell] independently applied differential carrier phase coupled with an 'on-the-fly' (OTF) cycle resolution algorithm for a series of test flights on transport aircraft. A. Brown, et al. (1993) of NAVSYS presented preliminary flight test results using carrier phase together with ground-based pseudolites to provide improved redundancy for OTF cycle resolution. Rowson, et al. (1994) [Rowson], Hundley, et al. (1995) [Hundley] of Wilcox Electric, and Van Graas, et al. (1995) [Van Graas, c] applied state-of-the-art, high-precision code receivers in DGPS architectures to safely perform automatic landings on transport-class aircraft.

Despite the rather impressive accomplishments to date, the fundamental goal of a GPS-based navigation system that provides the high accuracy and high integrity capabilities currently provided by ILS has not been achieved in this previous work.

1.3 Current Research

GPS research at Stanford University, supported by the Federal Aviation Administration (FAA), is focused on assessing the feasibility of satellite-based navigation for precision approach and landing. For reasons to be explained below, Stanford's Category III research effort is centered on combining kinematic carrier phase

measurements with ground-based pseudolites for the development of a navigation system capable of meeting the stringent ILS specifications for precision landing.

1.3.1 Kinematic Carrier Phase

Although both differential code and kinematic carrier phase technology have been proposed to meet the various challenges involved in precision approach and landing, the high precision of the GPS carrier phase provides the ultimate GPS navigation performance. Kinematic carrier phase offers two important advantages over code-based positioning for precision landing navigation:

1. Carrier phase positioning can provide the accuracy necessary to exceed the ILS accuracy specification of 2 ft vertical error (95% probability). Code-based positioning cannot provide the accuracy to meet the Category III ILS specifications with adequate time availability.

2. The high precision of carrier phase provides the leverage for RAIM in the sense that extremely tight detection thresholds may be set without incurring unacceptably high false alarm rates, thereby ensuring both *high integrity* and *high continuity*.

The high performance of carrier phase can only be achieved, however, if the integer cycle ambiguities can be accurately resolved for each space vehicle (SV).

1.3.2 Ground-Based Pseudolites

As noted earlier, the idea of using ground-based pseudolites for terminal area operations has a historical basis. Within the context of the application of kinematic carrier phase to Category III precision landing, the motivation for the use of ground-based pseudolites is two-fold:

11

1. When placed under the aircraft approach path, pseudolites provide the means for real-time cycle ambiguity resolution.

2. Pseudolites ensure the availability of redundant ranging measurements for RAIM.

1.3.3 The Integrity Beacon Landing System

The use of carrier phase measurements and ground-based pseudolites is combined in the Integrity Beacon Landing System (IBLS) concept (Figure 1.3). IBLS was conceived and developed at Stanford University [Cohen, b] as a high integrity solution to real time cycle ambiguity resolution for Category III precision approach. Two (or more) ground-based *GPS Integrity Beacon* pseudolites provide the basis for explicit estimation of cycle ambiguities during the approach. Integrity Beacons are simple, low power transmitters that broadcast L1 carriers modulated with unused PRN codes [Cobb]. The large geometry change that occurs during pseudolite overflight ensures the observability needed for cycle ambiguity estimation. Once cycle ambiguities have been initialized, real-time centimeter-level position fixes are possible through touchdown, rollout, and taxi—even when the aircraft is beyond the range of the Integrity Beacon.

1.4 Contributions

The fundamental goal of the Stanford effort is to design, build, and test a GPS-based navigation system that meets Category III precision landing sensor requirements. In response, kinematic carrier phase and ground-based pseudolites have been combined to form the foundation for the IBLS architecture.

Within this context, the central theme of the research documented in this thesis is the demonstration that GPS carrier phase measurements from spacecraft and ground-
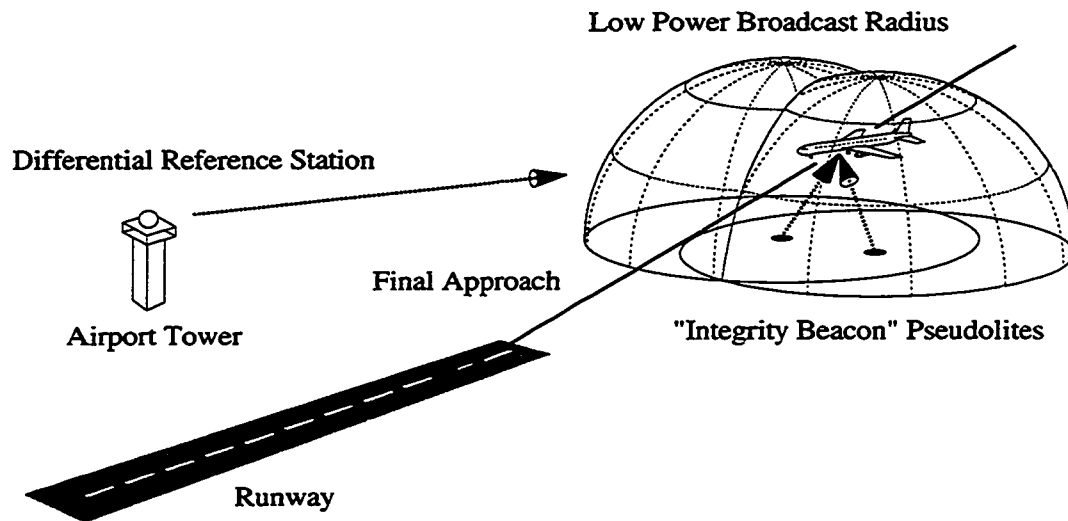
Figure 1.3: The Integrity Beacon Landing System (IBLS) Concept

based pseudolites can provide the basis for high navigation integrity. Contributions described in this thesis were made in the following specific areas:

1.4.1 Measurement Processing

Prototype algorithms for airborne kinematic carrier phase measurement processing, including high-speed algorithms for the first high-integrity airborne cycle ambiguity resolution, were developed, implemented, and flight-tested. Real-time algorithm performance was verified through an extensive battery of flight tests culminating in 110 successful automatic landings of a United Airlines Boeing 737-300. (Chapters 2 and 3)

1.4.2 Error Sources and Failure Modes

The error sources and fault modes for kinematic GPS navigation were characterized and their impact on navigation performance was assessed. Elements of the

navigation system architecture were evaluated and optimized based on failure detection capability and positioning accuracy in the presence of measurement error. (Chapter 4)

1.4.3 Fault Detection and Isolation

The theoretical framework of Receiver Autonomous Integrity Monitoring (RAIM) was generalized for application to carrier phase, cycle ambiguity resolution, and the wide range of navigation system failure scenarios. A new parity space algorithm was introduced to provide the capability for high integrity failure isolation. (Chapter 5)

1.4.4 Navigation Performance

The first quantitative definition of the parametric inter-relationship between GPS navigation system accuracy, integrity, continuity, and availability was achieved through analysis, simulation, and flight test. (Chapter 6)

CHAPTER 2

Kinematic Carrier Phase

The sub-centimeter precision of carrier phase measurements provides the potential for navigation at extremely high levels of accuracy and integrity. A number of basic architectural elements must be present, however, for kinematic (carrier phase) GPS navigation. These can be divided into space, ground, and airborne segments. The space segment is comprised of the GPS satellite constellation and the transmitted L-band signal. Carrier phase tracking receivers are present in both the ground and airborne segments of the kinematic architecture. The measured carrier phase observables from both the airborne and ground receivers are delivered to an airborne navigation computer (via digital data link for the ground measurements) for kinematic measurement processing.

2.1 Space Segment

The nominal GPS satellite constellation (Figure 1.2) consists of 24 space vehicles (SVs)–including three active spares–in circular, half-synchronous (11.97 hr) orbits. The satellites are distributed among six orbital planes inclined at 55 deg. Each plane contains four SVs. The right ascensions of the ascending nodes (Ω) of the planes are separated by 60 deg increments, and the nominal satellite phasing (Figure 2.1) within each plane has been selected to minimize the impact of a single satellite failure [Green]. At the time of
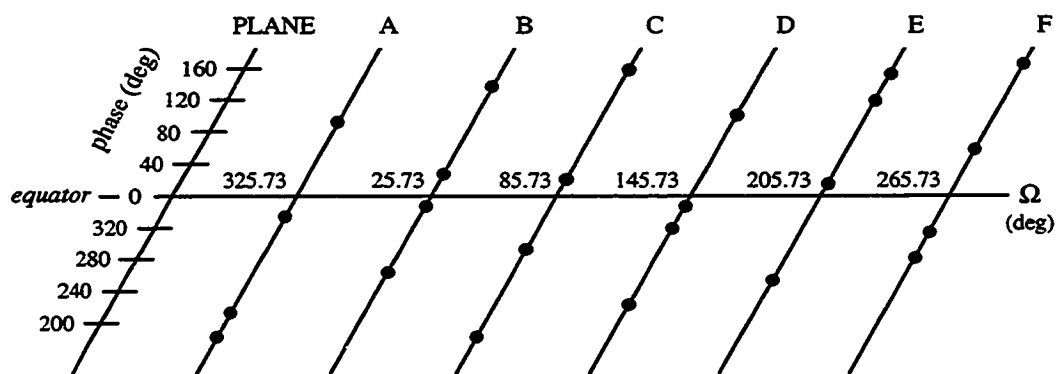
Figure 2.1: The GPS Satellite Constellation [Green]

writing, 24 Block II and IIA (second generation) and one Block I (first generation) spacecraft were in operation, the latter spacecraft in a 63 deg inclination orbit.

The GPS signal transmitted by the spacecraft is the sum of two L-band components: L1, with a center frequency at 1575.42 MHz, and L2, with a center frequency at 1227.6 MHz. The L1 carrier for each SV is modulated by a unique, week-long segment of a 37 week-long pseudorandom noise (PRN) code, known as the P (precision) code, which has a chip rate of 10.23 MHz. In addition, the quadrature component of the L1 carrier is modulated by 1 msec long PRN code, known as the C/A (clear acquisition) code, with a chip rate of 1.023 MHz. Both the in-phase and quadrature components of the L1 carrier are also modulated with a 50 bps data stream. The L1 signal transmitted by SV i may be expressed as [Spilker, b]

$$L1_i(t) = A_{P1}\, p_i(t)\, d_i(t) \cos \omega_{L1}t + A_{C/A}\, c_i(t)\, d_i(t) \sin \omega_{L1}t, \tag{2.1}$$

where $p_i(t)$, $c_i(t)$, $d_i(t)$ are the unique P and C/A codes and the navigation data, respectively, for satellite i. The constants A_{P1} and $A_{C/A}$ represent the relative signal amplitudes of the P and C/A codes. For L1, the C/A code power is roughly 3 dB greater

than that of the P code. The L2 signal, which does not incorporate the C/A code, can be expressed as

$$L2_i(t) = A_{P2} \, p_i(t) d_i(t) \cos \omega_{L2} t \,. \tag{2.2}$$

The dual frequency nature of the GPS signal provides the capability to remove the effect of the ionospheric delay [ICD-200], which is the largest natural error source affecting the GPS transmission. However, to ensure the "anti-spoofing" (A/S) capability of GPS and to limit positioning accuracy for users outside the U. S. military, the DoD has reserved the option of replacing the P code with an encrypted code, known as the Y code, at the same 10.23 MHz chipping rate [Parkinson, c]. Thus, the portion of the GPS signal accessible to the civil user is nominally only the quadrature (C/A code) component of the L1 transmission,

$$s_i(t) = A_{C/A} \, c_i(t) d_i(t) \sin \omega_{L1} t \,. \tag{2.3}$$

The pseudorandom C/A code modulation $c_i(t)$ effectively spreads the spectrum of the accessible L1 signal over a roughly 2 MHz bandwidth. Each spacecraft's C/A code is a unique Gold code sequence of +1 and −1. Because of the near-orthogonality property of Gold codes, code division multiple access (CDMA) is possible, in that the signal from any particular satellite can be recovered at the receiver by a correlation operation using a receiver-generated version of that satellite's C/A code [Spilker, b].

The digital data modulation $d_i(t)$ contains spacecraft health data, SV clock and ionosphere corrections, detailed orbital ephemeris information for SV i, and coarse orbital almanac information for all spacecraft in the constellation [ICD-200]. The ephemeris data provides the information on spacecraft location necessary to obtain a GPS position fix.

2.2 The GPS Receiver

The fundamental principle behind GPS satellite ranging is the measurement of the phase offset between the received PRN code for a given satellite and an identical code generated internally in the receiver. The measured code phase delay, known as the *pseudorange*, is the sum of the signal travel time and receiver's clock offset from GPS satellite time. With code phase measurements and navigation data from four or more SVs, enough information to compute position and clock offset from GPS time has been collected.

Figure 2.2 shows an example GPS receiver architecture consisting of an analog RF front end followed by a multi-channel digital signal processor. In the RF front end, the GPS signal is received, amplified, down-converted in frequency, and digitized for digital processing. Each digital signal processing channel is dedicated to a single satellite and consists of a delay-lock loop (DLL) and a phase-lock loop (PLL) as shown
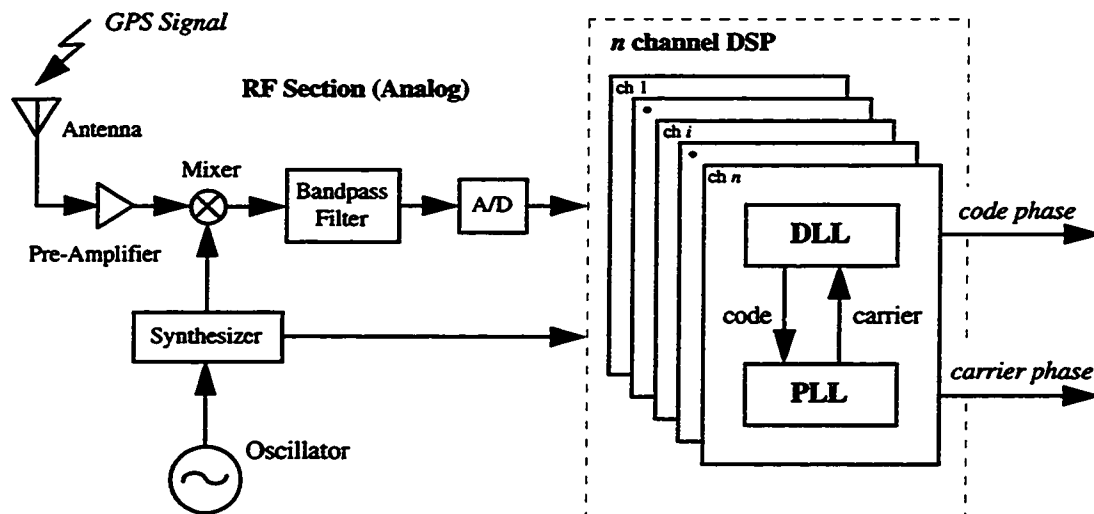
Figure 2.2: Basic GPS Receiver Architecture
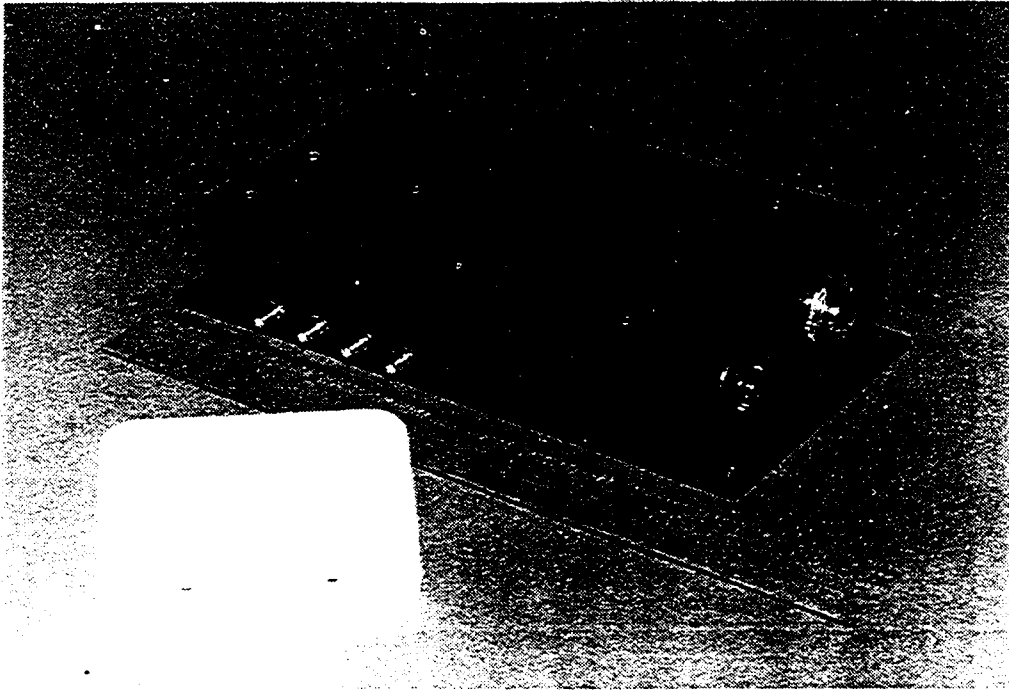
Figure 2.3: The Trimble TANS Quadrex Receiver

conceptually in Figure 2.2. The DLL generates an estimate of the code phase (pseudorange) and delivers a punctual C/A code signal to the PLL for reconstruction of the incoming carrier. The PLL provides a phase-locked carrier to the DLL for coherent detection of the modulated C/A code and generates an estimate of the carrier phase.

The precision of DLL code phase measurements is roughly on the order of 1 m (1σ)—less than 1% of the code chip length of 300 m. Similarly, PLL carrier phase measurements can also be made to a precision of roughly 1% of the carrier wavelength. Because the wavelength of the L1 carrier is approximately 19 cm, millimeter-level carrier phase measurement precision is possible. However, a constant integer cycle ambiguity is

present in each carrier phase measurement, and this ambiguity must be resolved before kinematic positioning can be done.

Figure 2.3 shows a photograph of the 6-channel Trimble Advanced Navigation Sensor (TANS) receiver and a hemispherical microstrip patch antenna/pre-amplifier. The basic TANS receiver architecture was modified by Cohen [Cohen, a] at Stanford University to provide the capability to generate carrier phase measurements. The resulting architecture, known as the TANS Quadrex, is the receiver used in the navigation experiments to be discussed in this dissertation.

2.3 The Carrier Phase Observable

The carrier phase measured at the receiver is the sum of the true range to the satellite, the cycle ambiguity, and a number of error sources. At an arbitrary GPS time t, the carrier phase observable $\phi_i(t)$ for spacecraft i can be mathematically expressed as

$$\phi_i(t) = \rho_i(t) + \tau^R(t) + N_i + \tau_i^{sv}[t - \rho_i(t)] - I_i(t) + T_i(t) + \varepsilon_i(t), \qquad (2.4)$$

where

$\rho_i(t)$ is the true range between the receiver and satellite i antenna phase centers,

$\tau^R(t)$ is the receiver clock offset from nominal GPS time,

N_i is the integer cycle ambiguity for satellite i,

$\tau_i^{sv}[t - \rho_i(t)]$ is the clock offset of SV i from nominal GPS time at the time of signal transmission (including S/A),

$I_i(t)$ is the carrier phase advance due to the presence of free electrons in ionosphere along the signal path,

$T_i(t)$ is the carrier phase delay due to tropospheric refraction, and

$\varepsilon_i(t)$ is the sum of multipath (signal reflections received by the antenna) and receiver noise errors.

Equation (2.4) has been nondimensionalized by implicitly expressing time in units of λ/c, where λ is the L1 wavelength (19.03 cm) and c is the vacuum speed of light, and distance in wavelengths (λ).

The error effects in the GPS observable are represented by the last four terms in equation (2.4). The effects of receiver noise and multipath will vary depending on the specific antenna and receiver hardware used, the line-of-sight to the spacecraft, and the localized geography (which is the source of signal reflections). In general, though, the magnitude of the resulting error is roughly only a few millimeters. The remaining three error effects, however, can be much larger. These error sources—ionosphere, troposphere, and spacecraft clock (including S/A)—can be referenced directly to a given spacecraft's transmission and their magnitudes are, therefore, independent of particular user site geography and receiver hardware. In addition, these errors are highly spatially correlated, in the sense that two receivers near each other (a few kilometers) will experience essentially identical errors. It is this spatial correlation of errors that is the basis for the concept of differential GPS.

2.4 Ground Segment

The ground segment of the kinematic GPS navigation system is comprised of a reference station and a number of ground-based pseudolites. The differential nature of kinematic GPS requires a reference station at a location which has been precisely surveyed relative to the airport runway, for removal of the spatially correlated errors discussed above. Additionally, as noted in Chapter 1, ground-based pseudolites must be present to provide the capability for explicit cycle ambiguity estimation. This notion will be discussed in much greater detail in Chapter 3.

2.4.1 Reference Station

The basic differential GPS reference station consists of a GPS receiver, to generate ranging measurements for all of the satellites in view, and a digital data link, to transmit the information to the aircraft for processing. The exact nature of the transmitted data can vary from one implementation to another. For example, the most straightforward data protocol is the direct uplink of the actual raw measurements collected at the reference GPS receiver. However, the transmit data structure can also take on the form of corrections to the measured ranges based on the predicted geometric range from the reference site to the satellite. The latter method, proposed by the Radio Technical Committee on Maritime Services (RTCM) Special Committee 104 [Kalafus], permits a smaller data rate and, therefore, relaxes the requirements on data-link bandwidth. However, if the ephemeris data used for a particular satellite is not identical in the aircraft and ground receivers, as is possible if one of the receivers decodes a new ephemeris message before the other, the aircraft cannot use the transmitted correction.

The effect of latency of received data, due to the finite data rate, is an important issue—regardless of the data protocol chosen. In practical applications, ground measurements are often projected ahead based on the past history of collected data to mitigate measurement latency. The specific prediction algorithm used can vary from simple polynomial fitting of a finite data window [Lawrence, a] to sophisticated adaptive filtering [Chou]. Measurement prediction can be implemented either in ground or airborne processing.

The IBLS breadboard reference station configuration used in flight testing is conceptually diagrammed in Figure 2.4. A Pentium-based PC collects raw carrier phase measurements from a six or nine channel TANS Quadrex receiver and RTCM SC-104 code phase corrections from a Trimble 4000 SSE survey receiver. Both GPS receivers are connected to the same microstrip patch receive antenna. Although only the TANS

22
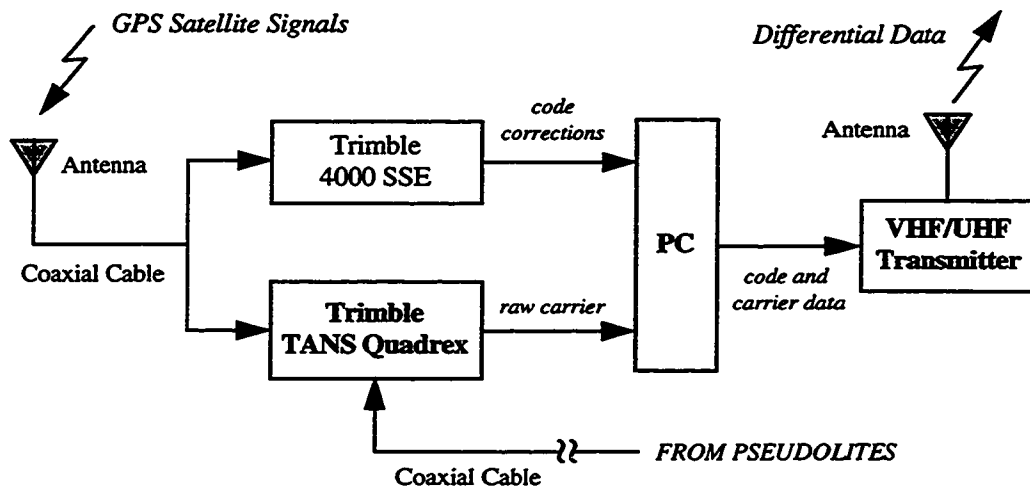
Figure 2.4: IBLS Breadboard Reference Station Configuration

receiver is actually necessary, the 4000 SSE is included in the breadboard configuration exclusively for convenience since it has the built-in capability to generate RTCM SC-104 code phase corrections. The PC creates raw carrier phase and code phase correction data packets for transmission to the aircraft via either a UHF or VHF transmitter. The 2400 baud capacity of the transmitter is sufficient to provide a 2 Hz measurement update rate to the aircraft.

2.4.2 Pseudolites

The use of ground-based pseudolites has been proposed for a number of different applications, including ranging augmentation for improved positioning accuracy and the transmission of differential corrections. Naturally, there has also been a corresponding variety of pseudolite designs [Cobb, *a, b*]. Traditionally, the primary difficulty with the use of pseudolites has been the large increase in signal power experienced by the airborne receiver as the pseudolite is approached which can result in jamming of the weaker satellite signal. This effect has become known as the 'near/far problem.'

The development of the *Integrity Beacon* pseudolite and its application to cycle ambiguity resolution has been a focal point for recent research at Stanford University. By limiting the required variation in range the pseudolite must accommodate during the approach, the near/far issue is avoided. Two classes of Integrity Beacon pseudolites have been developed under the direction of Cobb [Cobb, *b*]. These are the *Doppler Marker* and the *Omni Marker*.

The Doppler Marker [Cobb, *b*, Cohen, *e*], shown in Figure 2.5, is a simple low-power GPS signal transmitter. The transmitted signal consists of both L1 carrier and C/A code components, and thus appears to the airborne receiver to be simply an additional GPS satellite. For the signal to be received by an aircraft 1000 m away the transmit signal power is on the order of 1 μW. The volume within which the pseudolite signal is trackable by the airborne receiver is termed the pseudolite 'bubble.'

Figure 2.5: The Doppler Marker

The Omni Marker [Cobb, *b*, Cohen, *e*] is a somewhat more complicated design which also has potential application as a combined analog differential reference station and data link. The basic version consists of a receive antenna, a GPS receiver/transmitter, and a transmit antenna. The Omni Marker acquires the GPS signal for a given satellite through the receive antenna, strips off the PRN code, reapplies a new (unused) PRN code to the carrier, and rebroadcasts the signal via the transmit antenna. The outgoing and incoming code and carrier are phase coherent. At the aircraft, the differencing of this 'reflected' signal from the direct satellite signal provides an observable which has been implicitly differentially corrected. Thus the Omni Marker combines the role of a pseudolite as an additional ranging source (such as the Doppler Marker) with the additional capability to provide differential measurements directly. The Omni Marker concept is described in more detail in [Cohen, *e*].

The Doppler Marker, being the more straightforward application, will be used as the architectural basis for the underlying mathematical development to follow in the remainder of this dissertation. The mathematical development for the Omni Marker case is substantially similar but is somewhat more cumbersome in notation.

Should it be desired to use of pseudolites over a large variation in range, further design enhancements are necessary. A number of approaches aimed toward eliminating the associated near/far problem—including pulse modulation of the pseudolite signal—are summarized in [Cobb, *b*]. One possible application of such a 'terminal area' pseudolite, which will be discussed in more detail later, is to ensure the availability of redundant ranging measurements for autonomous fault detection aboard the aircraft.

2.5 Airborne Segment

The airborne segment of the kinematic GPS navigation system, as diagrammed in Figure 2.6, consists of three basic parts: a GPS receiver to generate ranging

measurements from all the satellites and pseudolites in view, a digital data receiver to retrieve the data transmitted by the reference station, and a measurement processor to convert the collected information into an aircraft position fix.

The airborne GPS receiver must be fed by two antenna inputs—one mounted on top of the fuselage to receive spacecraft signals and one mounted on the belly of the fuselage to receive pseudolite signals. The effective use of GPS measurements collected at two antennas requires knowledge of the displacement vector between the two antennas at each measurement epoch. This, in turn, is achieved by a one-time pre-flight calibration of the moment arm between antennas (in the aircraft body frame) together with real-time attitude determination. The IBLS breadboard navigation system uses a six or nine channel Trimble TANS Quadrex receiver on the aircraft together with an external attitude source. The origin of real-time attitude has varied depending upon the particular aircraft involved but has included both GPS-based attitude [Cohen, a] and inertial reference units.

A Pentium-based personal computer is the navigation processor used in the IBLS
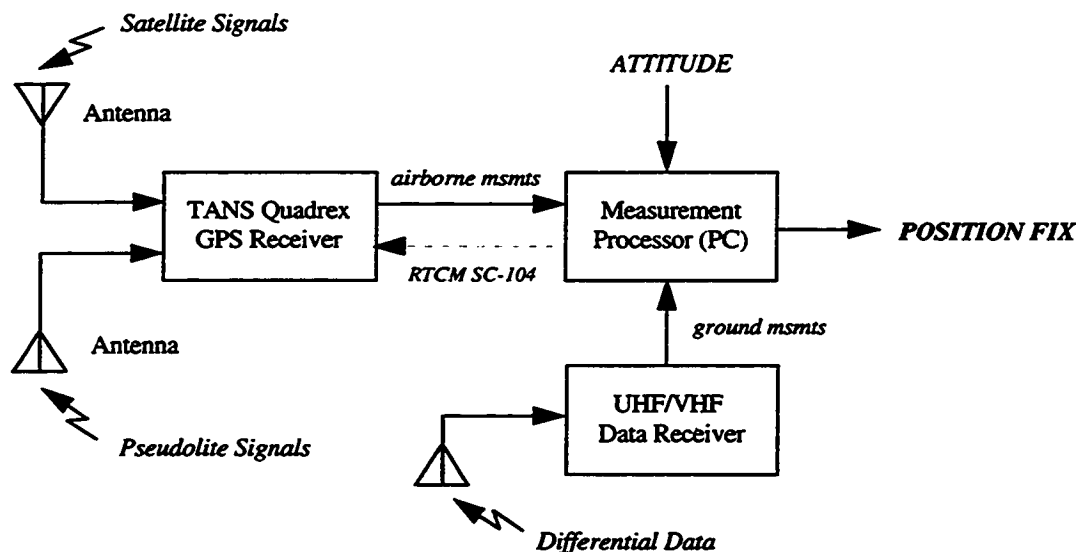
Figure 2.6: IBLS Breadboard Airborne Configuration

26

breadboard. The computer collects the digital raw carrier phase and code corrections transmitted from the reference station and received by an onboard UHF or VHF data receiver. The code phase corrections are subsequently forwarded to the TANS receiver which has the built-in capability to generate differentially-corrected code-based position fixes using RTCM SC-104 corrections. These code-based position fixes, which typically provide accuracy on the order of a few meters, are returned by the TANS receiver to the navigation processor. In addition, the TANS receiver also delivers satellite ephemeris data and raw carrier phase measurements to the processor. The measurement processing algorithms executing on the computer are centered on the use of the collected raw carrier phase measurements and the satellite ephemeris information, with the differential code-based position fixes primarily used for initialization purposes. The details of airborne measurement processing will be described in much greater detail in the remainder of this dissertation.

2.6 Single Difference Phase Observable

At a given true GPS time t, the raw aircraft-measured phase for satellite i can be expressed, following equation (2.4), as

$$\phi_i^a(t) = \rho_i^a(t) + \tau^a(t) + N_i^a + \tau_i^{rv}\left[t - \rho_i^a(t)\right] - I_i^a(t) + T_i^a(t) + \varepsilon_i^a(t). \tag{2.5}$$

This measurement is timetagged using *airborne receiver time*, $t^a(t)$, which is related to GPS time by

$$t^a(t) \equiv \tau^a(t) + t. \tag{2.6}$$

Similarly, at the reference station, raw phase measurements are timetagged using *reference receiver time*, $t^r(t)$, which is related to GPS time by

$$t^r(t) \equiv \tau^r(t) + t, \tag{2.7}$$

where $\tau^r(t)$ is the reference receiver clock bias. Thus, direct comparison of reference and aircraft phases is complicated by the fact that measurements with identical timetags will actually have been collected at different times.

The Trimble TANS receiver uses code phase measurements to perform internal position and clock bias estimation. The receiver is designed to use the resulting receiver clock bias estimate to adjust the receiver clock within ± 0.5 msec of GPS time. Therefore, the maximum difference between the true aircraft and reference measurement times is limited to 1 msec. However, considering that the satellite-to-aircraft range rate (doppler) can exceed 1000 m/s, deviations of greater than 1 m are still possible in direct comparison of the phases. A time-alignment procedure is necessary before the aircraft and reference phases can be compared.

Given aircraft and reference measurements with the same timetag, we can write

$$t^a(t) = t^r(t_0), \tag{2.8}$$

where t_0 is the true time of the reference phase measurement. The reference phase can thus be expressed as

$$\phi_i^r(t_0) = \rho_i^r(t_0) + \tau^r(t_0) + N_i^r + \tau_i^{sv}[t_0 - \rho_i^r(t_0)] - I_i^r(t_0) + T_i^r(t_0) + \varepsilon_i^r(t_0). \tag{2.9}$$

Performing a Taylor Series expansion of ϕ_i^r about t_0 yields,

$$\phi_i^r(t) = \phi_i^r(t_0) + \frac{d\phi_i^r(t_0)}{dt}(t-t_0) + \frac{1}{2}\frac{d^2\phi_i^r(t_0)}{dt^2}(t-t_0)^2 + \cdots. \tag{2.10}$$

Because the second and higher order terms contribute less than 0.1 mm, they can be neglected. Using equation (2.7) to express the derivative in reference receiver time, equation (2.10) can be rewritten as

$$\phi_i^r(t) = \phi_i^r(t_0) + \left[1 + \frac{d\tau^r(t_0)}{dt}\right]\frac{d\phi_i^r(t_0)}{dt^r}(t-t_0). \tag{2.11}$$

The term $d\tau^r(t_0)/dt$ is the ratio of the frequency offset of the TANS internal oscillator to L1 and is typically on the order of 10^{-6} or less. Its contribution to the phase expressed in (2.11) is extremely small (< 0.01 mm) and can be neglected. Using equations (2.6-8)

$$t - t_0 = -\left[\tau^a(t) - \tau^r(t)\right] - \left[\tau^r(t) - \tau^r(t_0)\right],\qquad(2.12)$$

where the second term in brackets represents the small change in the reference receiver clock bias in a time period less than 1 msec. Ignoring this term when (2.12) is substituted into (2.11) will result in an error smaller than 0.01 mm. Making this substitution and rearranging the result, the following is obtained:

$$\phi_i^r(t_0) = \phi_i^r(t) + \frac{d\phi_i^r(t_0)}{dt^r}\left[\tau^a(t) - \tau^r(t)\right].\qquad(2.13)$$

The *single difference phase* is defined as

$$\phi_i^a(t) - \phi_i^r(t_0) = \phi_i^a(t) - \phi_i^r(t) - \frac{d\phi_i^r(t_0)}{dt^r}\left[\tau^a(t) - \tau^r(t)\right].\qquad(2.14)$$

Performing the difference $\phi_i^a(t) - \phi_i^r(t)$ explicitly using (2.5) and (2.7), with t replacing t_0 in the latter equation, produces

$$\phi_i^a(t) - \phi_i^r(t_0) = \rho_i^a(t) - \rho_i^r(t) + \left[1 - \dot{\phi}_i^r(t_0)\right]\left[\tau^a(t) - \tau^r(t)\right] + N_i + v_i(t),\qquad(2.15)$$

where the notation ($\dot{}$) indicates differentiation in reference receiver time. The term $v_i(t)$ is primarily comprised of receiver noise and multipath experienced at both the aircraft and reference receivers but also includes smaller effects associated with the spatial decorrelation of ionospheric, tropospheric, and S/A errors. These latter contributions, which will increase in importance as the displacement between the aircraft and reference increases, will be discussed in more detail in Chapter 4.

The observation equation (2.15) provides the means for *implicit time alignment*, in the sense that aircraft and reference phases with identical timetags can now be differenced

directly. The tradeoff for this convenience is that the measured doppler at the reference receiver, $\dot{\phi}_i^r(t_0)$, must also be forwarded to the aircraft via the data uplink. However, the required doppler precision to keep time alignment error below 1 mm is only 5 Hz.

2.6.1 Satellite Ranging

The discrete time version of equation (2.15) for satellite i at time k is

$$\phi_{ik}^s = \rho_{ik}^a - \rho_{ik}^r + \left(1-\dot{\phi}_{ik}^{rs}\right)\tau_k + N_i^s + v_{ik}^s, \tag{2.16}$$

where the following definitions have been made:

$$\phi_{ik}^s \equiv \phi_{ik}^a - \phi_{ik}^r, \tag{2.17}$$

$$\tau_k \equiv \tau_k^a - \tau_k^r, \tag{2.18}$$

and N_i^s is the difference between aircraft and reference phase cycle ambiguities. Because the distance to the satellites is extremely large when compared to the displacement between the reference station and aircraft, the incoming satellite wavefronts are essentially planar (illustrated in Figure 2.7). The range difference for satellite i can therefore be mathematically expressed as the projection of the displacement vector from the reference to the aircraft, x_k, onto the line-of-sight (unit) vector to satellite i, e_{ik}. Thus, the linearized version of the observation equation (2.16) is
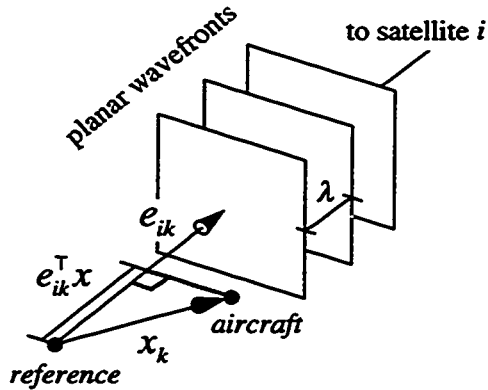
Figure 2.7: Planar Satellite Wavefronts

$$\phi_{ik}^s = -e_{ik}^\top x_k + \left(1-\dot{\phi}_{ik}^{rs}\right)\tau_k + N_i^s + v_{ik}^s. \tag{2.19}$$

30

The satellite line of sight vector e_{ik} is a simple and obvious function of the predicted satellite position (obtained from the broadcast ephemeris) and the approximate user position. With the application of a second order correction described in [Lawrence, c], which incorporates the prior knowledge of the code-based differential position, the error due to planarization of the spherical satellite wavefronts is limited to less than 1 mm over baselines as large as 10 km.

An example fifteen minute period of actual single difference phase data is shown in Figure 2.8 for one spacecraft. This data was collected across a short baseline (<1 m) on the rooftop of the Hansen Experimental Physics Building (HEPL) at Stanford University using two TANS Quadrex receivers. The apparent linear structure of the single difference phase observable is due to the dominance of the nearly constant drift in the clock bias τ_k (the relative frequency offset from L1) in the receivers. This clock bias
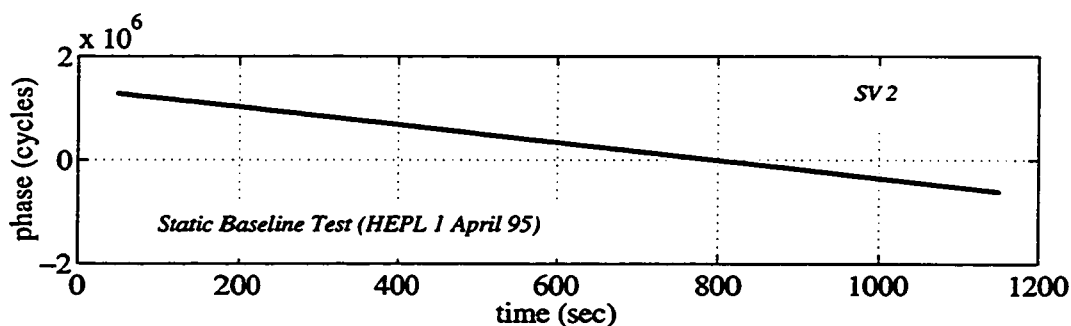
Figure 2.8: Single Difference Phase Observable
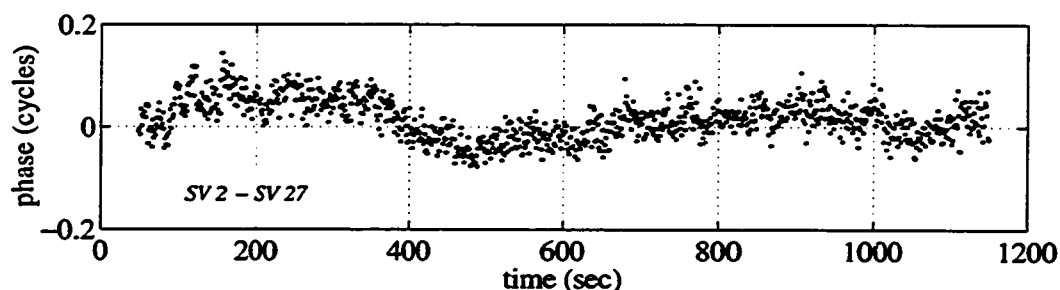
Figure 2.9: Double Difference Phase Observable

31

can be estimated when four of more satellites are in view (as will be discussed in Section 2.7.1) and is not considered a source of error. To observe the structure of the actual measurement error, v_{ik}^j, single difference phase measurements from two satellites can be subtracted from each other to obtain the *double difference phase* which effectively serves to remove the common mode clock bias term. The known baseline vector between the two antennas can also be removed, leaving only the difference between the single difference measurement errors for the two satellites. This simple double differencing procedure, however, is somewhat complicated by differing coefficients of the clock bias term for the two satellites resulting from the implicit time alignment formulation. A simple *generalized double difference* procedure to alleviate this difficulty is described in Appendix A. The resulting double difference phase error is shown in Figure 2.9. The error structure exhibits both a low-frequency component due to multipath and a white component primarily due to receiver noise. The double difference root-mean-square (rms) error is less than 0.05 cycles (λ), which corresponds to less than 7 mm rms error for the single difference phase error for each independent satellite. Although these error sources will be discussed in greater detail in Chapter 4, the millimeter level precision of carrier phase measurements can be readily seen in the results of this simple experiment.

2.6.2 Pseudolite Ranging

Anticipating the application of ground-based pseudolites in the near-vicinity of the aircraft and reference station (see Figure 2.10), linearity of pseudolite wavefronts cannot be assumed. The single difference observable for pseudolite i must instead be expressed in discrete-time form by
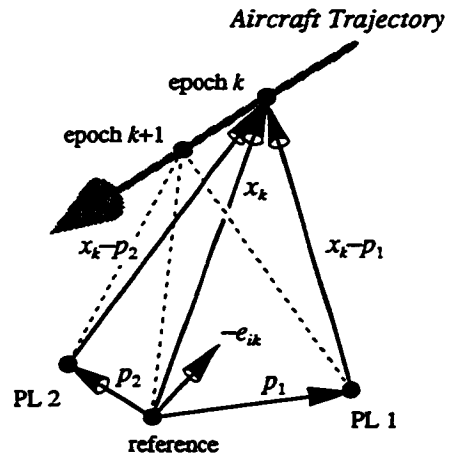
Figure 2.10: Pseudolite Geometry

$$\phi_{ik}^p = |p_i - x_k| + \left(1 - \dot{\phi}_{ik}^{Tp}\right)\tau_k + N_i^p + v_{ik}^p \qquad (2.20)$$

where p_i is the baseline vector from the reference station antenna to the transmit antenna of pseudolite i. Once the ground segment is installed, this baseline vector can be obtained to millimeter-level accuracy using established GPS static survey techniques. The constant N_i^p differs in character from its spacecraft counterpart N_i^s by the addition of a bias associated with the transmission of the pseudolite signal to the reference station via coaxial cable. In general, N_i^p will not be an integer.

2.7 Kinematic Positioning

The carrier phase measurements for n satellites at epoch k can be stacked as follows:

$$\begin{bmatrix} \phi_{1k}^s \\ \vdots \\ \phi_{nk}^s \end{bmatrix} = \begin{bmatrix} -e_{1k}^{\mathsf{T}} & 1 - \dot{\phi}_{1k}^{rs} \\ \vdots & \vdots \\ -e_{nk}^{\mathsf{T}} & 1 - \dot{\phi}_{nk}^{rs} \end{bmatrix} \begin{bmatrix} x_k \\ \tau_k \end{bmatrix} + \begin{bmatrix} N_1^s \\ \vdots \\ N_n^s \end{bmatrix} + \begin{bmatrix} v_{1k}^s \\ \vdots \\ v_{nk}^s \end{bmatrix}. \qquad (2.21)$$

If the cycle ambiguities are known and $n \geq 4$, it is possible to solve (2.21) to obtain the aircraft position relative to the reference station, x_k. In this event, given the pre-surveyed location of the reference station with respect to the runway (stored in an onboard database), centimeter-level position fixes can be forwarded to the aircraft autopilot.

For simplicity of notation, equation (2.21) is rewritten as

$$\phi_k^s = H_k^s u_k + N^s + v_k^s \qquad (2.22)$$

where H_k^s is the $n \times 4$ observation matrix and all other terms are $n \times 1$ vectors. In the general case where a prior estimate of the cycle ambiguity vector, $\bar{N}^s$, and its associated covariance, $\bar{P}_N$, are available, the least-squares solution to (2.22) is

$$\hat{u}_k = \left(H_k^{s\mathsf{T}} V_s^{-1} H_k^s\right)^{-1} H_k^{s\mathsf{T}} V_s^{-1} \left(\phi_k^s - \bar{N}^s\right) \qquad (2.23)$$

33

The weighting matrix

$$V_s^{-1} \equiv \overline{P}_N^s + I_n \sigma_\phi^2 \tag{2.24}$$

is the sum of the cycle ambiguity error covariance matrix and the measurement error covariance matrix. The form of the latter matrix implicitly assumes independent, identically distributed (i.i.d.) measurements with variance σ_ϕ^2. If $\overline{N}^s$ is an unbiased estimate of N^s and all errors are normally distributed, equation (2.23) provides an unbiased, maximum likelihood estimate of u_k.

The estimate error covariance is given by

$$\hat{P}_u = \mathrm{E}\left[(u_k - \hat{u}_k)(u_k - \hat{u}_k)^\mathsf{T}\right] = \left(H_k^{s\mathsf{T}} V_s^{-1} H_k^s\right)^{-1}. \tag{2.25}$$

Relation (2.25) expresses the quality of the position and clock estimate, $\hat{u}_k$, in terms of how well the cycle ambiguities are known $\left(\overline{P}_N\right)$, the precision of the carrier phase measurements $\left(\sigma_\phi\right)$, and the quality of the particular spacecraft geometry $\left(H_k^s\right)$.
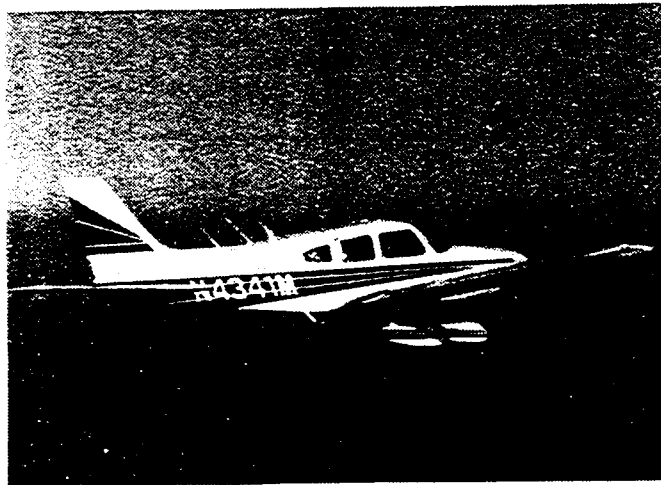
Figure 2.11: Piper Dakota with GPS Antennas

2.7.1 Flight Test Example

The first flight experiments with the IBLS breadboard architecture were performed on a Piper Dakota (Figure 2.11) at Palo Alto Airport in January, 1993. This aircraft is equipped with five GPS antennas: one on top of the fuselage for satellite ranging, one on the fuselage belly for pseudolite ranging, and three additional antennas (on the wingtips, and tail) to provide the capability for attitude determination using GPS [Cohen, a]. The aircraft and ground hardware used was identical to that described in sections 2.4 and 2.5, with the exception of the datalink, which was not present in these initial tests. Instead, data was stored onboard the aircraft and at the reference station for post-processing.

Using the collected data (carrier phase measurements and SV ephemeris), equation (2.23) was used to generate the post-processed kinematic GPS trajectory shown in Figure 2.12 (a) for one Piper Dakota approach, landing, rollout, and taxi to tiedown (note the different axis scales). For this trajectory, the cycle ambiguities
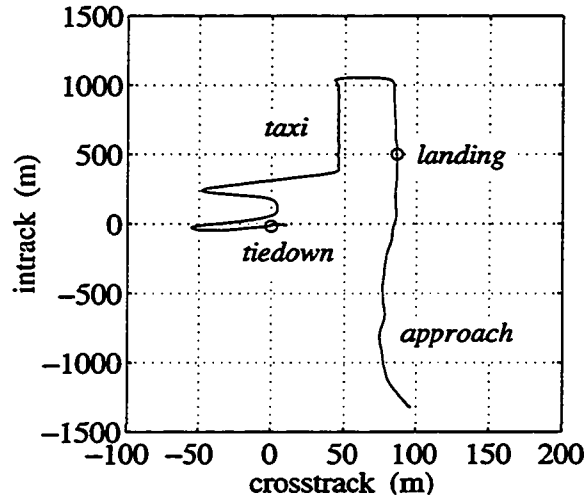
Figure 2.12 (a): Piper Dakota Kinematic GPS Trajectory
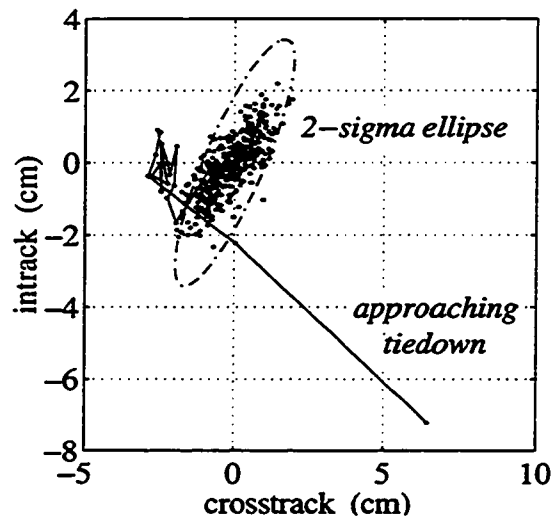
Figure 2.12 (b): Kinematic Positioning at Tiedown

35

were obtained using ground-based pseudolites and verified through a post-tiedown static GPS survey. While the means for obtaining cycle ambiguity estimates and associated covariances will be discussed in great detail in Chapter 3, it is instructive at present to examine the precision of kinematic positioning in this limit case example where the cycle ambiguities are known exactly $\left(\overline{P}_N = 0\right)$. The precision of the kinematic position fixes can actually be observed in Figure 2.12 (b) which shows a magnified view of the aircraft being pushed to its final tiedown position where approximately five minutes of carrier phase measurements from 6 SVs were collected at a 1 Hz sample rate. The tiedown coordinates were arbitrarily set to be the origin of the plot. As seen in the figure, the horizontal dispersion of position fixes does not exceed 3 cm and is consistent with the predicted 95% error ellipse based on the satellite geometry at the time and a measurement standard deviation of $\sigma_\phi = 1$ cm.

CHAPTER 3

Cycle Ambiguity Resolution

T he principal problem that must be solved before the benefits of carrier phase can be realized is the resolution of the unknown integer cycle ambiguities with high integrity. This chapter briefly describes a number of previously proposed methods of cycle ambiguity resolution, and then introduces a new, superior methodology, founded on the use of ground-based pseudolites, for *explicit estimation* of the cycle ambiguities. The mathematical basis of the new approach is detailed, and high-speed algorithms suitable for real-time airborne execution are derived. Flight test results are included to demonstrate algorithm performance.

3.1 Methods of Cycle Ambiguity Resolution

Cycle ambiguity resolution has been successfully and routinely achieved using several existing methods. However, as these methods have generally been developed to support static and kinematic survey applications, the definition of 'success' has always been quite different than that imposed by the stringent requirements and constraints associated with aircraft precision approach and landing.

37

3.1.1 Independent Initialization

The unknown cycle ambiguities can be initialized at any epoch k if an independent source of aircraft position, accurate to the centimeter-level, is available. If the known position is substituted for x_k in equation (2.21), the integers can be identified subject to the integer/clock bias observability conditions described in Section 3.2.1 below. The drawback of airborne application of this simple notion for cycle ambiguity resolution is obvious: no alternative navigation source exists to provide positions of such accuracy on final approach. Furthermore, if such a system did exist, kinematic GPS would likely not be needed at all.

3.1.2 Carrier-Smoothed Code

The use of carrier phase measurements to 'smooth' pseudorange (code) measurements has been suggested by a number of authors [Goad, Hwang, Hatch, a]. This approach is possible because of the complementary nature of the two measurements: The carrier phase provides a very precise ranging measurement but with an unknown bias, whereas the code provides an absolute, but much noisier measurement. In theory, if the code measurements were smoothed for a long enough time, the effective ranging accuracy would approach centimeter-level carrier phase limit and the cycle ambiguities would be implicitly known. In practice, however, the performance of the technique is fundamentally limited over short durations by code phase multipath errors which are both biased (non-zero mean) and colored (non-white) and over longer durations by the divergence of the code and carrier ionospheric errors [MacDoran]. These performance limitations inhibit the direct application of the carrier-smoothed code technique to cycle ambiguity resolution or Category III navigation (as defined by the ILS requirements). Nevertheless, carrier smoothing has typically provided good results in terms of positioning accuracy. For example, during precision approach flight trials, Paielli et al.

have demonstrated 2σ vertical accuracy of approximately one meter using this technique [Paielli].

3.1.3 Search Methods

When redundant satellite measurements are present, it is possible to check the consistency of a given candidate set of integer cycle ambiguities through the measurement residual. Based on this principle, a number of systematic integer search algorithms—sometimes called 'on-the-fly' (OTF) resolution algorithms—have been developed. (For example, see [Hatch, b] and [Euler].) These methods can be characterized, in a general sense, by the following simplified algorithm: A candidate integer set N_i^s is substituted into equation (2.22) in place of the true integer vector N^s, and the least-squares solution is obtained from equation (2.23) assuming that the cycle ambiguities are known exactly (as the candidate set). The measurement residual is then computed using

$$ r_k^i \equiv \phi_k^s - N_i^s - \hat{u}_k^i \tag{3.1} $$

where $\hat{u}_k^i$ is the least squares solution for candidate integer set i. The candidate integer set is then varied in a systematic fashion, and the process above is repeated until the magnitude of the residual r_k^i is minimized. Specific OTF techniques may differ in both the mechanics of the search and the selection of integer subspace to be searched, although the latter is generally obtained with the aid of code or carrier-smoothed code measurements. In addition, measurements of the GPS L2 carrier have also been used (together with L1 carrier phase measurements) with the goal of reducing the time required to complete the search; a description of this concept, which is known as 'widelaning,' may be found in [Paielli].

While OTF methods have been successfully used in several applications, including kinematic GPS survey [Frodge], their suitability for precision approach and landing is questionable for two reasons:

39

1. They are limited in availability since they implicitly rely on having redundant satellites in view. Additionally, as noted in Section 2.1, the L2 signal is not guaranteed to be available for civil use.

2. Navigation integrity and continuity are compromised because multiple solutions are possible.

Clearly, it is the second deficiency that is most serious. The reason this observation is true is quite simple: Any candidate set of integers that lies 'near' the range of the observation matrix H_k', as defined by equations (2.21, 2.22), will produce a least-squares residual that is near zero. Thus, in the presence of nominal measurement errors, a false integer set may actually have a comparable or even lower residual than the correct set. In this event, an OTF algorithm is faced with the option of either risking *integrity* by possibly selecting the wrong integer set or interrupting *continuity* by deliberately not choosing either, thereby aborting the remainder of the approach.

3.1.4 Motion-Based Methods

Satellite motion, together with redundant measurements, can provide the observability for *direct estimation* of the integer cycle ambiguities. This approach is ideally suited for GPS applications in which time is not a significant constraint; in fact, most GPS static survey systems are founded on this principle. Additionally, the use of redundant measurements for explicit motion-based estimation can ensure the *integrity* of cycle ambiguity resolution. This property—to be explained in detail in Chapter 5—makes motion-based estimation highly attractive for precision approach and landing applications. It is an unfortunate fact, however, that the rate of satellite motion is very slow in comparison with the time scales of most real-time applications—including precision approach and landing. As a result, satellite motion alone can provide only marginal leverage for real-time cycle ambiguity estimation [Lawrence, *c*, Van Graas, *b*].

In addition, five or more satellites are needed for direct estimation in non-static situations; thus, the availability of satellite-motion-based resolution is limited.

3.2 Real-Time Resolution using Pseudolites

The introduction of ground-based pseudolites can provide the advantages and alleviate the disadvantages of motion-based cycle ambiguity resolution. If the pseudolites are placed under the aircraft approach path, the large geometry change that occurs during aircraft overflight can provide, on a time scale of a few seconds, the observability needed for explicit cycle ambiguity resolution. Furthermore, if pseudolites are used, redundant satellites are not needed; thus, availability is ensured.

3.2.1 Observability

The geometric basis for cycle ambiguity resolution using ground-based pseudolites can be explained in relatively simple mathematical terms. For clarity of explanation, the (scalar) pseudolite and (vector) satellite single difference phase observables (2.20-22) are simplified as

$$\phi_k^p = |x_k| + \tau_k + N^p \tag{3.2}$$

$$\phi_k^s = H^s \begin{bmatrix} x_k \\ \tau_k \end{bmatrix} + N^s \tag{3.3}$$

where spacecraft motion during pseudolite overflight, time alignment terms, and measurement error have been ignored. These neglected effects are irrelevant to the present observability discussion. In addition, a single pseudolite located at $p_i = 0$ (the reference site) has been assumed.

Differencing the satellite phase at epoch k in equation (3.3) from the phase at epoch 0 (the first epoch during pseudolite overflight) and inverting, the following relation is easily obtained:

$$\begin{bmatrix} x_{k-0} \\ \tau_{k-0} \end{bmatrix} \equiv \begin{bmatrix} x_k - x_0 \\ \tau_k - \tau_0 \end{bmatrix} = H^{s-1} \phi_k^s , \tag{3.4}$$

where it has been assumed that four satellites are used so that H^s is invertible. The terms x_{k-0} and τ_{k-0} describe the *relative* trajectory and clock that are available from satellite ranging alone. Substituting these terms into the pseudolite observation equation (3.2) and collecting bias terms into a single constant b, the following expression is obtained:

$$\phi_k^p = \left| x_{k-0} + x_0 \right| + \tau_{k-0} + b . \tag{3.5}$$

The initial position (x_0) and bias (b) are now the only variables on the right-hand side of equation (3.5), and the variation in the pseudolite phase profile with respect to these parameters is given to first order by

$$\delta\phi_k^p = \frac{x_k^T}{|x_k|} \delta x_0 + \delta b . \tag{3.6}$$

Consider now the idealized linear pseudolite overflight trajectory shown in Figure 3.1 in which the radial (vertical) component of absolute position is constant and the crosstrack component is identically zero. In this case, equation (3.6) becomes

$$\delta\phi_k^p = \begin{bmatrix} \dfrac{x_k^i}{|x_k|} & 0 & \dfrac{x_0^r}{|x_k|} & 1 \end{bmatrix} \begin{bmatrix} \delta x_0^i \\ \delta x_0^c \\ \delta x_0^r \\ \delta b \end{bmatrix} . \tag{3.7}$$

The basis functions for intrack $\left(x_k^i / |x_k| \right)$ and radial $\left(x_0^r / |x_k| \right)$ absolute position are expressed graphically in Figure 3.2 versus intrack position (normalized by x_0^r). This figure shows that the effects on the pseudolite phase profile of radial and intrack absolute

42
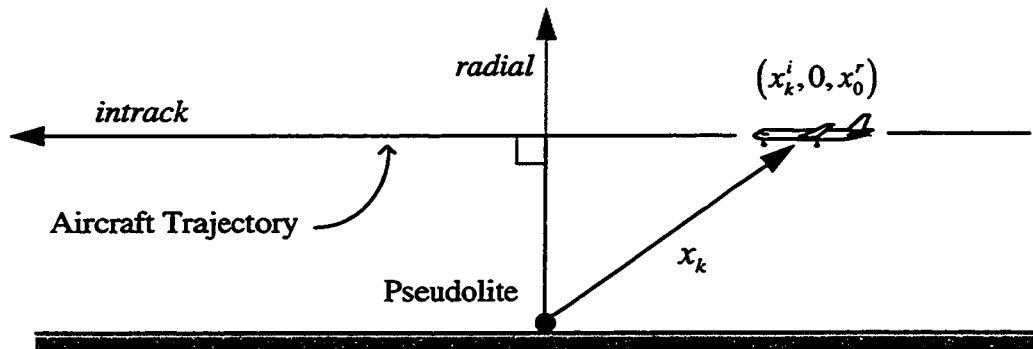
Figure 3.1: Pseudolite Overflight Geometry
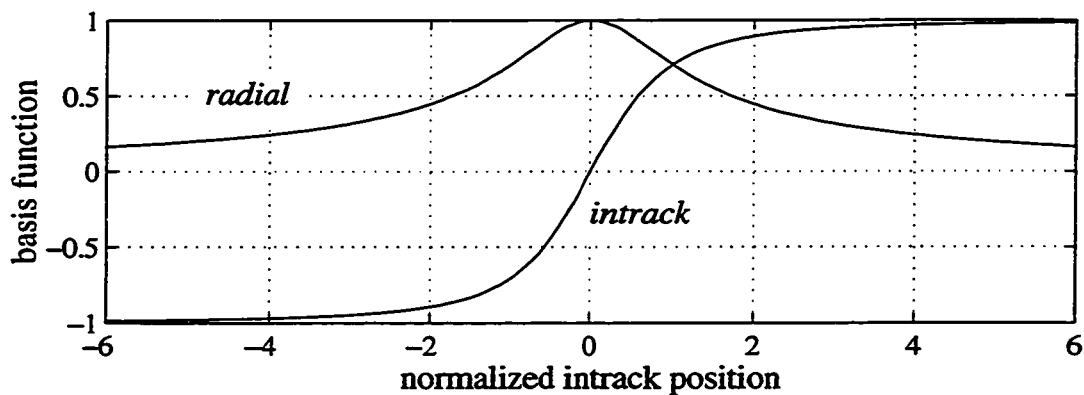
Figure 3.2: Absolute Position Basis Functions

position deviations are clearly distinguishable from each other and from the bias in equation (3.7) provided that there is a reasonably large relative geometry change—as determined by the normalized intrack position history. However, it is also clear from equation (3.7) that no information is present in the pseudolite phase profile concerning deviations in crosstrack absolute position. In fact, all trajectories consistent with the relative position history provided by equation (3.3), that have an initial intrack position of x_0^i and lie on the surface of a cylinder with radius x_0^r (with central-axis along the ground track), will also be consistent with the collected pseudolite phase measurements.

43

The introduction of a second pseudolite, however, can provide the means for full three-dimensional absolute positioning. Figure 3.3 shows an intrack view of the cylinders of absolute position (represented by the circular contours in the figure) using two pseudolites. For the

(intrack view)
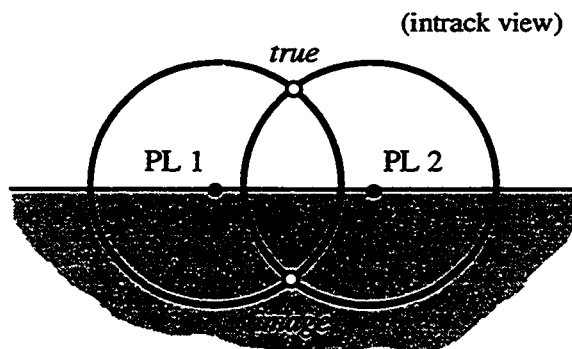
Figure 3.3: Contours of Absolute Position

idealized linear trajectory described above, the contours of absolute position intersect at two points: the true solution and an 'image' solution. In reality, since the image solution is underground, it can be easily rejected as false. Furthermore, it will be demonstrated in Section 3.3 that given even relatively poor initial knowledge of the aircraft position, convergence to the true solution can be ensured.

Cycle Ambiguity Observability

In practice, the information collected from the satellites and pseudolites during pseudolite overflight is processed to obtain cycle ambiguity estimates, $\overline{N}^s$, and estimate-error covariance matrix, $\overline{P}_N^s$, for all spacecraft in view. These are, in turn, substituted into equations (2.23-24) to provide absolute kinematic position fixes at any epoch after pseudolite overflight.

Within this context, the issue of cycle ambiguity observability must be addressed. Specifically, the single difference observation equations (2.19-20) contain both the cycle ambiguities $\left(\overline{N}_i^s \text{ and } \overline{N}_i^p\right)$ and a clock bias term (τ_k). Because the coefficients of the clock bias, $1-\dot{\phi}_{ik}^\sigma$ and $1-\dot{\phi}_{ik}^p$, are always nearly unity, and the clock bias is common to all ranging measurements, it is essentially impossible to distinguish the integer cycle ambiguity for a given channel from the clock bias. Fortunately, this is not required. To

44

resolve this issue, equations (2.19-20) are rewritten in terms of cycle ambiguity *differences* for the n_s satellite measurements and n_p pseudolite measurements, as follows:

$$\phi_{1k}^s = -e_{1k}^\top x_k + \left(1 - \dot{\phi}_{1k}^{rs}\right)\left(\tau_k + N_1^s\right) + \dot{\phi}_{1k}^r N_1^s + v_{1k}^s \tag{3.8}$$

$$\phi_{ik}^s = -e_{ik}^\top x_k + \left(1 - \dot{\phi}_{ik}^{rs}\right)\left(\tau_k + N_1^s\right) + \left(N_i^s - N_1^s\right) + \dot{\phi}_{ik}^r N_1^s + v_{ik}^s \quad \left(i = 2, n_s\right) \tag{3.9}$$

$$\phi_{ik}^p = \left|p_i - x_k\right| + \left(1 - \dot{\phi}_{ik}^{rp}\right)\left(\tau_k + N_1^s\right) + \left(N_i^p - N_1^s\right) + \dot{\phi}_{ik}^r N_1^s + v_{ik}^p \quad \left(i = 1, n_p\right) \tag{3.10}$$

For simplicity, the integer for satellite 1 has been chosen as 'master' for differencing; however, the choice of master satellite is arbitrary and does not in any way affect absolute positioning. If $N_1^s < 200$ cycles, all products of the form $\dot{\phi}_{ik}^r N_1^s$ in equations (3.8-10) are less than 0.2 mm and can be neglected. This condition is easily satisfied if an initial differential code phase measurement for satellite 1 is used to obtain an estimate of this cycle ambiguity, and then this estimate is subtracted from all subsequent carrier phase measurements for satellite 1. In fact, this procedure is implemented in the IBLS breadboard measurement processing algorithms for all satellites in view. Consequently equations (3.8-10) can be simplified as

$$\phi_{1k}^s = -e_{1k}^\top x_k + \left(1 - \dot{\phi}_{1k}^{rs}\right)\tau_k^1 + v_{1k}^s \qquad \text{(master satellite)} \tag{3.11}$$

$$\phi_{ik}^s = -e_{ik}^\top x_k + \left(1 - \dot{\phi}_{ik}^{rs}\right)\tau_k^1 + \Delta N_i^s + v_{ik}^s \qquad \left(i = 2, n_s\right) \tag{3.12}$$

$$\phi_{ik}^p = \left|p_i - x_k\right| + \left(1 - \dot{\phi}_{ik}^{rp}\right)\tau_k^1 + \Delta N_i^p + v_{ik}^p \qquad \left(i = 1, n_p\right) \tag{3.13}$$

where the following definitions have been made:

$$\tau_k^1 \equiv \tau_k + N_1^s \tag{3.14}$$

$$\Delta N_i^s \equiv N_i^s - N_1^s \tag{3.15}$$

$$\Delta N_i^p \equiv N_i^p - N_1^s. \tag{3.16}$$

The observation equations (3.11-13) will be used as the basis for the development of cycle ambiguity resolution algorithms in the following sections.

3.2.2 IBLS Architecture Considerations

A typical aircraft approach with IBLS will begin with carrier-smoothed code DGPS. Because the level of navigation performance provided during this phase of flight is consistent with the ILS requirements for Category I, the pseudolites can be placed at or near the present ILS Middle Marker site. This location corresponds to an aircraft overflight altitude of approximately 200 ft–the Category I DH. In order to avoid the near-far issue discussed in Section 2.4.2, the power output of the pseudolite is set conservatively low to provide a bubble radius of approximately 600 ft. (This near-far restriction can be removed if pseudolite signal pulsing is implemented [Cobb, b].) Once inside the bubble, the aircraft collects carrier phase measurements from both the satellites and pseudolites. At bubble exit, the cycle ambiguities are resolved, and centimeter level positioning is possible as long as four or more satellites are maintained.

Attitude and Moment Arm

One issue unique to the IBLS architecture which must also be addressed is the fact that the satellite signals are tracked via an antenna mounted on top of the aircraft fuselage whereas the pseudolite signals are received through an antenna mounted on the belly of the fuselage. Specifically, if x_k is defined to be the displacement vector from the reference station antenna to the aircraft top antenna at epoch k, then (3.13) must be more precisely expressed for $(i=1,n_p)$ as

$$\phi_{ik}^p = \left| p_i - \left(x_k + \ell_k \right) \right| + \left(1 - \dot{\phi}_{ik}^T \right) \tau_k^1 + \Delta N_i^p + v_{ik}^p \qquad (3.17)$$
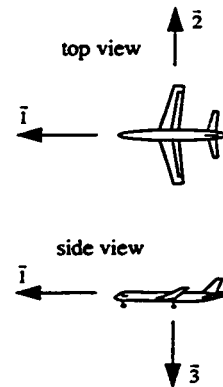
Figure 3.4:
Aircraft Body Axes

46

where ℓ_k is the 'moment arm' vector from the top antenna to the belly antenna. When x_k is expressed in local reference coordinates (for example, east, north, and up–ENU), then the moment arm will necessarily be a function of aircraft attitude.

Given a one-time preflight calibration of the moment arm ℓ^b, expressed in aircraft body axes (see Figure 3.4), and an aircraft orientation expressed as a rotation ψ in azimuth (from north), followed by a pitch of θ from the local horizontal, and roll angle φ, equation (3.17) can be rewritten as

$$\phi_{ik}^p = \left| p_i - \left(x_k + A_k \ell^b \right) \right| + \left(1 - \dot{\phi}_{ik}^{rs} \right) \tau_k^I + \Delta N_i^p + v_{ik}^p \qquad \left(i = 1, n_p \right) \qquad (3.18)$$

where the rotation matrix A_k is defined as

$$A_k \equiv \begin{bmatrix} \cos\theta\sin\psi & \cos\varphi\cos\psi + \sin\varphi\sin\theta\sin\psi & -\sin\varphi\cos\psi + \cos\varphi\sin\theta\sin\psi \\ \cos\theta\cos\psi & -\cos\varphi\sin\psi + \sin\varphi\sin\theta\cos\psi & \sin\varphi\sin\psi + \cos\varphi\sin\theta\cos\psi \\ \sin\theta & -\sin\varphi\cos\theta & -\cos\varphi\cos\theta \end{bmatrix}_k . \qquad (3.19)$$

This same procedure can also be used to translate the fuselage top-antenna position to an appropriate autopilot reference point such as the aircraft landing gear or ILS antenna.

3.3 Mathematics of Cycle Ambiguity Resolution

The estimation of cycle ambiguities involves the processing of pseudolite measurements (3.18) which are nonlinear in the aircraft position. The development of new cycle ambiguity estimation algorithms to process these measurements is driven by the high speed and high integrity needs of the envisioned real-time airborne application.

3.3.1 Nonlinear Least-Squares Estimation

Consider a general set of nonlinear observations collected over a given time interval

$$z = h(x) + v \qquad (3.20)$$

where z is a vector of M measurements, x is the vector of unknown parameters of dimension less than M, $h(\cdot)$ is the nonlinear relationship between the parameters and the measurements, and v is a zero-mean vector of random sensor errors whose elements are independent and identically distributed. Now if no prior information on the value of x is available, and x is observable only through (3.20), then a *least-squares estimate*, $\hat{x}$, which minimizes the cost function

$$J = (z - h(x))^{\mathrm{T}} (z - h(x)) \qquad (3.21)$$

is possible.

In addition to the familiar and intuitively reasonable cost function, least-squares estimators have a number of important properties:

1. When the measurement errors v are gaussian, the least-squares estimate $\hat{x}$ is also a *joint maximum likelihood estimate* of x [Jazwinski].

2. For the special case where the observation equation (3.21) is linear, the least-squares estimate of x is also a *minimum variance* estimate (or *unbiased estimate*), in that $\mathrm{E}\left[(x - \hat{x})^{\mathrm{T}}(x - \hat{x})\right]$ is minimized [Bryson].

3. The least squares problem is generally easily solved.

The most straightforward solution to the nonlinear least squares problem posed by (3.20-21) involves a Gauss-Newton iteration procedure as follows:

$$\hat{x}^{i+1} = \hat{x}^{i} + H^{i+}\left(z^{i} - h(\hat{x}^{i})\right) \qquad (3.22)$$

where

$$H^{i} \equiv \left.\frac{dh(x)}{dx}\right|_{x=\hat{x}^{i}} \qquad (3.23)$$

and

$$H^{i^+} \equiv \left(H^{i^T}H^i\right)^{-1}H^{i^T}. \tag{3.24}$$

The linear least-squares algorithm (3.22) is repeatedly executed until $\hat{x}^{i+1} - \hat{x}^i$ is negligible for the specific application under consideration. The convergence properties of (3.22) will depend on the specific nature of the function h and the quality of the initial guess $\hat{x}^0$.

In contrast to the 'batch' procedure just described, sequential least squares processing of nonlinear observations is also possible using a number of approximate approaches based on the linear Kalman Filter. These include, for example, the Extended Kalman Filter, the Iterated Extended Kalman Filter [Gelb], and the iterative algorithm of Haupt et al. [Haupt] While these algorithms can provide good results for many real-time non-linear applications where state updates are desired whenever a new set of measurements is collected, in the general case they do not minimize an identifiable least-squares cost function and do not provide a maximum likelihood estimate of the states. Thus, the resulting risk of excessive estimate error using sequential least-squares methods is inconsistent with the high level of navigation integrity required for precision approach and landing.

Within the context of the practical implementation of IBLS (see Section 3.2.2), where a finite number of nonlinear measurements are to be collected within the pseudolite bubble, cycle ambiguity estimation is well-suited for batch processing. Furthermore, centimeter-level positioning accuracy—which is possible only after the cycle ambiguities have been estimated—is not needed until after bubble exit.

3.3.2 Batch Algorithms

The resolution of cycle ambiguities by means of a Gauss-Newton batch algorithm requires the linearization of the pseudolite observation equation about the best current trajectory estimate. Given an approximate initial trajectory $\bar{x}_k$ obtained from code-based DGPS the pseudolite observation equation (3.18) can be expressed in terms of the deviation from the approximate trajectory: $\delta x_k \equiv x_k - \bar{x}_k$. Keeping first order terms only, the result is as follows:

$$\delta\phi_{ik}^p \equiv \phi_{ik}^p - \left| p_i - \left(\bar{x}_k^i + A_k \ell^b \right) \right| = -\bar{e}_{ik}^{\mathsf{T}} \delta x_k + \left(1 - \dot{\phi}_{ik}^{\mathit{rp}} \right) \tau_k^i + \Delta N_i^p + v_{ik}^p \quad \left(i = 1, n_p \right) \qquad (3.25)$$

where

$$\bar{e}_{ik}^{\mathsf{T}} \equiv \left[p_i - \left(\bar{x}_k^i + A_k \ell^b \right) \right]^{\mathsf{T}} \Big/ \left| p_i - \left(\bar{x}_k^i + A_k \ell^b \right) \right| \qquad (3.26)$$

is the approximate line-of-sight vector at epoch k from the aircraft top antenna to pseudolite i. For consistency, the satellite observations are also expressed in terms of the deviation from the approximate trajectory as follows:

$$\delta\phi_{1k}^s \equiv \phi_{1k}^s + e_{1k}^{\mathsf{T}} \bar{x}_k = -e_{1k}^{\mathsf{T}} \delta x_k + \left(1 - \dot{\phi}_{1k}^{\mathit{rs}} \right) \tau_k^1 + v_{1k}^s \qquad (3.27)$$

$$\delta\phi_{ik}^s \equiv \phi_{ik}^s + e_{ik}^{\mathsf{T}} \bar{x}_k = -e_{ik}^{\mathsf{T}} \delta x_k + \left(1 - \dot{\phi}_{ik}^{\mathit{rs}} \right) \tau_k^1 + \Delta N_i^s + v_{ik}^s \quad \left(i = 2, n_s \right) \qquad (3.28)$$

Defining $\delta\phi_k$ to be the vector of n_s satellite and n_p pseudolite measurements at epoch k as follows

$$\delta\phi_k \equiv \begin{bmatrix} \delta\phi_{1k}^s \\ \vdots \\ \delta\phi_{n_sk}^s \\ \delta\phi_{1k}^p \\ \vdots \\ \delta\phi_{n_pk}^p \end{bmatrix}, \qquad (3.29)$$

and E_k, the observation matrix at epoch k, as

$$E_k \equiv \begin{bmatrix} -e_{1k}^{\mathrm{T}} & 1-\dot{\phi}_{1k}^{rs} \\ \vdots & \vdots \\ -e_{n_sk}^{\mathrm{T}} & 1-\dot{\phi}_{n_sk}^{rs} \\ -\overline{e}_{1k}^{\mathrm{T}} & 1-\dot{\phi}_{1k}^{rp} \\ \vdots & \vdots \\ -\overline{e}_{n_pk}^{\mathrm{T}} & 1-\dot{\phi}_{n_pk}^{rp} \end{bmatrix}, \tag{3.30}$$

and stacking the measurements obtained at the n epochs inside the pseudolite bubble, the following matrix equation results

$$\begin{bmatrix} \delta\phi_1 \\ \vdots \\ \delta\phi_k \\ \vdots \\ \delta\phi_n \end{bmatrix} = \begin{bmatrix} E_1 & 0 & \cdots & 0 & 0 & \overline{I} \\ 0 & \ddots & 0 & \ddots & 0 & \vdots \\ \vdots & \ddots & E_k & \ddots & \vdots & \overline{I} \\ 0 & \ddots & 0 & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 0 & E_n & \overline{I} \end{bmatrix} \begin{bmatrix} \delta u_1 \\ \vdots \\ \delta u_k \\ \vdots \\ \delta u_n \\ N \end{bmatrix} + v. \tag{3.31}$$

where the following definitions have been assumed:

$$\overline{I} \equiv \begin{bmatrix} 0^{\mathrm{T}} \\ I \end{bmatrix}, \tag{3.32}$$

$$\delta u_k \equiv \begin{bmatrix} \delta x_k \\ \tau_k^1 \end{bmatrix}, \tag{3.33}$$

and

$$N \equiv \begin{bmatrix} \Delta N_2^s \\ \vdots \\ \Delta N_{n_s}^s \\ \Delta N_1^p \\ \vdots \\ \Delta N_{n_p}^p \end{bmatrix}, \tag{3.34}$$

and 0 is the null vector of length $n_s + n_p - 1$.

51

The linear least squares solution to (3.31) is obtained and the resulting estimate of the trajectory deviation $\delta \bar{x}_k$ is used to adjust the approximate trajectory: $\bar{x}_k^{new} = \bar{x}_k^{old} + \delta \bar{x}_k$. As directed by the Gauss-Newton algorithm (3.22), the process is repeated using the new trajectory and is continued through convergence.

In practice, the difficulty with the straightforward application of this procedure is due to the large size of the observation matrix in equation (3.31), which when $n = 50$ can exceed 500×200, requiring nearly 1 Mbyte of memory to store. More importantly, the total computation time for such a problem can surpass 60 sec on a Pentium-based computer. These difficulties, in particular the long computation time, prohibit the direct application of the simple batch least squares algorithm outlined above, and establish the need for the development of new, efficient, high-speed algorithms for nonlinear least-squares estimation.

3.3.2.1 Sparse Matrix Batch Least-Squares

The sparse nature of the observation matrix in (3.31) provides the basis for the needed improvements in both execution time and storage efficiency. For example, when $n = 50$, roughly only 50 kbytes of memory are required if the zeros in (3.31) are not stored. The sparsity of a matrix with substantially similar structure to (3.31) was exploited by Cohen [Cohen, a] in the treatment of a related cycle resolution problem associated with GPS-based attitude determination. The sparse matrix block Cholesky decomposition algorithm derived in [Cohen, a] was successfully applied with some modification to (3.31). Post-processing of previously collected raw flight data demonstrated substantial improvements in execution time, in that cycle ambiguities were resolved in 1 to 2 sec.

Despite the great improvements in speed, sparse matrix batch algorithms are relatively inflexible to certain conditions seen in practical airborne applications.

Specifically, consider the situation where the GPS receiver loses phase lock on a previously tracked satellite or pseudolite or acquires a new satellite or pseudolite while the algorithm is in the process of stacking phase data for batch processing. These events are, in fact, very common in practice, especially for low-elevation satellites and for pseudolites when the aircraft is near the bubble boundary. The inherently rigid structure of batch least-squares algorithms, however, makes accounting for these events very difficult in that a given satellite or pseudolite may occasionally have two or more cycle ambiguities that need to be resolved—only one of which will actually be applicable after bubble exit. Thus, the motivation is provided for the development of algorithms that provide the least-squares solution to (3.31) and retain the high speed characteristics of sparse matrix algorithms but are more adaptable to the conditions present in real-time airborne applications. These objectives can be met by a reformulation of the problem within the context of sequential information smoothing [Bryson].

3.3.2.2 Iterated Information Smoother

At any epoch k within the pseudolite bubble, the vector of satellite and pseudolite measurements can be written in vector form as

$$\delta\phi_k = \begin{bmatrix} E_k & \bar{I} \end{bmatrix} \begin{bmatrix} \delta u_k \\ N \end{bmatrix} + v_k , \tag{3.35}$$

where the measurement noise vector v_k is distributed as $v_k \sim N\left(0, \sigma_\phi^2 I_{n_s + n_p}\right)$ and the other variables in (3.35) are as previously defined. If no satellites or pseudolites are lost or acquired between epochs k and $k+1$, the aircraft kinematics can be expressed as a discrete Gauss-Markov process as follows:

$$\begin{bmatrix} \delta u \\ N \end{bmatrix}_{k+1} = \begin{bmatrix} \delta u \\ N \end{bmatrix}_k + w_k \tag{3.36}$$

where the process noise vector w_k is distributed as $w_k \sim N(0, W)$ and

$$W \equiv \lim_{\mu \to \infty} \begin{bmatrix} \mu I_4 & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.37}$$

The process noise covariance matrix W effectively invokes infinite process noise on the vehicle position and clock states in order to emulate the purely kinematic structure of (3.31) and zero process noise on the cycle ambiguities which are, of course, constant in time.

Given a state estimate-error covariance $\hat{P}_{k-1}$, the covariance time update $\overline{P}_k$ based on (3.36-37) is

$$\overline{P}_k = \hat{P}_{k-1} + W = \lim_{\mu \to \infty} \begin{bmatrix} \mu I_4 & \hat{P}_{uN} \\ \hat{P}_{uN}^T & \hat{P}_N \end{bmatrix}_{k-1}, \tag{3.38}$$

where $\hat{P}_{uN}$, $4 \times (n_s + n_p - 1)$, and $\hat{P}_N$, $(n_s + n_p - 1) \times (n_s + n_p - 1)$, are the right upper and lower sub-blocks of $\hat{P}$. This result can also be expressed in *information matrix* form as

$$\overline{S}_k = \begin{bmatrix} 0 & 0 \\ 0 & \hat{P}_N^{-1} \end{bmatrix}_{k-1} \tag{3.39}$$

where $\hat{P}_N^{-1}$ is the inverse of the lower right block of $\hat{P}$ that corresponds to the integer states. The information matrix measurement update can then be applied to obtain

$$\hat{S}_k = \begin{bmatrix} 0 & 0 \\ 0 & \hat{P}_N^{-1} \end{bmatrix}_{k-1} + \begin{bmatrix} E_k^T \\ I^T \end{bmatrix} [E_k \quad I]. \tag{3.40}$$

where, for clarity, the measurement covariance has been normalized by σ_ϕ^2. Using familiar matrix block inversion identities [CRC], equation (3.40) can be written explicitly in information form as follows:

$$\hat{S}_k = \begin{bmatrix} 0 & 0 \\ 0 & \hat{S}_N - \hat{S}_{uN}^T \hat{S}_u^{-1} \hat{S}_{uN} \end{bmatrix}_{k-1} + \begin{bmatrix} E_k^T \\ I^T \end{bmatrix} [E_k \quad I], \tag{3.41}$$

where the matrix $\hat{S}_u$, the upper left 4×4 block of the information matrix, is full rank at epoch k if four or more measurements (satellite and pseudolite) are available. In order for GPS positioning to be possible in any form, this condition must always be satisfied.

Defining the *information vector* as

$$y_k \equiv \begin{bmatrix} y_u \\ y_N \end{bmatrix}_k \equiv S_k \begin{bmatrix} \delta u \\ N \end{bmatrix}_k, \qquad (3.42)$$

where y_u is 4×1 and y_N is $(n_s + n_p - 1) \times 1$, the state update can be shown through a similar derivation to be given by

$$\hat{y}_k = \begin{bmatrix} 0 \\ \hat{y}_N - \hat{S}_{uN}^\mathsf{T} \hat{S}_u^{-1} \hat{y}_u \end{bmatrix}_{k-1} + \begin{bmatrix} E_k^\mathsf{T} \\ I^\mathsf{T} \end{bmatrix} \delta \phi_k . \qquad (3.43)$$

Under nominal circumstances, when satellites and pseudolites are not lost or acquired during the bubble pass, equations (3.41) and (3.43) can be applied sequentially, without modification, (with initial conditions $\overline{y}_1 = 0$ and $\overline{N}_1 = 0$) until the final bubble epoch n is reached. At this point, the information matrix $\hat{S}_n$ is invertible and cycle ambiguity estimates may be computed using (3.42). However, to ensure a solution consistent with a Gauss-Newton solution to (3.31), an updated trajectory must be obtained and the process must be repeated through convergence (see discussion following equation (3.34)). In order to obtain an identical updated trajectory to that provided by the batch solution to (3.31), backward smoothing of the collected measurements is necessary. This is easily done by interchanging the k and $k-1$ indices in equations (3.41) and (3.43) and regressing sequentially through the data with the 'initial conditions' $\overline{y}_n = 0$ and $\overline{S}_n = 0$. The information vector and matrix obtained from backward processing are added to the stored results from the forward pass according to

$$\hat{S}_k = \hat{S}_k^F + \overline{S}_k^B \qquad (3.44)$$

$$\hat{y}_k = \hat{y}_k^F + \overline{y}_k^B . \qquad (3.45)$$

The trajectory update $\delta\hat{x}_k$ can now be realized as a vector element of $\hat{S}_k^{-1}\hat{y}_k$. The entire process is then repeated until $\delta\hat{x}_k$ is negligibly small.

The utility of the information smoother algorithm is that, like the batch formulation of (3.31), no prior information is needed to begin. Not surprisingly, this property is also useful when new satellites or pseudolites are acquired or lost during stacking and processing. These events are accommodated in a relatively straightforward manner using the matrix block inversion methods described in Appendix B.

Furthermore, due to the efficient and compact structure inherent in sequential least squares processing, the iterated information smoother retains the high speed characteristics of the sparse matrix batch solution; for a stack size of $n = 50$, execution time on a Pentium processor is typically 1-1.5 sec.

3.3.2.3 Reduced-Order Iterated Smoother

It is possible to further increase speed of execution through the application of a *reduced-order* smoother formulation. The central assumption underlying this approach is that the satellite measurements provide information regarding the relative kinematic trajectory only, while the pseudolite measurements provide the means for absolute placement (in the local airport reference frame) of this relative trajectory. This notion is essentially identical to that used in the observability analysis of (3.2.1). In a strict sense, however, the cycle ambiguity estimates that result from this approach will be *approximate* because they are not obtained from the least-squares solution to the system of equations (3.31). Nevertheless, the underlying assumptions are not unrealistic, and thus, the improvement in speed afforded by the reduced order algorithm motivates its further development.

The vector of satellite phase measurements at epoch k can be expressed as

$$\delta \phi_k^s = E_k^s \delta u_k + \begin{bmatrix} 0 \\ \Delta N_2^s \\ \vdots \\ \Delta N_{n_s}^s \end{bmatrix} + v_k^s \ . \tag{3.46}$$

Similarly the pseudolite phase measurements can be written as

$$\delta \phi_k^p = E_k^p \delta u_k + \begin{bmatrix} \Delta N_1^p \\ \vdots \\ \Delta N_{n_p}^p \end{bmatrix} + v_k^p \ . \tag{3.47}$$

Subtracting the integer vector from both sides of (3.46), solving for δu_k, and substituting the result into (3.47) yields

$$\delta \phi_k^p - E_k^p E_k^{s+} \delta \phi_k^s = \begin{bmatrix} -E_k^p E_k^{s+} & I_{n_p} \end{bmatrix} \begin{bmatrix} 0 \\ N \end{bmatrix} - E_k^p E_k^{s+} v_k^s + v_k^p \ , \tag{3.48}$$

where N is defined in (3.34) and

$$E_k^{s+} \equiv \left(E_k^{s^T} E_k^s \right)^{-1} E_k^{s^T} \ . \tag{3.49}$$

The reduced-order observation equation (3.49) can be written in compact form as

$$\delta \phi_k^R = E_k^R N + v_k^R \ , \tag{3.50}$$

where E_k^R is the observation matrix in (3.48) with the first column removed, and the error vector is distributed as

$$v_k^R \sim N(0, V_k), \qquad V_k \equiv \sigma_\phi^2 \left[I_{n_p} + E_k^p \left(E_k^{s^T} E_k^s \right)^{-1} E_k^{p^T} \right] \ . \tag{3.51}$$

The stack of measurements (3.50) collected over n epochs in the pseudolite bubble can be processed by sequential static least-squares techniques. Since no prior information on the integer state vector N is assumed, the following information algorithm is used:

$$\hat{S}_n^N = \sum_{k=1}^n \left[\hat{S}_k^N + E_k^{R^T} V_k^{-1} E_k^R \right]$$ (3.52)

$$\hat{y}_n^N = \sum_{k=1}^n \left[\hat{y}_k^N + E_k^{R^T} V_k^{-1} \delta\phi_k^R \right]$$ (3.53)

where $\left[\hat{S}_k^N\right]^{-1}$ is the integer covariance and $\left[\hat{S}_k^N\right]^{-1}\hat{y}_k^N$ is the estimate of the integer vector N at epoch k. Once all n epochs are processed, the resulting cycle ambiguity estimates are substituted into (3.46), and a weighted least-squares fit (based on the integer covariance) is used to obtain the trajectory update. The entire processes is then repeated until the magnitude of the trajectory update is negligibly small.

For a stack size of $n = 50$, execution time of the reduced order-information smoother algorithm is typically 0.5-1 sec on a Pentium processor–almost one second faster than either the sparse matrix batch algorithm or the iterated information smoother. It is noted again, however, that the increase in speed is gained at the expense of not obtaining the actual least squares solution to (3.31); the resulting cycle ambiguity error compared with the least squares solution is typically on the order of 0.001 L1 cycles (<1 mm). Although both the iterated information smoother and reduced-order smoother algorithms have been used extensively in real-time flight testing, the former algorithm is currently implemented in the IBLS breadboard architecture because it provides greater flexibility with respect to envisioned incremental architecture changes (such as the application of Omni Marker pseudolites described in Section 2.4.2).

3.4 Convergence

Clearly, it is crucial that convergence be guaranteed for an algorithm intended for use in real-time zero-visibility precision landing operations. For a given pseudolite overflight, the convergence characteristics of the simple Gauss-Newton iteration procedure described in the previous sections will depend exclusively on the quality of the

initial guess trajectory. In order to assess the limits of convergence of the Gauss-Newton iteration procedure, a computer simulation of the IBLS navigation system was developed, which to the greatest extent possible, incorporated the actual real-time flight algorithms. The satellite constellation orbital parameters were obtained from a recent GPS constellation almanac; the satellite orbits were propagated analytically using a straightforward Keplerian orbit model. GPS satellite outages were included in the simulation using the Phlong and Elrod [Phlong] satellite state availability model, which incorporates allowances for both scheduled downtime for satellite maintenance (such as stationkeeping maneuvers) and unscheduled downtime due to satellite system failures.

The baseline IBLS architecture assumed three pseudolites in a triangular layout at the ILS middle marker; the motivation for this number and configuration is given in Section 5.2.4.1. The corresponding altitude of pseudolite overflight was approximately 200 ft. In addition, a pseudolite bubble radius of 600 ft, a 7.5 deg elevation mask, and a 5 Hz update rate were used. A standard 3 deg glideslope approach at San Francisco International (SFO) Runway 28 was repeatedly simulated in the presence of varying satellite geometry and availability. Finally, the error in the initial guess trajectory was systematically varied with the goal of generating a convergence map for the Gauss-Newton iteration procedure.

Figure 3.5 (a) shows the region of algorithm convergence with respect to crosstrack and vertical errors in the initial guess trajectory for all satellite geometries. (Also shown in the figure are the crosstrack positions of the pseudolites—indicated by the 'o' symbols in the figure.) The results in this figure are interpreted as follows: The aircraft position during the approach (indicated by the '+' symbol) is perfect; there are no deviations from the nominal approach path. However, the aircraft's *knowledge* of its crosstrack and vertical position components can be in error; the size of this error is specified by the assumed aircraft position given by the horizontal and vertical axes in the

59

figure. For comparison, the approximate NSE performance of code-based DGPS—the actual source of the initial guess trajectory—is also indicated in the figure. Clearly, even extremely rare DGPS errors of large magnitude lie well within the computed crosstrack-vertical convergence boundary. A similarly large convergence margin for intrack DGPS errors (not shown in the figure) was also obtained from the simulation.

In practice, of course, it is improbable that the aircraft will fly precisely between the pseudolites at the nominal altitude. However, in order to ensure a safe approach and landing, the aircraft Total System Error (TSE)—the actual deviation from the nominal flight path— must not be too large. For example, at a nominal altitude of
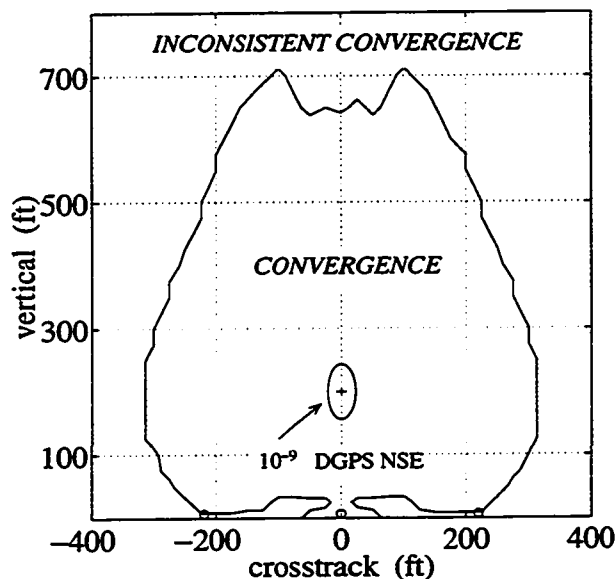
Figure 3.5 (a): Convergence Map (Nominal Approach)
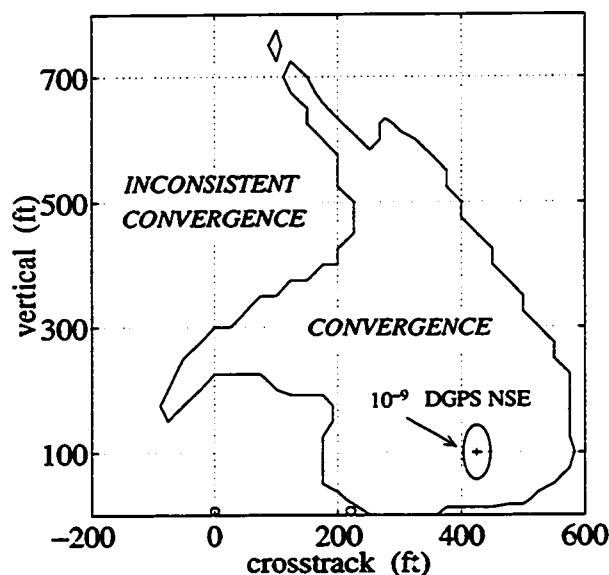
Figure 3.5 (b): Convergence Map (Limit-Case Approach)

200 ft, the proposed outer containment surface boundary for a GPS-based approach has a half-width and half-height of roughly 400 ft and 100 ft respectively [Kelly]. A breach of this containment surface is permitted to occur in only one out of ten million approaches. Therefore, in order to bound the convergence performance of the Gauss-Newton iteration procedure, the simulation described above was repeated with the actual aircraft position coincident with the containment surface boundary. The result obtained from simulation for the worst-case aircraft offset is shown in Figure 3.5 (b), in which the actual aircraft position deviates from nominal by 400 ft in crosstrack and is also 100 ft below nominal altitude. The simulation results show that, even in this limiting case, a large convergence margin is still present.

3.5 Flight Test Verification

The performance of the real-time information smoothing algorithms has been demonstrated through extensive flight testing. While a more thorough summary of overall navigation accuracy results obtained from the various stages of flight testing is given in Chapter 6, the subsequent sections of this chapter provide an experimental verification of the capability for real-time airborne cycle ambiguity resolution.

3.5.1 Piper Dakota Approach

As noted in Section 2.7.1, the first flight tests of IBLS were performed on a Piper Dakota at Palo Alto Airport in January, 1993. Since a datalink was not present in these initial tests, data was stored onboard the aircraft and at the reference station for post-processing. The purpose of these initial flight trials was simply to verify the overall concept and the prototype carrier phase processing algorithms.

With this goal in mind, the flight experiments were carried out with two ground-based (Doppler Marker) pseudolites placed under the approach path, such that the altitude

at pseudolite overflight was approximately 100 m. The ratio of pseudolite bubble radius to the radius of closest approach was set to approximately 3:1 for both pseudolites. During the pseudolite overflight, four channels on each of the TANS receivers (reference and airborne) were used to track satellites, and the remaining two channels were dedicated to tracking the two pseudolites. The pseudolite and spacecraft carrier phase measurements were collected and stored at a rate of 1 Hz at the aircraft and ground station. The aircraft attitude was obtained from an independent onboard GPS-based attitude system [Cohen, a] and was also stored to allow for post-flight computation of the moment arm corrections (3.18-19). After the last approach and landing, the aircraft taxied to its tiedown position at which point an independent static GPS survey of the baseline from the reference station antenna to the aircraft antenna was performed. This baseline provided a source of truth for the aircraft position at tiedown accurate to approximately one centimeter and, thereby, provided the means for independent verification of cycle ambiguities for the last approach.

The data collected during pseudolite overflight was post-processed in the laboratory using the iterated information smoother algorithm. The guess trajectory used to initialize the Gauss-Newton iteration procedure was obtained using code-based DGPS position fixes, which were also stored aboard the aircraft during flight. No prior knowledge of the cycle ambiguities was assumed in the processing of the carrier phase data.

The cycle ambiguity estimation results for a typical approach are shown in Figure 3.6 for an arbitrary satellite (PRN 15). This figure shows the integer (cycle ambiguity) estimate error and estimate error covariance history generated during the final forward sweep of the information filter. In a rather straightforward way, this figure demonstrates, through experimental means, both the observability for direct cycle ambiguity resolution using ground-based pseudolites and the actual performance of the cycle ambiguity
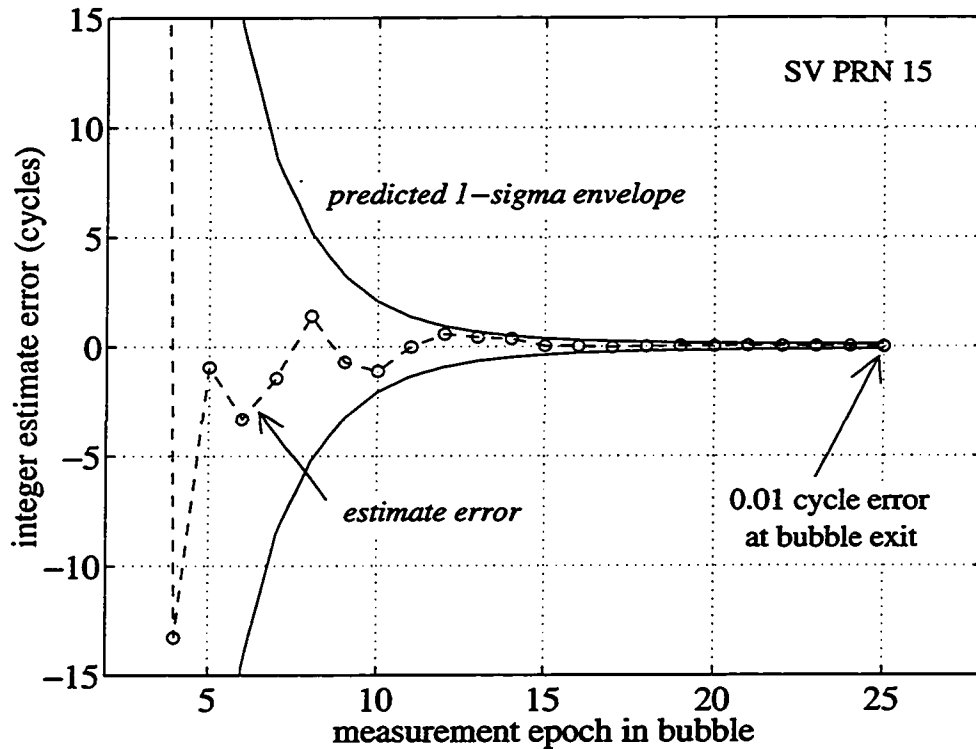
Figure 3.6: Piper Dakota Pseudolite Overflight

resolution algorithm. After only three measurement epochs, the system is overdetermined and integer estimates are available. After eight epochs, the integer for SV 15 is known with a standard deviation of five L1 cycles. Finally, at bubble exit, the integer error standard deviation is 0.1 cycles and the actual integer error is 0.01 cycles. The error in the resulting post-bubble position fixes (see Section 2.7.1) was inferred by the independent survey at tiedown to be less than 2 cm [Cohen, *b*].

3.5.2 Boeing 737 Automatic Landings

The IBLS concept, including its prototype cycle ambiguity resolution algorithms, were also ultimately tested within the context of their actual intended application–

providing reliable real-time navigation output for aircraft automatic landing. Under FAA sponsorship, a United Airlines Boeing 737-300 (Figure 3.7) was equipped with the IBLS breadboard system to verify its capability to provide the aircraft autopilot with navigation input suitable for automatic landing. The flight tests were carried out in October 1994 at the NASA Ames Crows Landing flight test facility in central California.

The IBLS flight and ground hardware used in the tests was substantially identical to that described in Chapter 2. Two ARINC 743 GPS antennas were mounted on the aircraft fuselage (one on the top and one on the belly). Additional minor modifications were necessary in the airborne system in order to provide the aircraft autopilot with an analog GPS navigation output that emulated the nominal ILS glideslope and localizer signal [Cohen, g]. Real-time aircraft attitude for the two-antenna moment arm correction was obtained from one of the 737's two onboard inertial units.

On the ground, two Doppler Marker pseudolites were situated under the approach path 3.5 km from the runway threshold, corresponding to a pseudolite overflight altitude
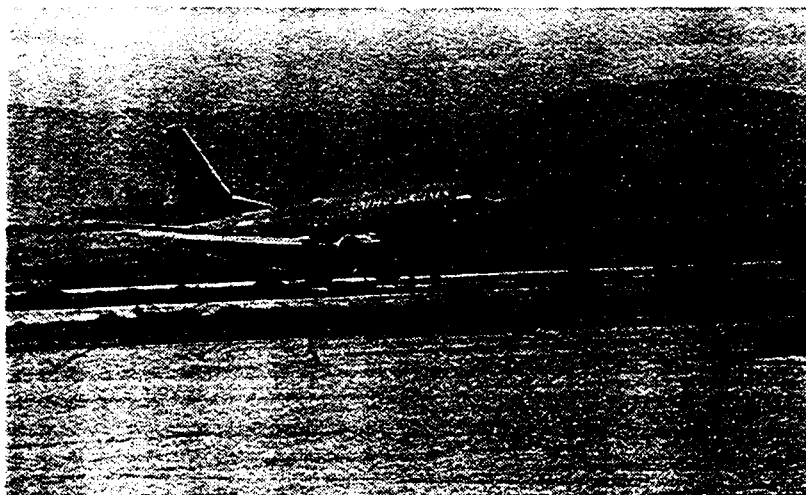
Figure 3.7: United Boeing 737 GPS at Crows Landing

of 600 ft. As was the case in earlier flight tests, the ratio of pseudolite bubble radius to the radius of closest approach was set to approximately 3:1 for both pseudolites. A reference station data-link update rate of 2 Hz was used; this provided the aircraft with 40-50 measurement epochs for cycle ambiguity resolution during a typical pseudolite overflight.

An NdYAG laser tracker at Crows Landing provided range, azimuth, and elevation to a retro-reflector mounted on the 737 nose landing gear and was used to obtain the 'true' aircraft position history during each approach. In actuality, the specified laser range accuracy of ±1 ft (1σ) and elevation and azimuth accuracy of ±0.2 mrad (1σ) provided position fixes considerably less accurate than the expected centimeter-level IBLS performance (suggested by the Piper Dakota post-flight static survey results discussed above).

In total, 111 approaches were executed over three days of flight tests. Because
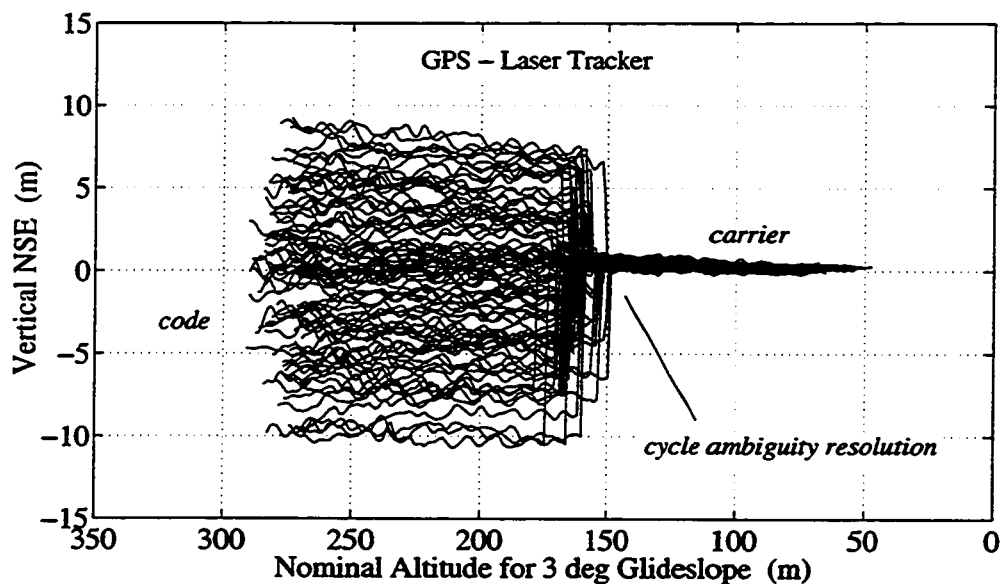
Figure 3.8: GPS-Laser Vertical Position Differences for Boeing 737 Autolands

only four channels of the TANS airborne and ground receivers were allocated for satellite tracking (the remaining two channels were dedicated to tracking the pseudolites), the resulting usable satellite geometries were relatively poor throughout the testing. The effect of this condition, although present during the entirety of each approach, was most notable in positioning prior to pseudolite overflight, where carrier-smoothed code DGPS was employed. During this phase of flight, the resulting vertical positioning error was conspicuously larger that actually achievable (95% error ~ 1-2 m) using a practical DGPS architecture, in which at least seven channels would typically be allocated to satellite tracking.

The IBLS vertical position error (NSE) history—as measured by the laser tracker—for ninety-five representative approaches is shown in Figure 3.8. On each of these approaches, cycle ambiguity resolution was successfully accomplished within approximately one second after bubble exit using the real-time reduced-order algorithm described in Section 3.3.2.3. The exceptional improvement in accuracy exhibited after cycle ambiguity resolution—resulting from the transition to kinematic carrier phase positioning—is clearly evident in the plot. The measured 95% vertical error (ensemble of all of the approaches) after cycle ambiguity resolution was 0.2 m, substantially better than the ILS vertical accuracy specification of 0.6 m. Furthermore, this error statistic agreed quite well with the laser tracker error specifications listed above, lending support to the hypothesis that error of the IBLS navigation output during the approach was smaller than the laser error.

CHAPTER 4

Error Sources and Fault Modes

An understanding of the likely conditions under which the kinematic navigation system operates is the first necessary step toward a quantitative assessment of navigation performance–including integrity. In a general sense, there are two rather broad categories of navigation operation; these are *normal error conditions* and *system failure*. Although it is clearly desirable that the first of these categories applies the large majority of the time, the severe requirements on system integrity demand an understanding (to the extent possible) of the nature and effect of navigation failures.

4.1 Error Sources

Because of the differential nature of kinematic GPS, the largest sources of ranging error are almost entirely eliminated. The remaining errors can be separated into three classes: *receiver-specific errors* such as receiver noise and multipath, *spatial decorrelation errors* which may include troposphere, ionosphere, S/A, and satellite clock error, and *architecture-induced errors* such as latency and moment arm errors. The goal of the subsequent sections is to characterize each of these error sources and develop descriptive models for later use in the evaluation of navigation performance.

4.1.1 Receiver Noise

The magnitude of the white receiver noise for raw carrier phase measurements, although typically very small, will vary depending on the signal power and the phase-lock loop (PLL) bandwidth. To ensure that receiver tracking errors are negligible, the PLL bandwidth must accommodate the anticipated aircraft dynamic response to disturbances such as wind gusts. The TANS receivers (in the IBLS breadboard), for example, use a phase-lock loop bandwidth of approximately 10 Hz to provide, with significant margin, the capability to track the translational dynamics experienced in commercial carrier aircraft.

The standard deviation of single difference carrier receiver noise can be expressed in L1 cycles as

$$\sigma_r = \sqrt{\frac{B_{PLL}}{C/N_0}} \frac{\lambda_{L1}}{2\pi} \tag{4.1}$$

where B_{PLL} is the phase-lock loop bandwidth, C is the carrier signal power, and N_0 is the white noise power spectral density [Carlson]. White, gaussian receiver noise can then be described independently for any given satellite by the following model (at epoch k):

$$v_k^r \sim N(0, \sigma_r^2), \qquad E[v_k^r v_l^r] = 0 \quad (k \neq l). \tag{4.2}$$

To quantify the actual receiver noise for the TANS receiver as a function of signal strength, a controlled laboratory experiment was performed using a GPS signal generator (Wellnavigate GS-100). Figure 4.1 shows the raw phase error (RMS) results of the experiment. The signal-to-noise ratio (SNR), expressed in terms of AMUs (amplitude measurement units), is a direct output from the Trimble receiver proportional to $\sqrt{C/N_0}$. Fifteen minutes of data were collected at signal strengths of 5, 10, 15, and 20 AMUs. Signals below 5 AMUs were not studied because these weak signals are never used in the actual navigation system (since robust phase-lock is not consistently achieved). As
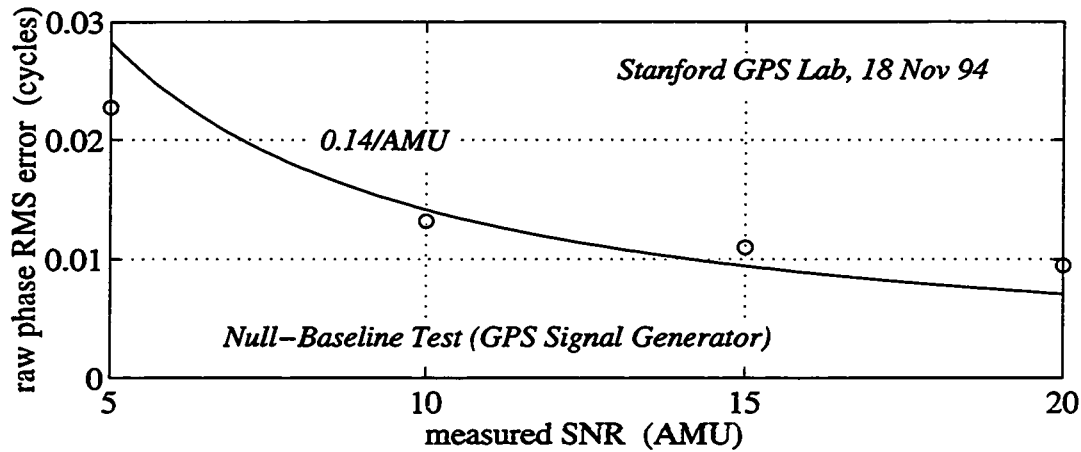
Figure 4.1: Receiver Noise Variation with Signal Strength

expected, the experimental results roughly exhibit the behavior ($\propto 1/\text{AMU}$) predicted by (4.1).

In order to aid in multipath attenuation, the Trimble patch antennas (used in the breadboard navigation system) have lower gain at low elevations. As a result, signal strengths will vary depending on the elevation of the satellite being tracked. Based on the experimental results, however, a simple, conservative error statistic for application in single difference phase receiver noise model (4.2) is $\sigma_r = 5$ mm (0.025 cycles).

4.1.2 Multipath

This error source is the result of unwanted signal reflections received at the user antenna. Because it is highly dependent on the physical environment, satellite geometry, and vehicle motion, carrier multipath is difficult to characterize in a general sense. For illustrative purposes, however, we consider a two hour data set collected across a short static baseline on the HEPL rooftop at Stanford University. Figure 4.2 shows the double difference phase error history due to multipath and receiver noise for the data set. The
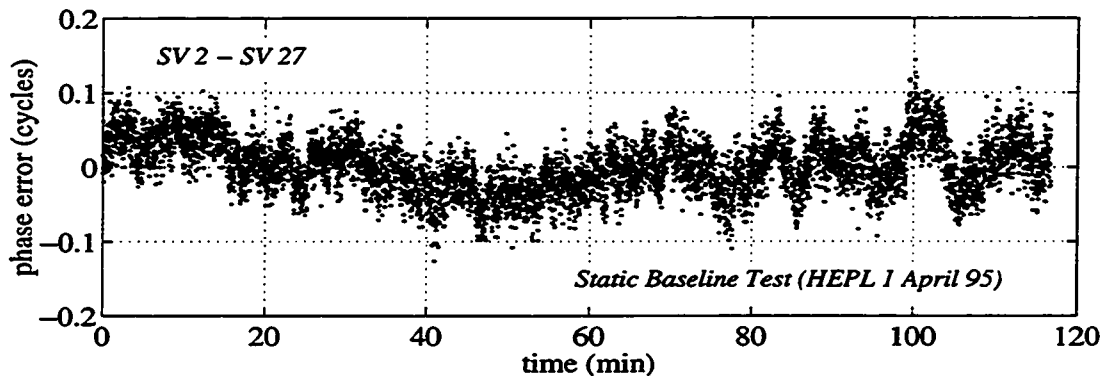
Figure 4.2: Multipath and Receiver Noise (Double Difference)

colored nature of the multipath error in this data, clearly visible in Figure 4.2, is expressed quantitatively in the power spectral density (PSD) plot in Figure 4.3. This plot shows that for the data set under consideration, the PSD is essentially flat for frequencies above 0.05 Hz. In addition, the histogram of double difference error for the entire two hour data set is shown in Figure 4.4. The double difference error roughly exhibits gaussian behavior with an RMS error (standard deviation) of approximately 7 mm (corresponding to a 5 mm single difference error).

Although the single difference RMS error for multipath can be reasonably assumed to be approximately 5 mm [Cohen, a], the temporal character of multipath can vary significantly depending on the specific implementation. For example, the multipath error experienced by an airborne receiver, due to the aircraft motion, would likely have a
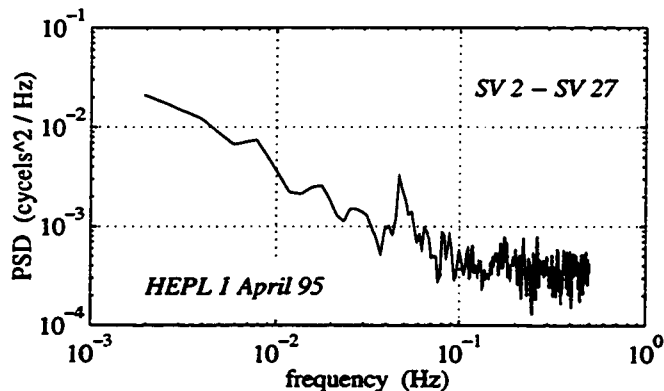
Figure 4.3: Multipath and Receiver Noise PSD
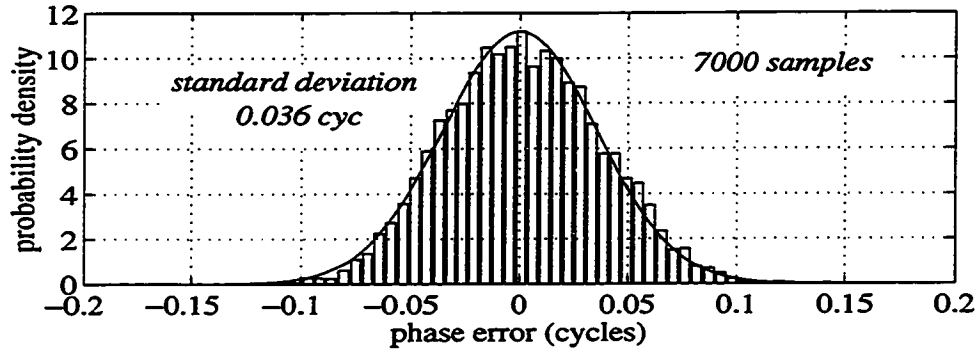
Figure 4.4: Histogram of Multipath and Receiver Noise Errors

significantly flatter (or 'whiter') power spectrum than that in Figure 4.3.

Within the context of the IBLS navigation architecture, multipath errors can be approximately modeled as the sum of three distinct components:

1. *Band-limited noise at low frequencies*—consistent with the timescale of an aircraft approach ($\geq$ 60 sec). The reference station is the primary source of this slowly varying noise component. Because these errors are bias-like over the timescale of a typical approach, they will be estimated in the cycle ambiguity resolution process and implicitly removed from subsequent (post-bubble) position fixes. Thus, this noise component does not contribute to position error and need not be further considered.

2. *Band-limited noise at intermediate frequencies*—approaching the timescale of a pseudolite overflight (~10-60 sec). The reference station multipath is again the primary source of this noise component, but aircraft multipath may also contribute. An abstract but reasonable worst-case model of this multipath component for IBLS is based on the use of two random biases for a given approach. First, for each ranging source i, a bias error selected from

$$b_i^{(1)} \sim N\left(0, \sigma_m^2\right), \qquad E\left[b_i^{(1)} b_j^{(1)}\right] = 0 \quad (i \neq j) \tag{4.3}$$

71

is applied inside the pseudolite bubble. Then after bubble exit, this first bias is removed and a second bias distributed as

$$b_i^{(2)} \sim N\left(0, \sigma_m^2\right), \qquad E\left[b_i^{(2)} b_j^{(2)}\right] = 0 \quad (i \neq j) \qquad (4.4)$$

is applied. The net result of this two-bias model can also be interpreted as a multipath-based cycle ambiguity estimate error of

$$b_i^m \equiv b_i^{(1)} + b_i^{(2)} \sim N\left(0, 2\sigma_m^2\right), \qquad E\left[b_i^m b_j^m\right] = 0 \quad (i \neq j). \qquad (4.5)$$

3. *White noise.* The aircraft is the primary origin of this component, but the reference station will be a source as well. The discrete white multipath noise model can be expressed (at epoch k) as

$$v_k^m \sim N\left(0, \sigma_{mw}^2\right), \qquad E\left[v_k^m v_l^m\right] = 0 \quad (k \neq l). \qquad (4.6)$$

The appropriate values for σ_m and σ_{mw} will, of course, be dependent on the specific implementation issues (such as antenna siting). For the purposes of present study, however, we assume for simplicity that the intermediate frequency multipath noise and white multipath noise components are of roughly equal power. To account for an RMS multipath error of 5 mm (see experimental results above), $\sigma_m = \sigma_{mw} = 3.5$ mm is used.

4.1.3 Troposphere

Signal refraction in the troposphere causes an effective delay in signal transmission that is experienced in equal magnitude in both code phase and carrier phase measurements. This tropospheric delay can be substantially (but not totally) accounted for through the differential nature of kinematic GPS. The majority of the residual differential troposphere delay—which can exceed 1/2 L1 cycle for a 200 ft aircraft altitude—can be removed through troposphere modeling. A thorough discussion of tropospheric effects and models can be found in [Spilker, c] and (with emphasis on precision landing

applications) in [Blomenhofer]. Furthermore, the development of troposphere models suitable for application to the IBLS architecture is detailed by Lawrence in [Lawrence, c].

To assess the residual (post-model) tropospheric error for satellite ranging, we begin with a simplified differential troposphere delay model from [Lawrence, c]

$$\Delta T_i^s = \frac{h(\eta - 1)}{\sin el_i},$$ (4.7)

where h is the aircraft altitude, η is the local index of refraction of the air (typically $\sim 1.0003 \pm 0.00005$) and el_i is the elevation of satellite i. In general, then, the modeling of differential tropospheric delay requires a reasonable knowledge of the aircraft altitude (which is available using differential GPS positioning without a troposphere model) and the local atmospheric characteristics (temperature, pressure, and humidity) that influence the index of refraction of air. The residual tropospheric error (after application of the model) can thus, to first order, be described by

$$\delta(\Delta T_i) = \frac{h \, \delta\eta}{\sin el_i}$$ (4.8)

where $\delta\eta$ is the error in the local index of refraction used in the model. Because the altitude during a pseudolite bubble pass is approximately constant, these residual satellite tropospheric errors are bias-like during pseudolite overflight and will be estimated in the cycle ambiguity resolution process. As the aircraft approaches the runway, however, the altitude and residual troposphere error decrease, resulting in an apparent error in the cycle ambiguity estimates. Modeling $\delta\eta$ to as normally distributed with zero mean and standard deviation of roughly 0.00005 [Spliker, c], and pseudolite overflight altitude of 200 ft, the apparent cycle resolution error $\left(b_i^T \right)$ for satellite i is also normally distributed with zero mean and covariance

73

$$E\left[b_i^T b_j^T\right] = \frac{\sigma_T^2}{\sin el_i \, \sin el_j} \qquad (4.9)$$

where $\sigma_T = 3$ mm.

For pseudolite ranging, the simplified differential troposphere delay model [Lawrence, c] is given by

$$\Delta T_i^p = r_i (\eta - 1) , \qquad (4.10)$$

where r_i is the actual range to the pseudolite. The net effect of the model error $r_i \delta\eta$ can be roughly assessed by observing that the range history for a short pseudolite overflight can be expressed approximately as

$$r_i(t) \approx \left[2 - f_v(t)\right] h , \qquad (4.11)$$

where h is the altitude of pseudolite overflight and f_v is the vertical (radial) basis function in Figure 3.17. As a result, the pseudolite integer error will be on the order of $2h\delta\eta$. However, since the pseudolite integers are not used after bubble exit, this error can be ignored. In addition, the satellite integer estimates will be in error an amount sufficient to produce a vertical *position error* on the order of $h\delta\eta$. For a 200 ft pseudolite overflight altitude, the total vertical RMS position error amounts to roughly only 3 mm.

4.1.4 Ionosphere

The free electrons in the ionosphere affect the nominal satellite signals primarily by causing a transmission delay in the code phase and an equal and opposite advance in the carrier phase. Although these errors are nominally mitigated through the use of a differential architecture, a residual error (due to spatial decorellation of the ionosphere) may exist across longer baselines. Fortunately, these errors are in general significantly smaller than the differential troposphere delay (pre-model) described in the previous section.

A simple approximate model of the absolute ionospheric delay (or advance) given in [Klobuchar] is

$$I_i = I_i^v M(el_i),$$ (4.12)

where I_i^v is the zenith ionospheric delay for satellite i–proportional to the total electron content (TEC) in the ionosphere along the line joining the satellite and the earth center–and M is an obliquity mapping function which accounts for the increase in TEC from its zenith value as the satellite elevation is decreased. A representative obliquity function is defined in [Klobuchar] as

$$M(el_i) \equiv 1 + 2 \left[\frac{96 - el_i}{90} \right]^3,$$ (4.13)

where elevation is expressed in degrees.

Both the vertical ionospheric delay and the spatial gradient of vertical delay are strongly dependent on time-of-day, geomagnetic latitude, season, and solar ultraviolet activity (which varies according to an 11-year cycle). However, based on a representative set of global contours of vertical ionospheric delay given in [Klobuchar] (for a year of high solar activity), 3 mm/km can be used as a somewhat conservative value for the 1-σ spatial decorrelation gradient.

The absolute ionospheric spatial decorrelation will, of course, depend on the distance between the reference station and the aircraft. However, like satellite differential tropospheric delay, these errors are bias-like during pseudolite overflight and will be estimated in the cycle ambiguity resolution process. As the aircraft approaches the runway, decorellation from the estimated biases occurs, resulting in an apparent error in the cycle ambiguity estimates. Consequently, the appropriate distance scale for ionospheric decorellation in the IBLS architecture is the separation between the pseudolites and the runway–approximately 1 km.

The effective cycle resolution error $\left(b_i^I \right)$ for satellite i is simply modeled as

$$b_i^I \sim N\left(0, \frac{\sigma_I^2}{M\left(el_i\right)} \right), \qquad E\left[b_i^I b_j^I \right] = 0 \quad (i \neq j) \tag{4.14}$$

where $\sigma_I = 3$ mm.

4.1.5 Selective Availability and Satellite Clock Errors

Selective availability (S/A) is an error source intentionally induced by the DoD which can be implemented by two methods: dithering the spacecraft clock and/or broadcasting erroneous ephemeris data. The former will clearly have the same effect on user ranging as nominal (non-S/A) satellite clock errors. Estimated statistics of S/A clock dithering component can be found in [Chao], and example mathematical models describing the phenomenon are available in [Chou, Braasch, Lear]. In a differential architecture, the only important effect of either the clock dithering component of S/A or nominal satellite clock errors is manifested in the *latency* of the reference station data used at the aircraft. This effect is discussed in the next section.

At present, it is unclear if the induced ephemeris error method of S/A is being implemented by the DoD, although recent experimental evidence [Braasch, Zumberge] suggests that it is not. For example, recently published statistics of navigation data ephemeris error obtained at the Jet Propulsion Laboratory [Zumberge] show that the broadcast orbit RMS error is less than 5 m. Considering the 1 km baseline from the pseudolites to the runway, the effective contribution to ranging error is less than 0.25 mm RMS and can be neglected.

4.1.6 Latency

Reference measurement latency, an error source present in all differential architectures (with the exception of the Omni Marker concept described in Section 2.4.2),

is related to the age of the measurements collected at the reference station when they are finally processed aboard the aircraft. In general, the latest reference station measurement timetag will lag behind the most recent aircraft receiver measurement timetag. For cycle ambiguity resolution, where the processing is executed on a stored measurement stack, this effect is irrelevant. However, after bubble exit, when punctual position fixes are required, the reference station measurements must be projected forward in time to the latest aircraft receiver timetag. The means for facilitating this projection can vary widely depending on the desired positioning accuracy, available data rate (bandwidth), and allowable computational load.

At present, the IBLS airborne measurement processor applies a simple quadratic fit [Lawrence, a] to a window of reference carrier phase measurements in order to project reference station measurements forward up to 1/2 second (consistent with the current reference data update rate of 2 Hz). The projection error using this technique has been demonstrated by Lawrence [Lawrence, a] to be less than 5 mm RMS. Thus, a simple model of latency error to be applied to carrier phase measurements at any epoch k after pseudolite overflight is given by

$$v_k^L \sim N(0, \sigma_L^2), \qquad E[v_k^L v_l^L] = 0 \quad (k \neq l). \tag{4.15}$$

where $\sigma_L = 5$ mm. The model (4.15) is applied independently to each satellite ranging measurement.

4.1.7 Moment Arm Errors

This class of error, unique to the pseudolite measurements, consists of three components: moment arm calibration error, attitude error, and fuselage flexure. Using optical survey techniques and existing aircraft design drawings/specifications the moment arm calibration between the phase centers of the top and belly antennas can be obtained to

77

within roughly 1 cm RMS in each of the three spatial axes. The effect of this calibration error can be modeled directly as a constant bias error in the post pseudolite position fixes.

Attitude error during pseudolite overflight is a contributing factor when the moment arm is long--as might be the case on large commercial aircraft. For a large moment arm length (fuselage diameter) of 20 ft and an RMS attitude error of 0.05 deg [Cohen, c], the RMS moment arm error is 5 mm. In addition, fuselage flexure during pseudolite overflight, on the order of 2 cm RMS [Ashley], also induces an effective moment arm error.

The combined effect of attitude errors and fuselage flexure can be approximately modeled as a 1.5 cm RMS moment arm calibration bias plus white noise (1.5 cm RMS) applied to the pseudolite phase measurements. The former component is present to account for any low frequency attitude and flexure errors which, during the timescale of a typical pseudolite overflight, would be essentially constant. The latter component can be cast directly in terms of RMS post-bubble position error (or equivalently, a moment arm calibration bias) using equation (3.7). For a typical pseudolite overflight with 40 measurement epochs, the resulting covariance matrix produces an RMS post-bubble position fix error of 7.5 mm.

The root-sum-square (RSS) of these three moment arm error components can be used to provide the total moment arm error contribution, which can be modeled directly as a constant bias error in the post pseudolite position fixes with

$$\delta x_{arm} \sim N\left(0, \sigma_{arm}^2\right). \tag{4.16}$$

where $\sigma_{arm} = 2$ cm. The moment arm calibration error model (4.16) can be applied independently in each of the three spatial axes.

ERROR SOURCE	MEASURED EFFECT (mm RMS)	REMARK
Receiver Noise	5	white
Multipath	5	colored
Troposphere	$3/\sin(el)$	1 km decorellation
Ionosphere	$3 \cdot M(el)$	1 km decorellation
S/A (ephemeris)	< 1	not implemented
Latency (incl. S/A clock)	5	1/2 sec projection
TOTAL	**1–1.5**	elevation dependent

Table 4.1: Satellite Single Difference Phase Error Budget

4.1.8 Error Budget

Table 4.1 briefly summarizes the various error magnitudes for differential satellite carrier phase ranging. Taking the RSS of all error sources, the effective carrier phase ranging error is 1-1.5 cm RMS, depending on the satellite elevation. The contribution of the moment arm errors described above is referenced directly to a 2 cm RMS *position* bias error and is not included in the table.

4.2 Navigation System Failures

In order to quantify navigation *integrity risk*, it is first necessary to identify—to the extent possible—the navigation system failure modes and evaluate their effects on the navigation system output. It is important to note, however, that many types of navigation failure do not contribute to integrity risk. For example, an airborne navigation system power failure or external jamming of the GPS signal result only in a *continuity risk*, because they are always detectable at the aircraft. An integrity risk is incurred if a

navigation system failure, or an unusually large ranging error, attributable to a familiar source such as multipath, causes a large position error that is potentially undetectable by any form of monitoring.

A top-level fault tree illustrating a number of failure classes that represent possible contributions to IBLS integrity risk is shown in Figure 4.5. Allocation of integrity risk begins by considering separately the three navigation segments: airborne, ground, and space. The various failure classes for each of these segments are described in the remaining sections of this chapter. The means for *detection* of these failure classes will be addressed in detail in Chapter 5.
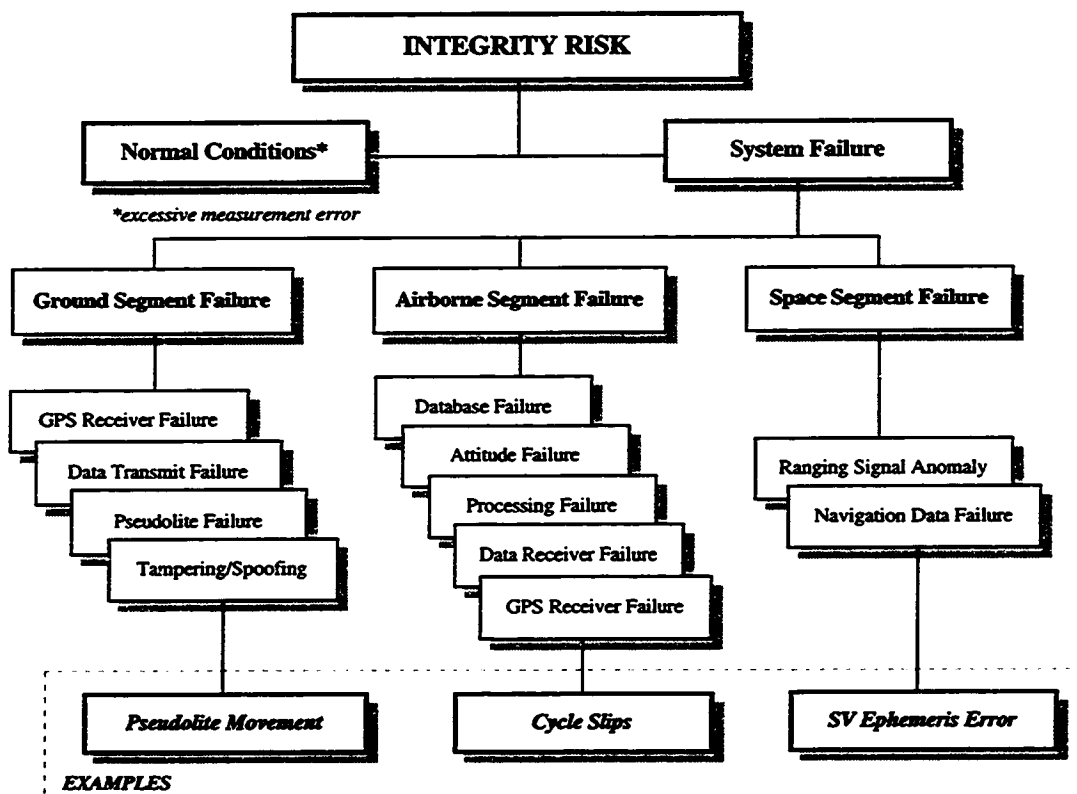
Figure 4.5: Top-Level IBLS Navigation System Fault Tree

In addition, three specific failure modes of diverse origin and character—one from each of the segments—is selected for more detailed analysis. Mathematical models for these failures will be developed for the quantitative evaluation of navigation performance through Monte Carlo simulation in Chapter 6.

4.2.1 Airborne Segment Failures

Airborne segment failure for IBLS can be broadly cast into the five functional categories shown in Figure 4.5: database, attitude, processing, data receiver, and GPS receiver failures. Errors in the onboard *database* can cause a variety of effects, some of which are benign and some potentially hazardous. Within the context of the IBLS architecture, this database would include the ground survey parameters identifying the locations of the reference station and pseudolites with respect to the runway as well as the aircraft moment arm calibration. It is noteworthy that errors in the ground survey portion database are mathematically equivalent to physical manipulation of the actual hardware on the ground. This particular failure will be discussed in more detail in Section 4.2.2. It is possible that validation by independent flight test can ensure overall database error probabilities as low as 10^{-8} [Davis].

Attitude system integrity is not only crucial for IBLS applications, but also for other precision landing systems (to correctly project antenna position to the landing gear) as well as for other phases of flight. The *navigation processing* function uses attitude, the airborne database, and GPS receiver outputs to generate aircraft position and perform autonomous integrity verification. A processing interruption (such as a real-time software failure) that results in the absence of position fix output generally contributes only to continuity risk. However, a failure in the integrity verification function would contribute directly toward integrity risk. Although the likelihood of such a failure is very difficult to

81

quantify, the rigorous procedures required for flight critical software certification [DO-178B] are designed to ensure that these events are extremely rare.

Failures in the airborne *GPS or datalink receivers* may result in either the lack of data or erroneous data (pseudorange measurements, carrier phase measurements or satellite ephemeris) being forwarded to the navigation processor. The former case, of course, corresponds to a continuity risk; the latter, to integrity risk. For a kinematic GPS system such as IBLS, the most commonly expressed concern in terms of integrity risk failures is associated with *cycle slip* events within the GPS receiver.

Cycle Slips

In the event of momentarily low signal strength on a given satellite (or pseudolite) signal, a GPS receiver channel may lose phase-lock on the satellite signal and, instead, track the noise component of the signal. When the satellite signal is reacquired, the elapsed number of integer cycles are not necessarily properly accounted for; or equivalently, the cycle ambiguity has changed. Such an event is known as a *cycle slip*. Figure 4.6 illustrates the effect on double difference phase error due to a slip of one cycle which was induced numerically in the post-processing of previously collected static
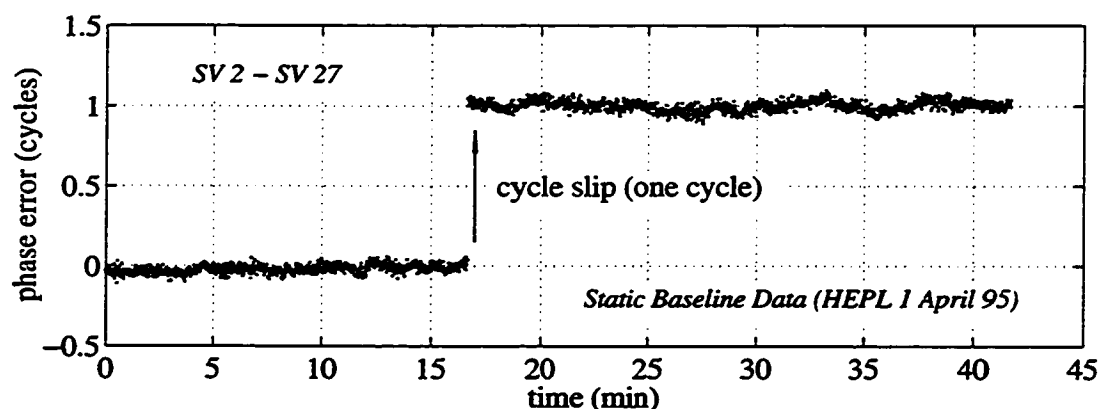
Figure 4.6: Cycle Slip of One Cycle (Double Difference Phase)

baseline data. Although the discussion has thus far deliberately been centered on failures of airborne origin, it is important to note that once the single difference phase observable is constructed in the navigation processor, a cycle slip in the reference receiver will be indistinguishable from an airborne receiver cycle slip.

In general, the probability of a cycle slip event increases with increasing phase-lock loop (PLL) bandwidth, increasing exposure time, and decreasing signal strength. As a first layer of protection against cycle slips, the low signal strength and weak phase-lock conditions under which cycle slips are likely to occur are nominally identified at the signal processing (phase-lock loop) level. Most cycle slips are detected in this manner; however, the remaining probability of an undetected cycle slip is not possible to compute analytically. The application of simplified first-order phase-lock loop theory [Spilker, a] suggests that the probability of a cycle slip occurrence is extremely rare—even for the lowest signal strengths actually encountered in practice. Furthermore, over the entire history of IBLS breadboard flight testing (nearly one thousand approaches), no cycle slip event, undetected at the signal processing level, has ever occurred during the course of an aircraft approach.

Based on accumulated results of null-baseline laboratory testing, cycle slips (undetected at the signal processing level) can be assumed to occur at a rate of roughly one in 10^5 pseudolite overflights (of approximately 10 sec duration). Cycle slips can further be assumed to occur only for satellites and pseudolites whose elevations are between -25 deg and +25 deg; this is consistent with the existing IBLS implementation, which uses Trimble Microstrip Patch antennas having approximately 5 dB lower gain at 10 deg elevation than at zenith. Given the occurrence of such a cycle slip event (CS) on a satellite or pseudolite during the approach, both null-baseline laboratory tests and first-order phase lock loop theory [Spilker, a] suggest that the likelihood of slips of magnitude larger than one decreases as slip magnitude increases. The following geometric sequence

has been chosen as a simple approximate model for this effect for use in further analysis in Chapter 6:

$$P(CS \; size = \ell | CS) \; = \; \left\{ (1/2)^\ell, \quad \ell = 1,2,3, \cdots \right\}. \tag{4.17}$$

4.2.2 Ground Segment Failures

A ground segment failure in the IBLS architecture, as indicated in Figure 4.5, can be classified as either a GPS receiver, pseudolite, or data transmission failure, or as tampering or spoofing of the ground-based navigation function. The discussion of airborne GPS receiver failures above, including cycle slips, is applicable to the *reference GPS receiver* as well. The airborne datalink receiver, however, is replaced with a ground-based data transmitter. Analogous to the aircraft case, failures in the ground GPS receiver or data transmitter may result in either the lack of data (continuity risk) or erroneous data (integrity risk) at the aircraft. As will be discussed in Chapter 5, the latter condition can nominally be detected through the use of a ground-based integrity monitor.

A *pseudolite failure* is manifested in either the transmission of an anomalous signal or the lack of signal output altogether. The continuity threat due to the latter condition can, of course, be alleviated through sufficient pseudolite hardware redundancy. Depending on the specific nature of the anomalous transmission, the former condition can result in either integrity or continuity risk. A continuity risk will occur if the abnormal signal is untrackable by the reference station or aircraft GPS receivers. If the signal is trackable, an integrity risk may occur, although the differential nature of kinematic GPS will, in general, serve to mitigate the effect of the failure on position estimation.

It has recently been suggested that a DGPS-based landing system may be susceptible to *intentional tampering or spoofing* [Klass]. Tampering and spoofing can take on a number of forms. For example, a reasonably sophisticated spoofer could

possibly alter the carrier phase measurements being transmitted by the datalink to the aircraft or even independently broadcast a separate, false satellite or pseudolite signal from an alternate site. Such spoofing techniques could potentially produce erroneous and possibly hazardous position estimates at the aircraft. It is perhaps more likely, however, that intentional tampering or spoofing would take the form of physical manipulation of the ground hardware. The effect of physical movement of the reference receive antenna or a pseudolite transmit antenna is equivalent to that of the spoofed datalink and pseudolite broadcast scenarios, respectively. Furthermore, unintentional errors in antenna siting, perhaps after maintenance work, and ground survey errors in the airborne database (as noted in the previous section) are also covered by this scenario.

Ground Antenna Movement/Spoofing

In the IBLS architecture, a displacement (Δx_R) of reference station receive antenna to an unsurveyed location will, in general, not cause aircraft position error. For an arbitrary spacecraft i, the effective phase error $e_i^T \Delta x_R$ will be estimated as a bias in the cycle ambiguity estimate. Because the satellite motion is very slow, the effective reference phase error is essentially constant during the course of the approach and is removed when the cycle ambiguity estimate is applied to perform a post-bubble position fix.

Movement of a pseudolite, however, will result in absolute position error. Figure 4.7, for example, illustrates that the effect of crosstrack pseudolite movement is an error in the crosstrack-vertical absolute position. Although the sensitivity of vertical error (the most critical component) to pseudolite movement can be lowered somewhat by reducing the nominal separation distance between the pseudolites, pseudolite movement must be considered an integrity risk. (Specific architecture implementations designed to mitigate this risk will be discussed in Chapter 5.)

absolute position error

(intrack view)

PL 1   PL 2

actual pseudolite locations ——————————   —————— nominal PL 2 location
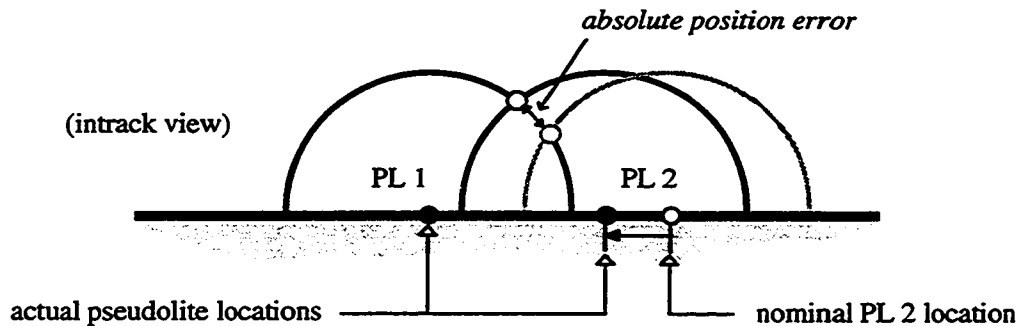
Figure 4.7: Effect of Crosstrack Pseudolite Movement

It is noteworthy that mis-siting or movement of existing ILS antennas (or spoofing of the ILS signal-in-space) can also lead to potentially hazardous conditions. However, as such an event has apparently not been observed in a total of more than 200 million approaches since 1959 [Davis] it is not unreasonable to assume that such occurrences will also be relatively rare within the context of the IBLS architecture. Specifically, for the purposes of developing a simple mathematical model of mis-siting failure (to be applied in Chapter 6), a somewhat conservative prior probability of mis-siting error of 10^{-6} per approach is assumed. Given that a pseudolite siting failure has occurred, the displacement probability density is assumed to be gaussian, with zero mean, and independent in all three spatial dimensions. Consistent with the notion of maintenance-related mis-siting, the standard deviation used for horizontal error is set at 1 m. As large vertical siting errors are more difficult to physically achieve, the associated standard deviation is set at 0.5 m.

4.2.3 Space Segment Failures

A satellite failure represents an integrity risk during the transitional period between failure inception and detection at GPS Master Control. At this point, the affected spacecraft is declared unhealthy and the situation is communicated to GPS users

through the GPS navigation data message. Consequently, if the received data is read punctually at the receiver, integrity risk is no longer present, although navigation availability (or continuity, if a spacecraft is declared unhealthy while an aircraft approach is underway) may be affected. Estimates of satellite availability, including down-time due to hard failures and stationkeeping maneuvers, are derived in [Durand, Phlong].

As viewed from the perspective of the GPS user, satellite failures fall into two functional categories: ranging signal failures and navigation data failures. A failure on the spacecraft's onboard clock is the predicted most common source of *ranging anomalies* [Gower]. The effects of such ranging errors are nominally eliminated through the use of differential GPS architectures. However, an unusual type of satellite range error was recently exhibited by SV PRN 19 [Nordwall]. The symptom was a residual pseudorange bias—not removed by DGPS—when non-identical receivers were used at the reference station and aircraft. Clearly, such a failure represents an integrity risk. Kinematic carrier phase architectures (such as IBLS), however, are not affected by this failure mode.

The class of *navigation data failures* which represent a potential integrity risk to both code and carrier-based architectures is ephemeris error, although rather large errors (in the hundreds of meters) are needed to produce a noticeable inaccuracy in positioning.

Satellite Ephemeris Errors

Among the possible origins of integrity-threatening ephemeris errors are S/A ephemeris errors of exceptionally large magnitude, orbit determination or data upload error at the MCS, or a data receive or transmission error at the spacecraft. For the differential GPS user, it will be shown in Chapter 5 that a spacecraft ephemeris error can be mathematically expressed as an equivalent ranging error that scales linearly with the displacement between the aircraft and reference station (x) according to $\delta e_i^T x$, where δe_i

is the error in the line-of-sight vector to satellite i. In the IBLS application, the average ranging error during the bubble pass $\overline{\delta e_i^T x}$ is calibrated and removed through the cycle ambiguity resolution process. For example, consider an ephemeris error on a high elevation satellite as illustrated in Figure 4.8. In this case, the estimated kinematic trajectory is rotated about the approximate point of pseudolite overflight, and the deviation of the estimated (false) trajectory from the true trajectory increases as the aircraft approaches the runway.

Given the degree of redundancy and fault-tolerant design incorporated into the GPS Operational Control Segment in general [Francisco] and the active monitoring of broadcast navigation data by the MCS, in particular, a major broadcast ephemeris error is extremely unlikely. Furthermore, the envisioned implementation of the Wide Area Augmentation System (WAAS) [Kee, Enge], will provide an additional layer of detection against latent ephemeris failures. Thus, for modeling purposes, a frequency of one latent ephemeris failure in 10^6 aircraft approaches will be assumed. The conditional probability density of satellite position error, given that a satellite ephemeris error has occurred, is modeled as gaussian, independent, and identically distributed with zero mean and 100 km standard deviation in each of the three spatial dimensions.
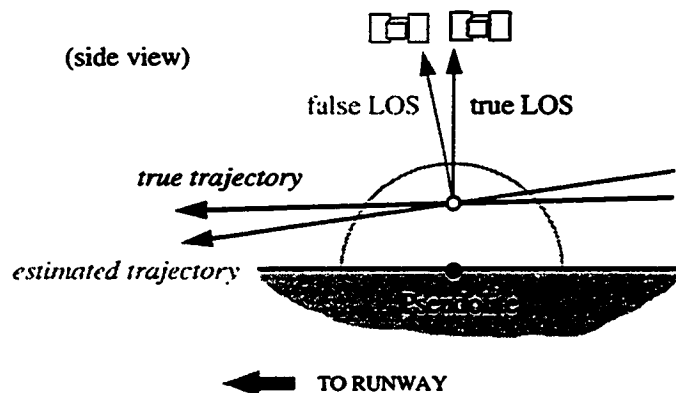
Figure 4.8: Effect of Satellite Ephemeris Error

CHAPTER 5

Fault Detection and Isolation

Although failures in a well-designed navigation system will be relatively rare, the capability for *detection* and *isolation* of failures is necessary in order to satisfy the stringent requirements for navigation system integrity and continuity during precision approach and landing (see Table 1.1). The detection function generates an alarm in the event of a hazardous navigation failure and is, therefore, of paramount importance in the reduction of integrity risk. The isolation function provides the potential means for specifically identifying and removing a detected system failure, thereby reducing continuity risk.

A top-level functional block diagram showing the roles of fault detection and isolation is given in Figure 5.1. In the large majority of approaches, of course, a navigation failure will not occur. The existence of a detection function, however, opens up the possibility of *false alarm*. In this case, a failure has not occurred but an alarm has been issued nonetheless, resulting in a loss of continuity. In the case of an actual navigation system failure (or rare-event large normal condition error), the failure will generally be detected, and isolation will be attempted. If the detected failure is clearly distinguishable from other failure modes, it can often be successfully isolated and removed. If the correct failure is removed, integrity and continuity are maintained.

89
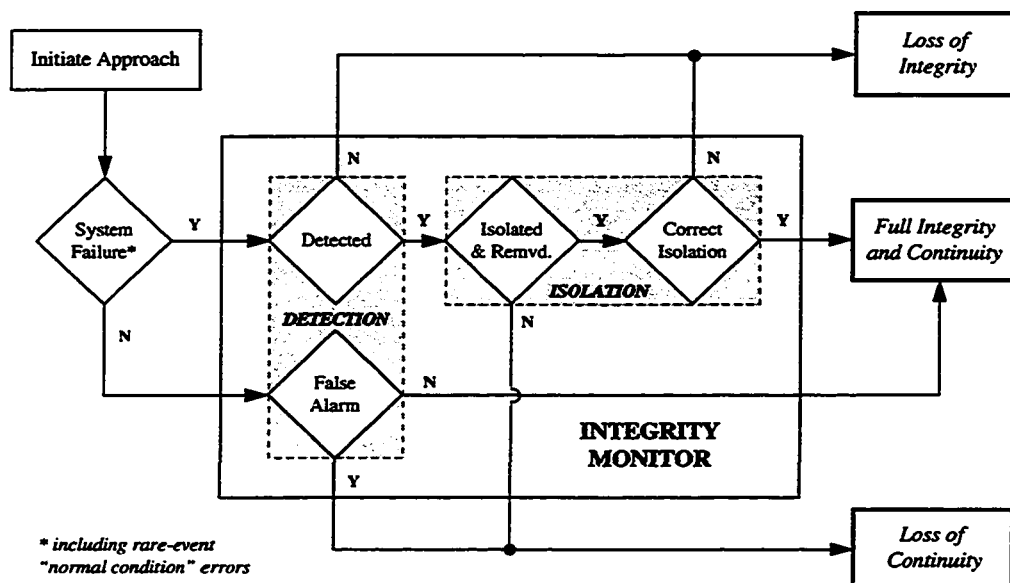
Figure 5.1: Functional Block Diagram of an Integrity Monitor

However, if the failure is not detected, or if detection does occur but the wrong failure is removed, navigation integrity is lost. A final possibility occurs when a failure has been detected, but isolation is not deemed possible; in this event, continuity is lost. For simplicity, we define the term *integrity monitor* as any given system which provides the means for fault detection and isolation.

5.1 Methods of Integrity Monitoring

All methods of integrity monitoring rely, in one way or another, on checking the consistency of redundant information. In the existing ILS application, the integrity monitoring function is delegated to two separate systems: *ground monitoring* and *redundant airborne sensors*. The ground monitor is present to detect failures in the ILS signal-in-space (glideslope and localizer). In the event of a detected failure during a Category III approach, a backup glideslope or localizer transmitter is engaged within two

seconds [ICAO]. Redundant sensors onboard the aircraft, often implemented as two or more independent tracks of ILS receiver/inertial reference unit (IRU) pairs, provide the means for the detection of airborne receiver or IRU failures. If three or more parallel sensor tracks are implemented, the faulty track can be removed, and the continuity of both the positioning and fault detection functions preserved.

The integrity monitoring concepts established for ILS (ground monitoring and redundant airborne sensors) can also be applied to GPS-based precision landing architectures. This is especially true because, for the foreseeable future, the implementation of GPS for precision approach and landing will likely be in the form of an ILS-lookalike architecture [Miller]. In such a system, the airborne GPS sensor will function simply to replace those inputs (glideslope and localizer) formerly supplied by the ILS. (In fact, such an architecture was used in the Boeing 737 flight trials described in Section 3.5.2.) Clearly, the existing level of airborne sensor redundancy present with ILS is preserved for the GPS application as well. Thus, a majority of airborne segment navigation failures will continue to be detectable and isolatable by monitoring the consistency of the parallel airborne sensor tracks. In addition, the monitoring of inertial attitude—needed in IBLS for the antenna moment arm correction—can be achieved without impact on the existing aircraft autopilot by simply cross-checking redundant inertial attitude inputs at a given airborne GPS receiver.

The difference in integrity risk between parallel track implementation of the new GPS and existing ILS hardware will, of course, be due to the differing failure probabilities and characteristics of airborne GPS and ILS receivers. For example, rare-event multipath errors may represent an integrity risk for code differential GPS architectures if all airborne receivers respond in a similar fashion to the multipath. This situation may not be unrealistic given the common geometry-based origin of the multipath for all the receivers.

The ground monitor concept of ILS will likely also have a role with GPS-based landing systems in the detection and isolation of ground and space segment failures. The simplest conceivable GPS ground monitoring architecture uses a GPS receiver and data receiver. The monitor GPS antenna is placed at a precisely surveyed site so that the resulting differential position fixes and/or range measurements at the monitor site (based on the broadcast reference data) can be directly compared to the known quantities. In the event of a discrepancy, the primary reference station is taken off line and the backup is switched on. A number of more sophisticated and robust GPS ground monitor architectures are described in [Braff, Markin, LIP].

Although the effects of ranging errors originating at the spacecraft are usually eliminated due to the very nature of differential positioning, and a majority of other space and ground segment failures can be detected by ground monitoring, integrity risk is still present. For code-based architectures, one example of residual integrity risk may be found in the SV 19 phenomenon described in Section 4.2.3. In this case, ground monitoring may or may not be able to detect the ranging error, depending on the actual receivers used at the reference station, monitor station, and aircraft. In addition, spoofing of the reference datalink or GPS spacecraft signals may elude detection at the monitor site if the spoofed signal is specifically directed only at an incoming aircraft.

IBLS and Receiver Autonomous Integrity Monitoring

The Integrity Beacon Landing System, through its use of carrier phase measurements and ground based pseudolites, provides the *built-in* capacity for high navigation integrity. The use of carrier phase measurements, for example, excludes rare-event code multipath and SV 19 failures as integrity risks. Of greater significance, however, is the intrinsic capability for comprehensive fault detection using *receiver autonomous integrity monitoring* (RAIM). The great precision of carrier phase and the redundant ranging measurements from the pseudolites provide the leverage for an

INTEGRITY MONITOR	DETECTS FAILURES IN
Ground Monitor	Ground and Space Segments
Redundant Airborne Sensors	Airborne Segment
RAIM	*Ground, Airborne, and Space Segments*

Table 5.1: Methods of Integrity Monitoring

unprecedented level of RAIM performance. These concepts will be discussed in detail in the sections that follow.

Although ground monitoring and redundant airborne sensors can be implemented, respectively, as a preliminary defense against ground/space segment failures and airborne segment failures, the transition from a familiar (ILS) to a new (GPS) navigation system for aircraft precision landing requires a careful and comprehensive approach toward integrity monitoring. In this respect, highly effective carrier phase RAIM is of paramount importance. RAIM, like redundant airborne sensors, enables the final integrity decision to be made at the aircraft. However, unlike redundant airborne sensors, RAIM provides the capability to detect failures originating in all three navigation system segments (see Table 5.1). A number of specific benefits that can be realized through the application of RAIM will be discussed in Section 5.2.4.2. In a more general sense, however, carrier phase RAIM can serve a principal layer of defense against unknown new failure types resulting from the transition from ILS to GPS.

Within this context, a highly effective integrity monitoring architecture consistent with the expected ILS-lookalike implementation of IBLS might include:

1. *Ground monitoring*, as the preliminary layer of protection against ground and space segment failures.

2. *Redundant airborne sensors*, as the preliminary layer of protection against airborne segment failures.
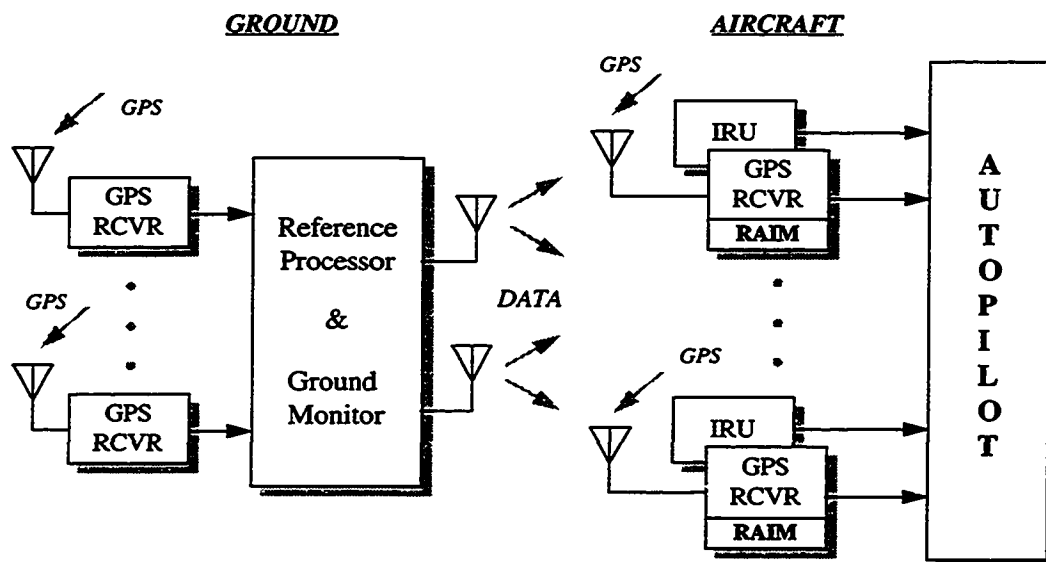
Figure 5.2: Example Integrity Monitoring Architecture

3. *Carrier Phase RAIM*, as the main layer of protection against failures in all segments.

Figure 5.2 illustrates a conceptual integrity monitoring architecture incorporating all three of these elements; in this figure, the airborne GPS receivers are assumed to have the embedded capability for measurement processing. It is also possible, however, that the broad fault detection capability of carrier phase RAIM can be used, in future architectures, to relax the requirements for ground-based monitoring and redundant airborne sensors.

5.2 RAIM-Based Fault Detection

Integrity risk (I) and continuity risk (C) —as may be easily verified by inspection of Figure 5.1—can be explicitly defined in mathematical terms as

$$I \equiv P(MD) + P(MI|DF)P(DF) \qquad (5.1)$$

$$C \equiv P(FA) + P(NI|DF)P(DF) \qquad (5.2)$$

where the following event definitions have been made:

MD is a missed detection,

FA is a false alarm,

DF is a detected failure,

MI is an incorrect isolation (or *mis-isolation*), and

NI represents the case where a failure has been detected but isolation is not deemed possible.

Because it is necessary that integrity risk be much lower than continuity risk (see Table 1.1), the detection function can be considered to be the primary duty of the integrity monitor. Thus, it is appropriate to address the detection problem first; the discussion of the fault isolation is deliberately postponed until Section 5.3. Fault detection can be classified as a composite binary hypothesis problem [Van Trees, Daly]–the two possible states of the navigation system being, of course, normal operation and system failure. As such, only the probabilities of missed detection and false alarm are relevant in this discussion.

RAIM-based fault detection is based on verifying the consistency of redundant GPS measurements. Although a number of related methods of RAIM implementation have been proposed [Brown, a, b], the use of the norm of the least squares residual vector as a test statistic is the most direct.

5.2.1 The Least Squares Residual

Consider the generalized linear observation equation

$$z = (H - \delta H)u + \delta z \qquad (5.3)$$

where u is the $m \times 1$ vector of parameters to be estimated, z is the $n \times 1$ $(n > m)$ measurement vector, H is the $n \times m$ observation matrix, and δz and δH are the unknown errors in the measurement vector and observation matrix (e.g. satellite ephemeris errors), respectively. The general relation (5.3) is applicable to both post-bubble kinematic positioning discussed in Chapter 2 and the cycle ambiguity resolution problem of Chapter 3. In the latter case, for example, the measurement vector z corresponds to the vector of stacked single-difference phase measurements collected during pseudolite overflight as indicated in equation (3.31). In general, the desired position vector, x (3×1), is a vector element of u:

$$u = \begin{bmatrix} \vdots \\ x \\ \vdots \end{bmatrix}.$$ (5.4)

For the specific case of cycle ambiguity estimation, the vector x can be most conveniently interpreted as the position fix at bubble exit.

Assuming, for the moment, independent and identically distributed (i.i.d.) measurements, the maximum likelihood, least squares solution to (5.3) is

$$\hat{u} = H^+ z,$$ (5.5)

where

$$H^+ \equiv \left(H^T H \right)^{-1} H^T.$$ (5.6)

and $\hat{u}$ is the estimate of the true state u. The least squares estimate error is defined as

$$\delta u \equiv \hat{u} - u = \begin{bmatrix} \vdots \\ \hat{x} \\ \vdots \end{bmatrix} - \begin{bmatrix} \vdots \\ x \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \delta x \\ \vdots \end{bmatrix}$$ (5.7)

and is given by the relation

$$\delta u = \begin{bmatrix} \vdots \\ \delta x \\ \vdots \end{bmatrix} = H^+ \left(\delta z - \delta H u \right). \tag{5.8}$$

The position error, of course, can also be extracted explicitly through the linear transformation

$$\delta x = C \delta u, \tag{5.9}$$

where, in the general case,

$$C \equiv \begin{bmatrix} 0 & \cdots & 0 & I_3 & 0 & \cdots & 0 \end{bmatrix}. \tag{5.10}$$

The least squares residual vector r is a measure of the consistency of the measurement vector z subject to the observation matrix H, and is expressed simply as

$$r \equiv z - H\hat{u}. \tag{5.11}$$

Using equations (5.3) and (5.5) the residual vector definition (5.11) can also be expressed directly in terms of measurement and observation matrix errors as

$$r = \left(I - HH^+ \right) \left(\delta z - \delta H u \right). \tag{5.12}$$

For the case of correlated (or non-i.i.d.) measurement errors with covariance matrix W, relations equivalent to (5.8) and (5.12) can be easily obtained if both sides of (5.3) are simply pre-multiplied by a 'whitening' matrix $W^{-\frac{1}{2}}$ [Walter]. This situation is particularly appropriate to the case of post-bubble position estimation given prior estimates of the cycle ambiguities (as described in Section 2.7). For simplicity of notation, however, the mathematical development will continue, without loss of generality, for the i.i.d. case.

The simplest form of RAIM fault detection is based on the use of the magnitude of the residual vector $\|r\|$ as a statistical indicator of possible navigation failure [Sturza,

Parkinson, a]. Under normal error conditions (NC), δH is negligible and δz is gaussian and i.i.d. with zero mean and standard deviation σ_z. In this case, it can be shown (Appendix C) that the quantity $\|r\|^2/\sigma_z^2$ is a χ^2 distributed random variable with $n - m$ degrees of freedom (DOF),

$$\frac{\|r\|^2}{\sigma_z^2} \sim \chi^2(n-m) , \tag{5.13}$$

where n is the number of measurements and m is the number of unknowns $(n > m)$. A residual threshold R can be set analytically using (5.13) to achieve any desired probability of false alarm under normal error conditions (NC),

$$P(FA|NC) = P(\|r\| > R|NC) = \frac{1}{2^{\frac{n-m}{2}}\Gamma(\frac{n-m}{2})} \int_{\frac{R^2}{\sigma_z^2}}^{\infty} s^{\frac{n-m}{2}-1} e^{-\frac{s}{2}} ds , \tag{5.14}$$

where the integral in (5.14) is the *incomplete gamma function*. In Figure 5.3, the generalized residual threshold R/σ_z is plotted against $n - m$ (DOF) for false alarm probabilities of 10^{-2}, 10^{-5}, and 10^{-8}. Although $n - m$ is typically less than 5 for kinematic (post-bubble) positioning, during cycle ambiguity estimation (pseudolite overflight) a large number of redundant measurements–often greater than 100–will be present. *Correlated* measurement noise during pseudolite overflight will, in general, cause biases in the cycle ambiguity estimates (see Section 4.1.2), but will also result in false alarm rates somewhat lower than those analytically predicted [Masson].

In the event that the position error
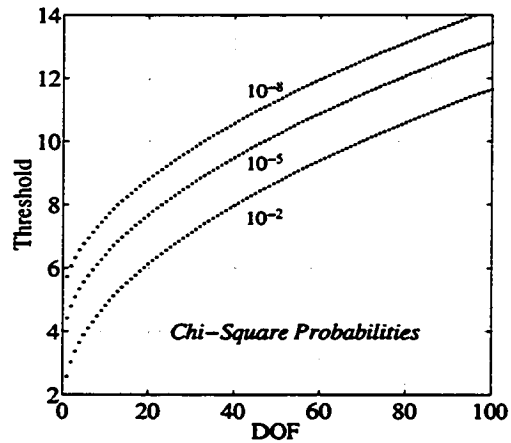
Figure 5.3: RAIM Detection Thresholds

exceeds a predefined *protection radius*, a, but $\|r\| < R$, a missed detection has occurred. Thus the probability of missed detection can be defined explicitly as

$$P(MD) = P(\|r\| < R, \|\delta x\| > a).$$ (5.15)

It can be shown (Appendix C) that, in a deterministic sense, $\|r\|$ and $\|\delta x\|$ are independent, since δz can be chosen to give a very large $\|\delta x\|$ and a very small (even zero) $\|r\|$. This is true because $\|r\|$ and $\|\delta x\|$ result from the projections of $\delta z \in R^n$ onto two orthogonal subspaces of R^n. In fact, this independence also applies in a statistical sense if the measurement error truly exhibits the assumed i.i.d. gaussian characteristics associated with the normal error conditions case. However, given the possibility of failure in the navigation system, this simple i.i.d. error model does not apply, and in general, a correlation between $\|r\|$ and $\|\delta x\|$ will exist. It is precisely the degree of this correlation, within the context of the envisioned navigation system failure modes, that must be quantified to demonstrate the integrity
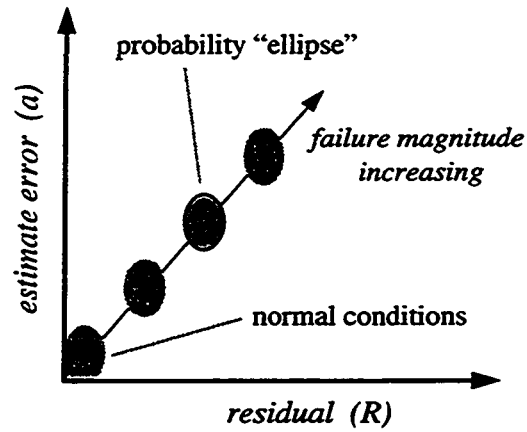
Figure 5.4 (a): Estimate Error vs. Residual
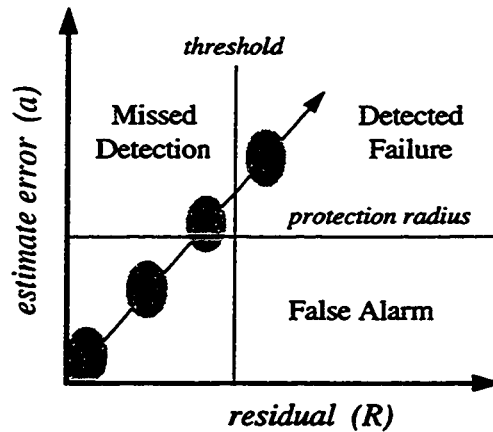
Figure 5.4 (b): Basic RAIM Algorithm

monitoring capability of RAIM-based fault detection.

Figure 5.4 (a) is a conceptual plot of position error versus the least squares residual. The (roughly ellipse-like) probability dispersion nearest the origin represents the case of i.i.d. gaussian normal condition errors such as multipath and receiver noise. We now define a general $(n \times 1)$ failure mode vector d as

$$\delta z - \delta H u \sim N\left(d, \sigma_z^2 I_n\right), \tag{5.16}$$

where σ_z^2 is the normal measurement error variance. Because both the mean residual (5.12) and mean position error (5.8) magnitudes will scale proportionately with the failure magnitude $\|d\|$, the normal condition error ellipse will slide up the failure mode axis whose slope is given by $\| CH^+d \| / \| (I - HH^+)d \|$. In Figure 5.4 (b), a horizontal line constraint is drawn to represent the protection radius a. Note that it is possible, for small failure magnitudes, that the accuracy specification not be breached. Also shown in this figure is the residual threshold R.

The resulting RAIM fault detection algorithm is a simple one: Check the residual statistic to see if it is larger than the threshold. If so, a system failure is declared. Given this simple algorithm, four outcomes are possible. Under normal operation, the position error does not exceed the protection radius and the residual is smaller than the threshold. If the
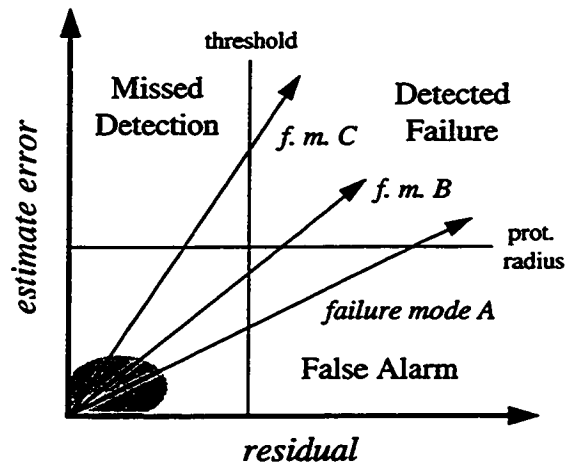
Figure 5.5: Integrity Risk and RAIM

position error does not exceed the protection radius, but the residual is larger than the threshold, a false alarm has occurred. When both the protection radius and residual threshold have been breached, a detected failure has occurred. Finally, a missed detection happens when the position error is larger than the protection radius, but the residual is smaller than the threshold.

In the general case, of course, more than one failure mode exists. This situation is conceptually illustrated in Figure 5.5, where a family of three fault modes are shown. It is seen in this figure that the characteristic slopes of failure modes A and B are low enough to ensure the RAIM detectability of hazardously large fault magnitudes. Failure mode C, however, represents an integrity risk in that the missed detection region is penetrated. In this case, for failures of intermediate magnitude, the estimate error exceeds the protection radius, but the residual is lower than the threshold.
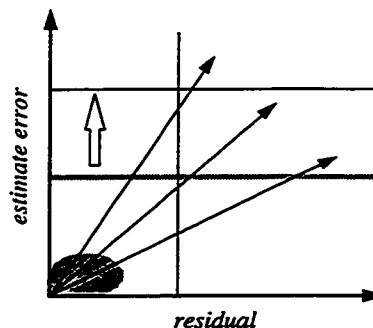
(a) Integrity and Accuracy
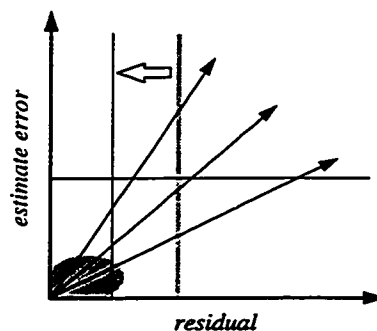
(b) Integrity and Continuity
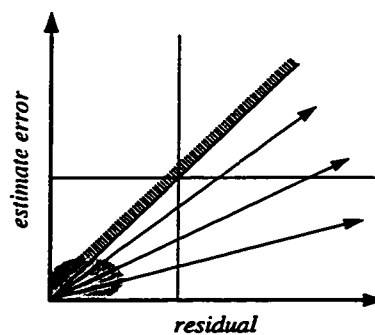
(c) Integrity and Availability

Figure 5.6: Mitigating Integrity Risk

5.2.2 Mitigating Integrity Risk

It is important to note explicitly that integrity risk can always be reduced at the expense of any of the other navigation performance parameters (accuracy, continuity, and availability). For example, *accuracy* can affect integrity risk in two ways. First, if the *measurement precision* is increased (σ_z is lowered), then for the same probability of false alarm, the residual threshold can be reduced. The result is a narrowing of the missed detection region. Second, the *protection radius* can be relaxed directly to mitigate the integrity risk associated with failure mode C. This situation is illustrated in Figure 5.6 (a).

As depicted in Figure 5.6 (b), *continuity* of the navigation function can also be traded for integrity. For a given measurement precision, a reduction in the detection threshold will result in a narrower missed detection region but will also cause an increase in false alarm rate—thereby increasing continuity risk.

Finally, integrity risk can also be mitigated at the expense of *availability*. Because, for a given failure (d), the failure mode slope $\| CH^+d \| / \| (I - HH^+)d \|$ is a function of satellite geometry—through the observation matrix H—it is possible to reject geometries on the basis of the maximum predicted mode slope [Chin]. In an approximate sense, those geometries whose maximum failure mode slopes are greater than the slope of the line passing through the origin and the intersection of the threshold and protection radius can be declared to be unavailable for navigation. (See Figure 5.6 (c).)

5.2.3 Fault Detection with Carrier Phase

When *code phase* measurements are used for such high accuracy and high integrity applications as aircraft precision approach and landing, the navigation performance parameter tradeoffs discussed above can be quite severe. In contrast, when

(a) Nominal View   (b) Expanded View
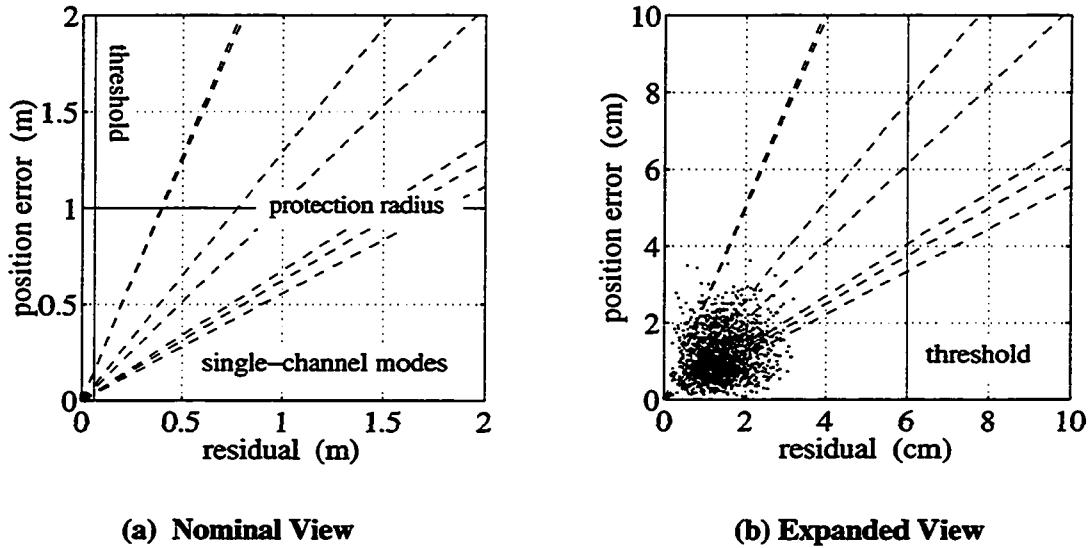
Figure 5.7: RAIM with Carrier Phase

carrier phase measurements are used, these tradeoffs are favorable since very low rates of missed detection can be achieved simultaneously with very low rates of false alarm. This concept has been experimentally validated using data collected across a short static baseline on the HEPL rooftop at Stanford University. During the experiment, seven satellites were available for positioning. (The baseline vector and cycle ambiguities were obtained through a prior static survey.) Figure 5.7 shows the position error and least squares residual for approximately 1000 position fixes which were executed over a 20 min interval. Because no failures occurred during the period under consideration, all of the data points are tightly packed into the normal error condition region near the origin ($\sigma_z \approx 0.8$ cm). The residual threshold in Figure 5.7 was set to provide a desired false alarm rate of one in 10^7 conservatively assuming $\sigma_z = 1$ cm. In addition, a protection radius of 1 m was assumed. The resulting missed detection region exhibited in Figure 5.7 is extremely narrow and is notably distant from the single-channel failure modes associated with each of the seven satellites in view.

5.2.4 RAIM Application to IBLS

The application of RAIM fault detection to the IBLS architecture can be separated into the two computational phases associated with a given approach; these are *cycle ambiguity resolution* and subsequent *kinematic positioning*.

5.2.4.1 Cycle Ambiguity Resolution

Because the stacked observation equation (3.31) is identical in form to equation (5.3), the discussion of Section 5.2.1 can be applied directly to the cycle ambiguity estimation problem. In this case the state vector u contains the cycle ambiguities and position fixes at each measurement epoch in the bubble. By appropriate choice of the matrix C (5.10), the vector x can be selected to correspond to the bubble-exit position fix.

On a typical approach, after pseudolite overflight and the completion of cycle ambiguity estimation, the residual is computed using equation (5.11). If the magnitude of the residual exceeds the predetermined threshold, a cycle ambiguity resolution failure is declared. Although the threshold could be a variable parameter, in order to ensure consistency with the ILS continuity specification (Table 1.1), it will ideally be set to generate a probability of false alarm, $P(FA|NC)$, of less than 10^{-6} per approach. The navigation performance, including integrity risk, afforded by the application of RAIM-based fault detection for the IBLS cycle resolution process will be quantitatively assessed in Chapter 6.

It is important to explicitly state that the nonlinearity present in the cycle ambiguity estimation problem (resulting from the pseudolite observation equation) has been deliberately ignored in the fault detection analysis thus far. The reason for this omission is based on the fact that only very large failure magnitudes would appreciably affect the linearity of fault signatures (i.e. the failure "directions" in Figure 5.5). The validity of this assumption is illustrated by example using Piper Dakota flight test data
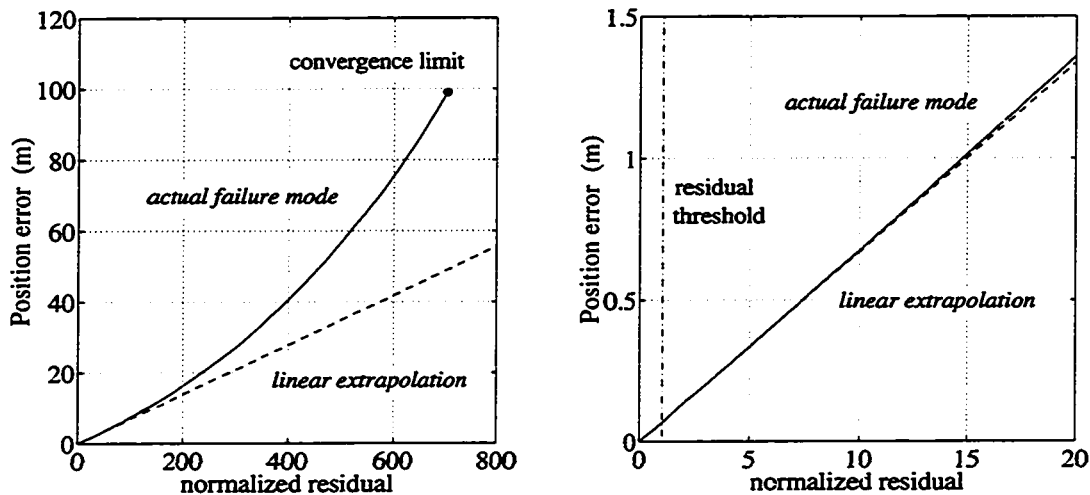
Figure 5.8: Fault Signature Nonlinearity (Cycle Slip Example)

collected at Palo Alto Airport. The data collected during pseudolite overflight was post-processed with a cycle slip intentionally induced on an arbitrary satellite channel at approximately the middle of the pseudolite bubble pass. Figure 5.8 shows the fault signature behavior as the cycle slip size is increased from zero to 370 cycles. In this plot, the residual axis is normalized by the threshold set to give $P(FA|NC) = 10^{-7}$. For slip magnitudes of 371 and above, the cycle ambiguity resolution algorithm did not converge, providing perhaps the most obvious means for the detection of the large failure. For cycle slip sizes less than 371 cycles, the fault signature does exhibit the effect of system nonlinearity; however, when compared on the scale of the residual threshold, it is clear that the true nonlinear fault signature is nearly indistinguishable from its linearized counterpart.

In general, the least squares residual can be used to detect a wide variety of failure modes during cycle ambiguity resolution. However, two failure modes discussed in Chapter 4 require some special consideration. These are *movement or mis-location of the ground-based pseudolites* (Section 4.2.2) and *satellite ephemeris errors* (Section 4.2.3).

Consider first the two pseudolite configuration illustrated in Figure 4.7. Because both pseudolites provide the capability for estimating *intrack* position, movement of either pseudolite intrack will result in an inconsistency between the measured carrier phase histories for the two pseudolites. This inconsistency will be evident from cycle ambiguity estimation in terms of the least squares residual. Although a quantitative assessment of the overall detectability of pseudolite movement will be made in Chapter 6, for the present discussion, the movement of pseudolites along the intrack axis can be considered to be a nominally detectable failure.

However, movement of a pseudolite in the *crosstrack* or *vertical* directions is, in general, not detectable by checking the consistency of the two pseudolite measurement histories, because both pseudolites are needed to define the absolute trajectory in the crosstrack-vertical plane. Some measure of the carrier-based absolute position quality at bubble exit is available in comparison with a pure differential code position fix. Although the code position fix is immune to pseudolite movement, its lower accuracy denies the capability for the reliable detection (with low false alarm rate) except for large (>5 m) position errors. The means for robust detection of crosstrack or vertical pseudolite movement using carrier phase RAIM (applied to cycle ambiguity estimation) requires the addition of a *third pseudolite*.

Additional pseudolites can also aid in the detection of ephemeris failures. For example, the 'trajectory rotation' about the axis joining the pseudolite transmitters illustrated in Figure 4.8 (due to an ephemeris error on a high-elevation satellite) may be undetectable using two pseudolites. In addition, the use of code phase measurements can again provide only marginal capability for the detection of this failure type. A code phase ground integrity monitor with a large displacement (>1 km) from the reference station can observe such failures provided that the axis joining the reference and monitor antennas lies in the plane of trajectory rotation.

FLIGHT PATH

PL 3

PL 1   PL 2

(a) Triangular

PL 4   PL 3

PL 1   PL 2

(top view)

(b) Rectangular
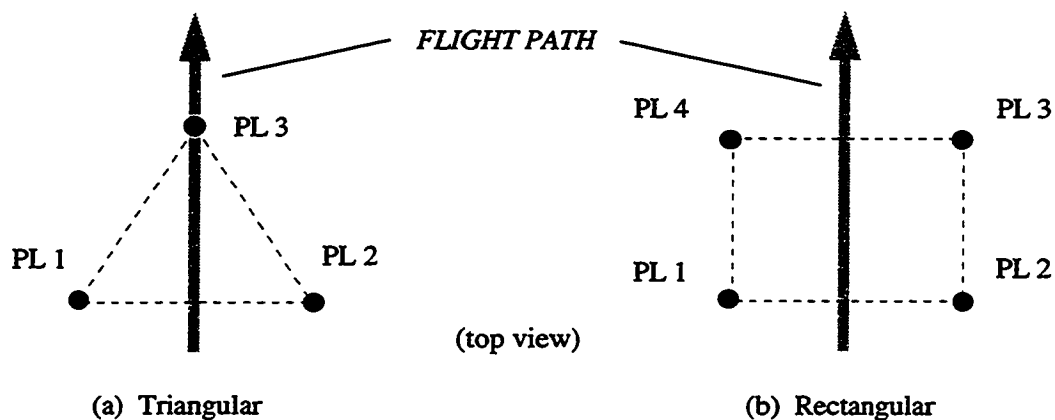
Figure 5.9: Pseudolite Geometric Configurations

The effectiveness of carrier phase RAIM in the cycle ambiguity resolution process in the detection of this class of rare failures is improved if a third pseudolite, non-collinear with the two nominal pseudolites, is placed under the approach path. Two example pseudolite layouts that provide the capability to detect both ephemeris failures and crosstrack-vertical pseudolite mis-siting are shown in Figure 5.9. The triangular configuration (a) is the simplest geometric layout that provides this fault detection capability. The optimization of pseudolite separation in this triangular layout will be discussed in Chapter 6. Alternatively, the rectangular configuration (b) can provide the additional capability to ensure fault detection availability even in the event of a single pseudolite outage.

5.2.4.2 Kinematic Positioning

Two fundamental questions arise in regard to the application of carrier phase RAIM after cycle ambiguity resolution:

1. What is the role of RAIM during this phase of flight?

2. Are redundant measurements available to provide the capability for RAIM after the aircraft exits the pseudolite bubble?

With respect to the first question, we begin by noting that there is substantial benefit in terms of overall system integrity resulting from the application of IBLS when compared to simpler alternatives such as differential code or carrier-smoothed code systems employing ground integrity monitors [Rowson]. Specifically, the detectability of rare-event multipath errors and SV 19-type failures is questionable using the latter architectures. The use of carrier phase in IBLS ensures that these failures will not be present. In addition, when compared to the OTF cycle resolution techniques (search methods) employed by alternative carrier phase architectures [Van Graas, a], the IBLS-capability for direct motion-based cycle ambiguity estimation and its associated built-in RAIM availability (discussed in the previous section) alleviates the integrity risk of erroneous cycle ambiguity resolution.

In addition, as noted in Section 5.1, the availability of RAIM during pseudolite overflight also offers the means for the detection of latent failures in each of the navigation system segments, including for example, the tampering/spoofing scenarios described in Section 4.2.2. In the event that a spoofing transmission is directed at the aircraft exclusively, neither ground monitoring nor redundant sensor tracks aboard the aircraft can provide an adequate defense. With the IBLS architecture, such latent failures are precluded when cycle ambiguity resolution is deemed to be successful by the RAIM check at bubble exit. During the remaining short exposure time of the approach (~1 min), the inception of failure types which are undetectable by means of (carrier phase) ground monitoring or redundant airborne sensor tracks is, by any estimation, likely to be extremely rare. One potential scenario, however, that does motivate the use of RAIM after cycle ambiguity resolution is an airborne cycle slip event (undetected in the receiver signal processing) occurring simultaneously on multiple sensor tracks.

In addition to addressing the question of the role of RAIM after bubble exit, it is important to note that even if RAIM is deemed necessary during this phase of the approach, the availability of redundant satellite measurements is not guaranteed (question 2 above). A quantitative assessment of the availability of RAIM for kinematic positioning was performed by Monte Carlo simulation. Seven representative international cities approaches were selected for the simulation: San Francisco, Chicago, New York, London, Amsterdam, Frankfurt, and Tokyo. In the simulations, a current ephemeris was used to propagate the positions of the 24 Block II and IIA GPS spacecraft. In addition, the Phlong and Elrod satellite state availability model [Phlong] was used to simulate the effect of scheduled (stationkeeping maneuvers, etc.) and unscheduled (hard failures) spacecraft down time. An elevation mask of 7.5 deg was assumed.

The most fundamental measure of RAIM availability is simply the raw availability of having at least five satellites in view. The raw availability results obtained from simulation are given in the first column of Table 5.2. For an unaugmented GPS constellation, the fraction of time five of more satellites are in view is 99.4 %.

To assess potential improvement in the raw availability of RAIM, several augmentation schemes were also tested. First considered were ranging signals from three geosynchronous Inmarsat spacecraft, which are soon to be equipped with GPS Wide Area

AUGMENTATION	RAW AVAILABILITY (%)	TOTAL AVAILABILITY (%)
None	99.4	97.3
Three Inmarsats	99.98	99.94
One Pseudolite	99.97	99.90
Two Pseudolites	99.998	99.99

Table 5.2: Carrier Phase RAIM Availability for Kinematic Positioning

Augmentation System (WAAS) transponders, at longitudes of 15.5W, 55W, and 179E deg. A representative hard failure model was also applied to these spacecraft [Phlong]. As expected, the results show much improved raw availability (99.98%) resulting from an average of approximately one additional satellite in view at each of the selected sites. From a practical standpoint, however, it is important to note that the deployment of additional spacecraft or spacecraft payloads (such as Inmarsat) dedicated to providing greater GPS availability is not guaranteed.

Augmentation from additional ground-based *terminal area pseudolites* (separate from the nominal IBLS pseudolite configuration) was also considered. These long-range RAIM-augmentation pseudolites were assumed to have a hard failure probability of 10^{-4} per approach. Pseudolite azimuths were randomly selected uniformly between 0 and 360 deg, while elevations varied uniformly between -5 deg and +5 deg. The results again show much improved raw availability. Extrapolating the results in Table 5.2 suggests that raw availability can reach any desired level as the number of space-based or ground-based ranging sources is increased. Again, however, practical considerations must also be weighed before these numerical results become meaningful. At present, the most promising approach in mitigating the near/far issues associated with long-range pseudolites is based on pulsing of the transmitted signal [Cobb, b]. Although recently published flight test results are encouraging [Lawrence, b], the ability of pulsed pseudolites to provide robust performance for precision approach is unclear. In particular, the capability of aircraft receivers to consistently maintain carrier phase-lock on signals originating at low elevations (due to low antenna gains and fuselage blockage and diffraction effects) has not been conclusively demonstrated.

As was discussed in Section 5.2.3, the missed detection region is extremely narrow using carrier phase (see Figure 5.7). Although a hypothetical failure mode penetrating this missed detection region is mathematically possible to construct (for

example, an error vector in the range of H would suffice), given that such a mode must be related physically to a *real system failure*, the likelihood of its occurrence will be extremely low. In addition, the integrity risk arising from such failures can be further mitigated using the approach described in Section 5.2.2 by limiting the availability of navigation on the basis of the maximum single-channel failure mode slope for a given satellite geometry [Chin, Brown, b]. Specifically, an availability limit slope can be set such that for a given satellite geometry, if the worst-case mode slope exceeds the limit, RAIM (and navigation in general) is declared to be unavailable.
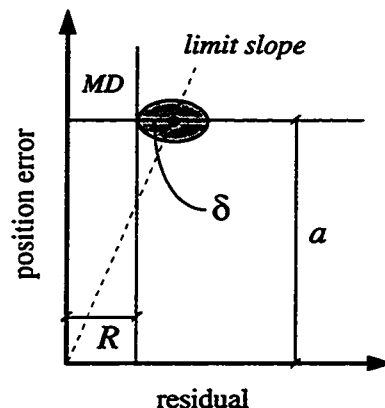
Figure 5.10: Availability Limit Slope

As illustrated in Figure 5.10, the availability limit slope is a function of the protection radius (a), residual threshold (R), and an integrity buffer (δ). Because of the high precision of carrier phase (see Figure 5.7) both δ and R are small, and it is expected that the availability limit slope is very steep. This assertion can be validated quantitatively using representative values for parameters a, R, and δ. A protection radius of 1.1 m is used for consistency with the existing Category III ILS requirements (Section 1.1.2). The residual threshold is taken to be 6.2 cm to generate a false alarm probability of 4×10^{-8} per epoch for the average case of seven satellites in view. This rate of false alarm conforms with the ILS Category III continuity requirement of 2×10^{-6} for each 15 sec interval of the approach assuming independent carrier phase measurements with 1 cm standard deviation collected at a 5 Hz rate.

If the integrity buffer is taken to be zero (for the moment), the availability limit slope is $110/6.2 \approx 18$. In this case, a cycle slip on a hypothetical worst-case satellite

(whose mode slope is equal to the limit slope), with worst case magnitude, may statistically result in a missed detection due to the effect of measurement noise. However, even in this worst-case situation, because of the high precision of carrier phase (illustrated, for example, in Figure 5.7), the position error may exceed the protection radius by a few centimeters. In reality, of course, a position error of 1.2 m will not be appreciably more hazardous than an error of 1.0 m. In this sense, it is legitimate to consider limit slopes as high as 18 using carrier phase.

A more conservative estimate of the availability limit slope, however, is also possible by choosing δ based on an assumed probability for a worst-case cycle slip. For example, we first assume the probability of cycle slip occurrence (undetected by receiver signal processing and occurring simultaneously on multiple sensor tracks) to be 10^{-4} for a given (post-bubble) approach of approximately one minute duration. Now, given that a cycle slip has occurred, the likelihood that the satellite with the worst-case mode slope has experienced the slip, and furthermore, that the slip size is equivalent to the worst case magnitude, is assumed to be 10^{-2}. Thus, the overall probability of a worst-case cycle slip is 10^{-6} per approach. In order to protect against such a cycle slip at each measurement epoch and provide a missed detection probability of 10^{-10} per approach, the integrity buffer δ is selected such that the residual at a given epoch will be smaller than R, and the vertical position error larger than a, with a probability of 10^{-4}. Applying a non-central χ^2 distribution [Abramowitz] to the residual statistic in the presence of the worst-case cycle slip bias, the integrity buffer is computed to be 3.1 cm. The resulting availability limit slope is thus approximately

$$\frac{a}{R+\delta} \approx 12 \ . \tag{5.17}$$

The total availability of satellite geometries having five or more satellites in view and maximum single-channel failure mode slopes smaller than 12 was also quantified

from the Monte Carlo simulation; these results are included in Table 5.2. Note that even for the unaugmented constellation, the total availability of RAIM after pseudolite overflight is approximately 97.3%. In a general sense, it is noteworthy that, due to the high precision of carrier phase, the total availability is bounded rather closely by the raw availability of having five satellites in view.

5.3 RAIM-Based Fault Isolation

The overall likelihood of a navigation system failure (SF) can be expressed as the following sum:

$$P(SF) = P(MD) + P(DF) \qquad (5.18)$$

It is clear from equations (5.2) and (5.18) that as the probabilities of missed detection and false alarm approach zero, continuity risk approaches the navigation system reliability limit $P(SF) \approx P(DF)$. Aside from improving component reliability in the navigation system, further mitigation of continuity risk below this limit must be accomplished through fault isolation and removal. It is important to note, however, that fault isolation can only be achieved at the expense of increased integrity risk because *mis-identification* becomes a possibility. This condition is conceptually illustrated in the block diagram of Figure 5.1 and mathematically expressed in equations (5.1) and (5.2), where we note that $P(M \cap DF) \equiv 0$ and $P(N \cap DF) \equiv 1$ if no means for fault identification exists.

We recall that the IBLS architecture, with carrier phase RAIM, provides the means for mitigation or detection of failure modes (SV 19, tampering/spoofing, limit-case code multipath, etc.) which are difficult or impossible to detect using alternative architectures that rely on ground monitoring and redundant airborne sensors alone. In addition, carrier phase RAIM provides the means for detecting failures originating in all three navigation system segments (ground, air, and space), and when used together with

ground monitoring and redundant airborne sensor tracks, provides an independent layer of protection against all failure classes, including unknown new failure types arising from the transition from ILS to GPS. However, because the continuity requirement for precision approach and landing is much less stringent than the integrity requirement, the need for fault isolation through RAIM is subordinate, given that such capability is provided by the ground integrity monitor and redundant airborne sensor tracks. In the general sense, then, RAIM-based fault isolation will likely not play as critical a role as RAIM fault detection in precision landing navigation (for example, see [WMOPS]). However, the capability for partial-isolation through RAIM can be a useful aid in the identification and removal of certain specific failure types—most notably, airborne receiver cycle slips—which may be difficult to directly isolate otherwise. In this context, and in the interest of exploring the limits of high integrity and high continuity kinematic GPS performance, methods of RAIM-based failure isolation are also presented.

In general, RAIM-based isolation is more difficult than detection because of the obvious need to be *explicit* about the nature of the detected failure. For detection, only the *magnitude* of the residual vector, r, was used. For isolation, the *direction* of r will be of paramount importance. In addition, isolation requires at least one more measurement than detection—six satellites instead of five for the case of kinematic positioning. Thus, after pseudolite overflight, the availability of the redundant measurements needed for RAIM-based isolation will be notably lower than for detection (unless a means for ranging augmentation is provided—e.g., terminal area pseudolites).

5.3.1 Parity Space

Since the $n \times 1$ least-squares residual vector r is orthogonal to the columns of the $n \times m$ observation matrix H (see Appendix C), the elements or r are not all independent. However, if we consider only the $(n - m)$–space orthogonal to the columns of H, often

referred to as *parity space*, the residual vector can be expressed in a more compact form. We begin with the singular value decomposition of the observation matrix as described in Appendix C:

$$H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S \\ 0 \end{bmatrix} V^\tau. \tag{5.19}$$

In this case, the $n \times n$ unitary matrix U has been partitioned into an $n \times m$ matrix U_1 and an $n \times (n-m)$ matrix U_2. The other elements of equation (5.19) are identical to those in equation (C.3) in Appendix C. Using the definition of the residual (5.12) and equation (C.8), the residual vector can now be expressed as

$$r = U_2 U_2^\tau (\delta z - \delta H u), \tag{5.20}$$

and the *parity vector* is defined to be

$$p \equiv U_2^\tau (\delta z - \delta H u) = U_2^\tau r. \tag{5.21}$$

Under normal error conditions, the measurement error is distributed as $N(0, \sigma_z^2 I_n)$. Therefore, since $U_2^\tau U_2 = I_{n-m}$, we have

$$\|p\| = \|r\| \tag{5.22}$$

and

$$p \sim N(0, \sigma_z^2 I_{n-m}). \tag{5.23}$$

Figure 5.11 illustrates the case where $n - m = 2$; the shaded error circle centered at the origin represents the normal condition error dispersion (5.23). This scenario, for example, is consistent with the case of positioning with six satellites.

Given the event of an actual failure d, the parity vector is distributed as

$$p \sim N(U_2^\tau d, \sigma_z^2 I_{n-m}). \tag{5.24}$$

Thus, as the failure magnitude $\|d\|$ is increased, the nominal error dispersion circle will slide up the projected failure axis $U_2^\tau d / \|d\|$ as shown in Figure 5.11. The signature axes for a number of hypothetical failure modes are indicated in the figure. Because the magnitudes of the parity and residual vectors are the same, the fault detection threshold R

Figure 5.11: Parity Space

takes the form of a circular boundary in the figure.

The principal utility of the parity vector and parity space is the simple and elegant framework it provides for the development of isolation algorithms. In a general sense, once a failure is detected ($\|p\| > R$), the failure may also be isolatable depending on how close p happens to be relative to the known failure mode signatures.

Generalized Likelihood Ratio

Formally, most failure isolation problems can be classified as composite multiple hypothesis problems [Van Trees]. This means that while the failure *modes* may be known, certain important information, such as the probability density of a failure magnitude (given the failure type) is often not known. For this type of problem, a frequently applied algorithm that is often effective is the Generalized Likelihood Ratio (GLR) test. Although this algorithm can be cast in forms suitable for either sequential [Willsky, a, b] or single-epoch least squares estimation [Daly], for clarity of explanation,

we will consider the latter case. This approach is consistent with application to post-bubble kinematic positioning.

The application of the GLR algorithm to the isolation of single channel failure modes proceeds as follows: The signature direction q_i of a failure on receiver channel i in parity space is given simply by the i-th column of U_2^T normalized by its

Figure 5.12: Isolation Integrity Risks

magnitude. Given a detected failure, the GLR algorithm isolates the failure mode i that maximizes the inner product $p^T q_i$. The result, of course, is simply the failure signature axis to which the given parity vector is closest. The justification behind this algorithm is that the identified failure mode is the one with maximum conditional density $f(p \mid FM_i)$ —where FM_i denotes a failure on channel i—with the implicit assumption that the failure magnitude is equal to its maximum likelihood value.

Other than its simplicity, an additional apparent benefit of the unmodified GLR test is that isolation is *always* possible—ensuring that continuity is maintained. However, this gain in continuity is obtained at the expense of increasing integrity risk. Two example cases in which the unmodified GLR algorithm results in low integrity isolation are shown in Figure 5.12, in which $n - m = 2$ (kinematic positioning with six spacecraft). In this event, the failure modes shown represent anomalies in a given row of the observation equation (5.3); thus, FM_i is a failure in the i-th measurement or the i-th row of the observation matrix. Referring to the figure, for parity vector A, failure mode 2 is isolated. However, note that $f(p \mid FM_2)$ is only slightly larger than $f(p \mid FM_1)$ as

indicated by the relative distances of point A to fault signature axes 1 and 2. The resulting risk of choosing the wrong failure mode and leaving the true failure in place is obviously rather high in this situation. For case B, again failure mode 2 is isolated. In this case, no other known failure modes are nearby, but $f(p \mid FM_2)$ is extremely small, even assuming the maximum likelihood failure magnitude. In this event, the risk of an unknown linear combination of the failure modes shown (i.e., multiple channel failures) may be quite high. In general, a maximum of $n - m - 1$ (which is 1 for the example) failed rows can potentially be isolated at one time using any parity space method, including GLR.

The specific needs associated with precision approach and landing applications clearly motivate the development of an alternative isolation algorithm targeted toward improving continuity subject to the constraint of maintaining an extremely high level of integrity.

5.3.2 Fault Isolation with Carrier Phase

In addition to the benefits provided to fault detection described in Section 5.2.3, the high precision of carrier phase, when compared to code, also greatly benefits the integrity of the failure isolation process since mis-identification will be less likely. The basic reason for this is that the ratio of failure magnitude to measurement noise will be much larger for carrier phase than for code, thus providing a much cleaner observation— through the parity vector—of the effect of the failure. This concept is illustrated in the simulation results shown in Figure 5.13, which shows the parity space for a typical geometry with six satellites in view using both code phase and carrier phase measurements. The single channel fault signatures are plotted, as is the fault detection threshold R corresponding to a per-epoch false alarm rate of 10^{-8} under normal error conditions. The parity vectors obtained from 1000 kinematic position fixes under normal

(a) **Measurement Dispersion using Code**

(b) **Measurement Dispersion using Carrier**
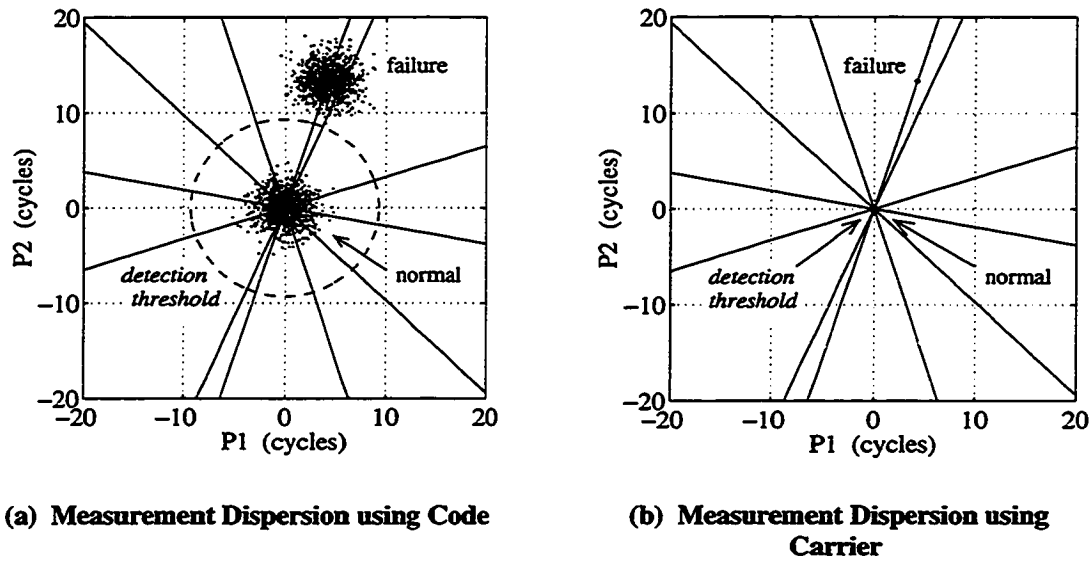
Figure 5.13: Example Parity Space for Positioning with Six Satellites

error conditions ($\sigma_z = 1$ cm for carrier and $\sigma_z = 30$ cm for code) are shown in the plots. An additional 1000 position fixes were executed with normal measurement errors superimposed on a bias channel failure of 30 cycles induced on channel 1. In a qualitative sense, it is plainly evident in the plot that the extremely tight parity vector dispersion due to the small carrier phase measurement noise provides an essentially unobscured view of the failure on channel 1, whereas the wide dispersion of the code phase noise makes it difficult to distinguish between adjacent failure signatures.

5.3.2.1 Fault Isolation in Kinematic Positioning

In order to maintain low integrity risk when RAIM-based fault isolation is implemented, it is necessary to accept a somewhat reduced overall improvement in continuity. In simple terms, under certain conditions (such as the two cases illustrated in Figure 5.12) isolation of a specific failure mode should not be attempted. In general, a high integrity algorithm should ensure that failures are isolated only when the parity vector lies very near one failure mode and is far from all others.

Given a detected failure, the most direct method for quantifying the appropriate isolation boundaries uses Bayes Rule:

$$P(FM_i|p) = \frac{P(FM_i)f(p|FM_i)}{\sum_{j=1}^{n} P(FM_j)f(p|FM_j)}. \tag{5.25}$$

For the present, we will consider only the single-channel failures $FM_1, FM_2, \ldots, FM_n$. (The effect of multi-channel modes will be addressed later in this section.) If (5.25) can be quantitatively evaluated, a simple high-integrity algorithm would isolate and remove failure mode i if

$$P(FM_i|p) > 1-\varepsilon, \tag{5.26}$$

where ε is a very small number ($< 10^{-9}$) set according to the desired level of isolation integrity $P(MI|DF)$.

The conditional density of p given a failure in channel i of magnitude γ is

$$f(p|FM_i,\gamma) = N(\gamma Q_i q_i, I_{n-m}), \tag{5.27}$$

where Q_i is the magnitude of i-th column of U_2^T and q_i is its direction. In general, γ is not known, and may in fact itself be a random variable. However, given a prior probability model for γ, equation (5.25) is tractable in the sense that it can be expressed independently of γ. For example, if no prior knowledge of failure magnitudes is available, we may model the failure magnitude as uniformly and identically distributed for each channel as

$$\gamma \sim \lim_{M \to \infty} U[-\tfrac{M}{2}, \tfrac{M}{2}]. \tag{5.28}$$

In this case, it is shown in Appendix D that equation (5.25) reduces to

$$P(FM_i|p) = \frac{\alpha_i \exp\left[-\frac{1}{2}\left\| p - (p^\tau q_i)q_i \right\|^2\right]}{\sum_{j=1}^{n} \alpha_j \exp\left[-\frac{1}{2}\left\| p - (p^\tau q_j)q_j \right\|^2\right]} , \qquad (5.29)$$

where

$$\alpha_i \equiv P(FM_i)/Q_i . \qquad (5.30)$$

The isolation algorithm defined by (5.26), (5.29), and (5.30) implicitly suppresses the problem associated with parity vector A in Figure 5.12, because (5.26) can only hold true for *at most* one failure mode. However, the algorithm does not address the issue of parity vector B. This is to be expected, of course, since the second problem is related to the existence of multi-channel modes, which were excluded from the above derivation by hypothesis.

For simplicity, we first consider the case of $n-m=2$. In this instance, multi-channel modes, although they will likely be exceptionally rare, cannot be isolated using RAIM. Suppose now that a failure has been detected and isolated to channel i using the algorithm described above. Qualitatively, in order for the algorithm to be effective against "B"-type events (in Figure 5.12) it should also verify that the parity vector is not too distant from failure mode i. Mathematically, the algorithm should choose *not* to isolate the i-th failure mode if

$$\left\| p - (p^\tau q_i)q_i \right\| > \rho , \qquad (5.31)$$

where ρ is a predefined threshold. In this case, the entire receiver would be declared inoperational.

The distance criteria (5.31) is essentially equivalent to reapplying the original detection algorithm with the hypothesis that $(p^\tau q_i)q_i$ is the fault vector. We note, however, that because we have prior information that a failure has in fact occurred (from

fault detection), it is necessary that $\rho < R$ to maintain integrity against mis-isolation. This of course, implies that

$$P(NI|DF) > P(FA|NC). \tag{5.32}$$

The composite algorithm, using both the Bayes criteria and distance criteria, would mitigate the integrity risk associated with fault isolation by declaring both events A and B to be non-isolatable. The performance of the single-channel RAIM isolation algorithm described above is evaluated through extensive simulation and experiment in reference [Pervan, a].

In the event a given channel i is isolated, the measurement can be efficiently removed using the stored covariance matrix $(H^T H)^{-1}$ by application of the measurement downdate (Appendix E)

$$\hat{x}_{new} = \hat{x}_{old} - \frac{(H^T H)^{-1} h_i}{1 - h_i^T (H^T H)^{-1} h_i} (z_i - h_i^T \hat{x}_{old}), \tag{5.33}$$

where z_i is the i-th measurement, h_i^T is the i-th row of the observation matrix H, and i.i.d. measurements have been assumed.

Extension to Multi-Channel Isolation

When $n - m > 2$, RAIM-based fault isolation can conceivably be used for isolation of multiple channel failures. For example, if three redundant measurements are available, combinations of failures of any two channels may be considered. However, in practical receiver applications, the likelihood of failure on two channels is extremely small when compared to a single channel failure (such as a cycle slip on a low signal strength satellite). In fact, given the degree of commonality between channels (e.g., the RF electronics and local oscillator), a two-channel failure may actually be less likely than a full receiver failure resulting in simultaneous failures on all channels. Thus, in the

context of the envisioned redundant airborne receiver architecture, the practical benefit (in terms of mitigating continuity risk) of providing the capability to isolate multiple-channel failures through RAIM–even when permitted by the degree of measurement redundancy–is likely to be inconsequential.

5.3.2.2 Fault Isolation During Cycle Ambiguity Resolution

The parity space methods described above can also in principle also be applied to the cycle ambiguity resolution problem. However, the large structure of equation (3.31) and the possibility of a failure occurring on any given channel at any given time during pseudolite overflight, instead strongly suggest the application of a *sequential* fault isolation algorithm. This method is, in fact, consistent with the sequential information smoother algorithms actually used in practice for cycle ambiguity estimation (see Sections 3.3.2.2 and 3.3.2.3). In this case, the *innovations* vector (see [Pervan, a]) is simply used in place of the parity vector in the fault isolation algorithms.

Care must be taken, however, with respect to the inherent nonlinearity in the observation equation. Although the effect of nonlinear fault signatures has a negligible effect on fault detection performance (Section 5.2.4.1), fault isolation algorithms such as those described above rely heavily on the assumption of linearity of the fault signatures. Thus, for very large failure magnitudes, a significant integrity risk may result.

For the case of airborne cycle slip isolation, however, this risk can be mitigated through a pre-RAIM consistency comparison of code and carrier based ranges. This check provides an alternative means for the direct isolation of the very large slips that are the cause of nonlinear fault signatures.

123

CHAPTER 6

Navigation Performance

T he performance of any GPS-based navigation system for aircraft precision approach and landing must ultimately be *quantified* in order to prove that the performance requirements have actually been satisfied. This task, however, is complicated by two factors: First, at the time of this writing, the actual requirements for GPS-based precision landing systems have not yet been universally agreed upon. Second, the envisioned integrity requirement, which in any case will be on the order of approximately one hazardous failure in one billion approaches, is extremely difficult to quantitatively demonstrate regardless of the precise specification eventually agreed upon. In this context, a *parametric* approach, based on Monte Carlo simulation and flight test, has been used to generate a quantitative assessment of the overall kinematic navigation system performance.

6.1 Quantifying Navigation Performance

The navigation system accuracy (A) specified in terms of its 95% error performance is mathematically expressed as

$$\left\{ A \mid P(\delta x \leq A) = 0.95 \right\}. \tag{6.1}$$

Navigation integrity and continuity risk are identified in equations (5.1) and (5.2), respectively. For the purposes of conservatively quantifying integrity and continuity risk and ensuring reasonable availability, RAIM-based fault isolation is not assumed. Furthermore, if a single track of airborne GPS avionics is considered (i.e., no isolation capability) equations (5.1) and (5.2) reduce to

$$I = P(MD) \tag{6.2}$$

and

$$C = P(FA) + P(DF). \tag{6.3}$$

The performance benefits of fault isolation through the application of multiple airborne sensor tracks will be discussed in Section 6.4.

Since the normal error conditions (NC) and system failure (SF) states are mutually exclusive, the integrity risk (6.2) can be expanded as

$$I = P(MD \mid NC) P(NC) + P(MD \mid SF) P(SF). \tag{6.4}$$

Because $P(NC) \approx 1$, we can write

$$I \leq P(MD \mid NC) + P(MD \mid SF) P(SF). \tag{6.5}$$

Now using the definition of the missed detection event (5.15) and the fact that the measurement residual and state estimate error are independent under normal error conditions, the first term on the right-hand side of (6.5) can be expressed as

$$P(MD \mid NC) = P(\|\delta x\| > a \mid NC) P(\|r\| < R \mid NC) \leq P(\|\delta x\| > a \mid NC). \tag{6.6}$$

Substituting (6.6) into (6.5) yields

$$I \leq P(\|\delta x\| > a \mid NC) + P(\|\delta x\| > a, \|r\| < R \mid SF) P(SF), \tag{6.7}$$

where the left-hand side is closely bounded by the right. The terms on the right-hand side of (6.7) correspond to the integrity risk incurred under normal error conditions and system failure, respectively.

In an analogous manner, the continuity risk in equation (6.3) can also be expanded as

$$C \leq P(\|r\| > R \,|\, NC) + P(\|r\| > R \,|\, SF)P(SF) . \tag{6.8}$$

The explicit separation of the normal condition and failure state terms in the continuity and integrity risk expressions in equations (6.7) and (6.8) provides the basis for separate consideration of the two operational states. In the sections that follow, the utility of this separation will become clear in that the normal condition terms can be quantified by analytical means, whereas the Monte Carlo simulation can be devoted entirely to sampling only failure cases.

Finally, the fraction (F) of time the navigation system is available before the approach has begun is related to the likelihood of having either poor satellite geometry (PG)—see Section 5.2.4.2—or a hard navigation system (airborne or ground) failure before the approach (FBA) has begun, or both. Assuming (somewhat conservatively) that these geometry and failure events are mutually exclusive, we can write

$$F = 1 - \left[P(PG) + P(FBA) \right] . \tag{6.9}$$

Equation (6.9) indicates that design for reliability (low overall system failure rates) is obviously an important factor for availability. Of course, given any GPS-based navigation architecture with built-in ground and airborne hardware redundancy, it is reasonable to expect that $P(FBA) \ll P(PG)$. In this case, equation (6.9) reduces to

$$F \approx 1 - P(PG) . \tag{6.10}$$

6.2 Fault-Free Performance

Navigation system accuracy (A) is the only performance parameter that can be quantified through direct experimental means. In this respect, the results of an extensive battery of flight tests of the breadboard IBLS navigation system will be reviewed. In addition, simulation-based covariance analyses can be used to complement the experimental results and assess performance under a variety of conditions difficult to realize in actual flight testing. Covariance analysis also provides the means to evaluate the normal condition (NC) component of integrity risk in equation (6.7).

6.2.1 Flight Test

From the first flight experiments in January of 1993 to the present, several hundred real-time approaches have been executed with the IBLS breadboard system (described in Chapter 2) using a two pseudolite ground configuration. An abbreviated history of the most important flight trials is given in Table 6.1.

In general, evaluation of navigation sensor error (NSE)–i.e., accuracy–has been limited by the accuracy of the reference ('truth') measurements. The most accurate source of truth used in any of the flight tests was established by means of an independent post-flight static GPS survey. This method was applied after tiedown for the Piper Dakota flight test at Palo Alto Airport (California) indicated in the first row of Table 6.1. Comparison with the real time kinematic positions obtained at tiedown indicated an RMS position fix error of approximately 2 cm [Cohen, *e*].

Although it is reasonable to expect that a positioning accuracy of 2 cm was maintained throughout the entire approach, further testing in the Piper Dakota was done to verify this hypothesis using other means of position verification. These included touch-and-go flight trials performed at Palo Alto Airport using a laser altimeter (Optech G150 Laser Rangefinder) to determine the aircraft vertical position with respect to the

runway together with a runway surface fit (accurate to roughly 5 cm) [Cohen, e]. The resulting RMS accuracy as compared with the laser altimeter output was approximately 5 cm— well bounded by the ±5 cm accuracy specification of the laser altimeter and the predicted runway surface fit error. Additional flight tests at Palo Alto Airport compared IBLS position output directly to the runway surface fit alone yielded a vertical $|\mu|+2\sigma$ NSE of roughly 15 cm. (Here μ is defined to be the mean error; the statistic $|\mu|+2\sigma$ is then roughly a measure of 95% error.) In this case, an additional dynamic error (~2 cm) in knowledge of the true GPS antenna position was present due to the flexibility and vibration of the aircraft during landing [Cohen, d].

In order to verify the full three-dimensional NSE of IBLS, flight trials were performed using the Piper Dakota and an NdYAG laser tracker at the NASA Ames Flight Test Facility in central California. Laser range accuracy was specified at ±1 ft (1σ) and

AIRCRAFT	LOCATION	TRUTH REFERENCE	NO. OF APRCHS	VERT. NSE (m)
Piper Dakota	Palo Alto, CA	static survey	1	0.02*
Piper Dakota	Palo Alto, CA	laser altimeter	6	0.05*
Piper Dakota	Palo Alto, CA	runway fit	15	0.15**
Piper Dakota	Crows Landing, CA	laser tracker	10	0.4**
King Air	FAA Tech. Cen., NJ	laser tracker	49	0.5**
Boeing 737	Crows Landing, CA	laser tracker	110	0.2**

* RMS error

** $|\mu|+2\sigma$ (95% error)

Table 6.1: IBLS Vertical NSE Flight Test Results

elevation and azimuth accuracy at ±0.2 mrad (1σ). For these tests, the Piper Dakota was equipped with a laser retroreflector (mounted on the tail). For the ten approaches for which data was collected, the vertical and horizontal $|\mu|+2\sigma$ NSEs were approximately 0.4 m—consistent with the expected magnitude of the laser tracker error.

In July of 1994, the IBLS breadboard was installed in a Beechcraft Super King Air 200 (see Figure 6.1) at the FAA Technical Center in Atlantic City, New Jersey, and was used in place of the nominal ILS glideslope and localizer receivers to provide the navigation source for 49 autocoupled approaches down to 50 ft altitude [Cohen, f]. During these approaches, the aircraft position was independently determined by a laser tracker (GTE precision automated tracking system) whose 1σ range accuracy was specified at ±1 ft and 1σ elevation and azimuth accuracy at ±0.1 mrad. The resulting vertical IBLS $|\mu|+2\sigma$ NSE was computed to be 0.5 m, again consistent with the laser error specifications.

Finally, as described in Section 3.5.2, in October of 1995 the IBLS breadboard

Figure 6.1: Beechcraft King Air at FAA Technical Center

129

was ultimately used to perform automatic landings of a United Airlines Boeing 737-300 at Crows Landing [Cohen, g]. In all 110 automatic landings were executed. The vertical $|\mu|+2\sigma$ NSE for these approaches, obtained by means of the laser tracker, was approximately 0.2 m.

6.2.2 Analysis and Simulation

Although the nominal two pseudolite configuration shown in Figure 1.3 is sufficient to provide exceptional accuracy performance—as evidenced by the flight test results just discussed—we recall from Section 5.2.4.1 that the use of three or more pseudolites, arranged in a non-collinear configuration, is advantageous in terms of improved RAIM fault detection capability, especially in the presence of satellite ephemeris and pseudolite movement failures. In this regard, the accuracy performance using a minimal three-pseudolite triangular configuration (see Figure 5.9) was evaluated through simulation-based covariance analysis.

In the analysis, an equilateral triangular layout was considered; assumptions included a 5 Hz measurement rate, a pseudolite far/near ratio of 3:1, and a 200 ft overflight altitude (at the crossing of the axis joining pseudolites 1 and 2 in Figure 5.9). Approach simulations were executed for a nominal 3 deg approach at San Francisco International Airport (SFO) Runway 28, using the satellite geometry and outage model described in Section 5.2.4.2 with a 7.5 deg elevation mask. Similarly, for the belly-mounted antenna, the elevation mask applied for lines-of-sight from the aircraft to the pseudolites was -7.5 deg—i.e., signals originating from elevations larger than -7.5 deg were assumed to be untrackable using the belly antenna. Initially, a simplified white-i.i.d.-gaussian measurement noise model with 1 cm standard deviation was implemented.

The 95% navigation sensor error (A) resulting from overflight of the triangular pseudolite configuration for 17,000 randomly selected satellite geometries (with spacecraft outages included) is shown in Figure 6.2 (a) as a function of pseudolite spacing (measured in terms of one-half the triangle side length). It is evident in the figure that the 95% error performance is relatively insensitive to pseudolite separations between 200 and 400 ft. For displacements below 200 ft, the geometric observability of the cycle ambiguities begins to decrease. In the limit, as spacing approaches zero, no crosstrack observability exists. For displacements above 400 ft, the elevation angle histories from the aircraft to pseudolites 1 and 2 are near the −7.5 deg mask, resulting in short tracking durations for these pseudolites and poor observability of absolute vertical position.

Given the acceptable range of 200–400 ft for the triangular layout, a number of practical considerations suggest that the pseudolite spacing be made as small as possible, the most obvious of which, of course, is that small pseudolite displacements require a smaller land area. In addition, however, we note that although
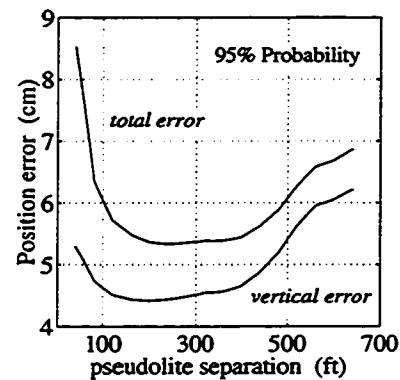
(a) Nominal 95% NSE
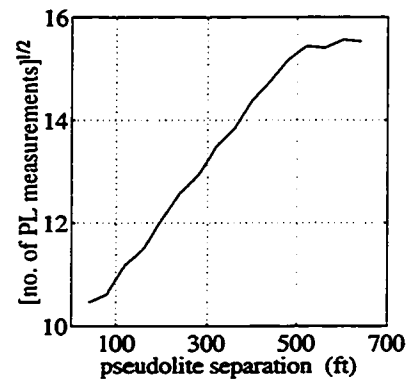
(b) Relative Noise Averaging
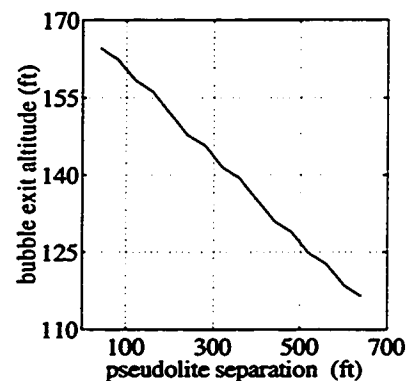
(c) Bubble Exit Altitude

Figure 6.2: IBLS Performance vs. Pseudolite Separation

accuracy performance is essentially identical for pseudolite separations of 200 and 400 ft, this level of performance is more dependent on the effect of measurement averaging (i.e., filtering of white noise) for the 400 ft case. This is most easily observed in Figure 6.2 (b), which is a plot of the square root of the total number of pseudolite measurements during the bubble pass (a relative measure of the white noise averaging effect) as a function of pseudolite separation. Together, Figures 6.2 (a) and (b) suggest that the performance for a 400 ft separation can be expected to be less robust to the presence of colored noise sources (such as multipath and differential troposphere) than that for a 200 ft separation. Finally, we also observe in Figure 6.2 (c) that the bubble exit altitude is nearly 20 ft higher for a 200 ft separation, thus providing an earlier alert height in the event of a detected failure in cycle ambiguity resolution. Within this context, a nominal pseudolite separation of 220 ft was selected as the baseline for more detailed analysis.

It is also interesting to note that while the three-pseudolite configuration provides improved fault detection capability (see Section 5.2.4.1) over a two-pseudolite configuration, the associated improvement in navigation accuracy (A) is due primarily only to the noise averaging effect of the additional measurements collected. For comparison, in the two-pseudolite case (using a half-separation distance of ~100 ft), $A \leq 7$ cm.

The 95% performance shown in Figure 6.2 (a) was computed assuming a direct aircraft overflight through the horizontal centroid of pseudolite configuration at the nominal (200 ft) altitude. In practice, of course, the aircraft will deviate somewhat from this nominal trajectory. The effect on 95% accuracy performance of such off-nominal trajectories is illustrated in Figure 6.3. Indicated in the figure is the specified 95% aircraft TSE bound at 200 ft altitude for a Special Category I (SCAT-I) approach. This requirement, which applies to differential satellite-based navigation systems (such as carrier-smoothed code) that are capable of providing sufficient performance down to 200
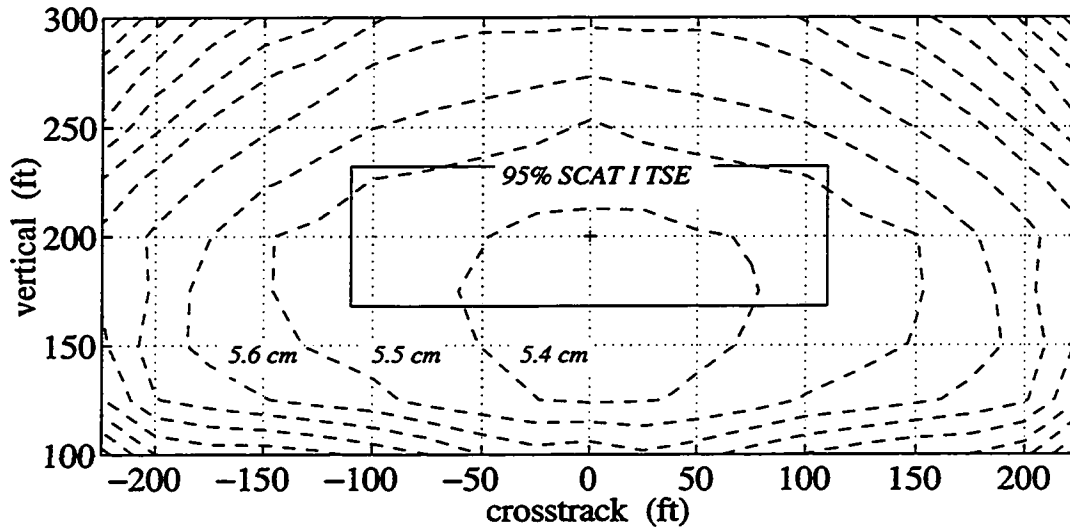
Figure 6.3: IBLS 95% Accuracy Performance for Off-Nominal Trajectories

ft altitude, stipulates that 95% of the time the aircraft trajectory must pass within the boundaries of the box shown in the figure. The IBLS performance contours in the figure show that the 95% accuracy performance penalty is on the order of millimeters even for trajectories well outside the SCAT-I TSE boundary.

In order to assess the navigation system performance beyond the 95% accuracy and quantify the integrity risk under normal error conditions, $P(\|\delta x\| > a \mid NC)$, the detailed error models derived in Chapter 4 were incorporated into the covariance simulation. The simulation output was stored directly as a histogram of post-bubble vertical position estimate error standard deviations $f(\sigma_x \mid NC)$ for all satellite geometries with five or more satellites in view (anticipating the post-bubble RAIM function needs). The probability of the vertical position estimate error exceeding a was then computed by a discrete application of the following gaussian convolution integral:

$$P(\|\delta x\| > a \mid NC) = \int_a^\infty \int_0^\infty N(\|\delta x\| \mid \sigma_x) f(\sigma_x \mid NC) d\sigma_x \, d\|\delta x\| . \qquad (6.11)$$

133
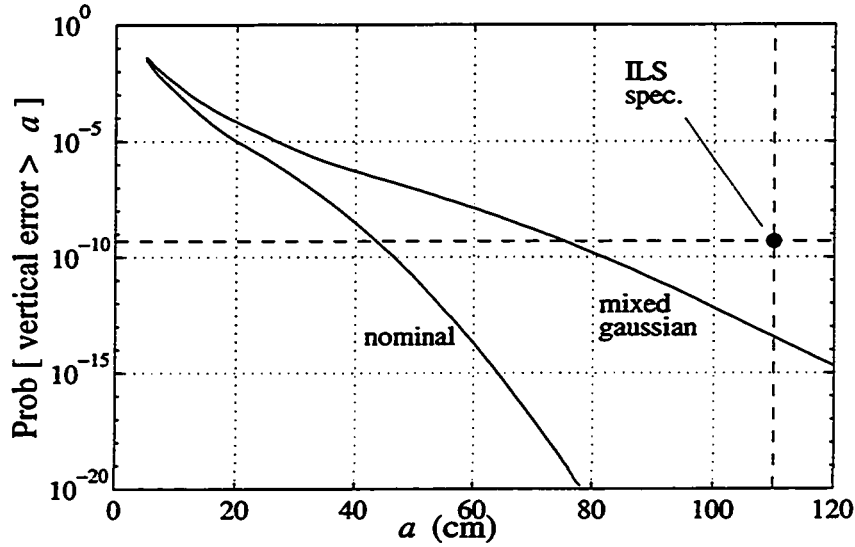
Figure 6.4: Integrity Risk Under Normal Error Conditions

The results are shown in Figure 6.4. The lower curve represents the integrity risk for a given vertical protection limit a assuming the nominal gaussian-based error models described in Chapter 4. It is clear from the curve that a protection radius below 0.5 m can be accommodated with a fault-free integrity risk of 10^{-10}.

In order to assess the sensitivity of fault-free integrity risk to the error model assumptions, a second application of the convolution integral (6.11) was performed using a more conservative 'mixed' gaussian model. In this case, the term $N(\|\delta x\| \,|\, \sigma_x)$ was replaced with

$$N_{mix}(\|\delta x\| \,|\, \sigma_x) \equiv (1 - \varepsilon_{mix})N(\|\delta x\| \,|\, \sigma_x) + \varepsilon_{mix}N(\|\delta x\| \,|\, 2\sigma_x) \qquad (6.12)$$

to effectively widen the 'tails' of the nominal gaussian-based error models. A mixture ratio of $\varepsilon_{mix} = 0.05$ was arbitrarily selected for comparison with the result obtained from the nominal error models. (The mixed gaussian result is shown in the upper curve of Figure 6.4.) Although the protection limit in this case increases to approximately 0.8 m

for an integrity risk of 10^{-10}, a significant margin is still present when compared to the 1.1 m specification for ILS. With the selected mixed gaussian model, the ILS vertical protection limit is achieved with a fault-free integrity risk of approximately 10^{-14}.

6.3 Failure-State Performance

In order to assess overall integrity risk, clearly it is necessary to evaluate navigation system performance in the presence of failures. In this regard, both flight experiments and computer simulations were specifically targeted toward this goal. Flight trials were performed to verify the real-time autonomous fault detection capability of IBLS. In addition, extensive simulation of aircraft approaches, with a number of diverse failure modes, was used to provide the basis for a parametric, quantitative assessment IBLS integrity risk (I).

6.3.1 Flight Test

Twenty-four test approaches were performed with intentionally induced failures to validate the RAIM fault detection function built into the real-time IBLS cycle ambiguity resolution algorithms. These approaches were conducted during two separate flight trials in the Piper Dakota at Palo Alto Airport using the nominal two-pseudolite IBLS breadboard system described in Chapter 2. At the bubble exit for each approach, the residual magnitude check (see Chapter 5) was executed to attempt to detect the induced failure.

On ten of the approaches, one of the pseudolites was deliberately moved either intrack, crosstrack, up or in an arbitrary spatial direction. The total displacement of the pseudolite transmit antenna from its nominal surveyed site varied between 0.5 and 2 meters. On a second set of ten approaches, cycle slips were induced numerically into the single difference carrier phase measurements prior to airborne processing. Cycle slip
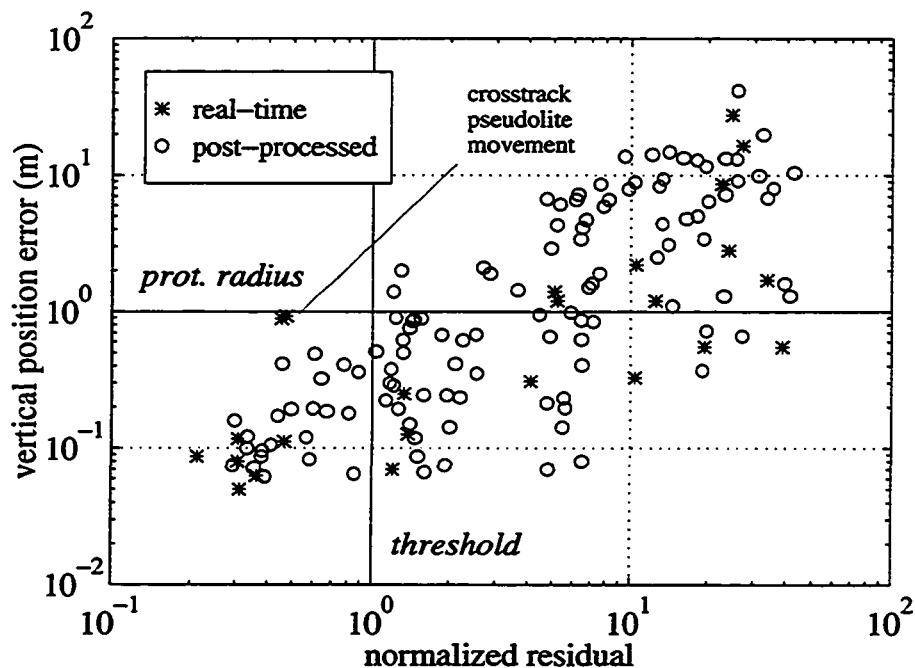
Figure 6.5: IBLS Autonomous Fault Detection Flight Test Results

magnitudes ranged from one to twenty cycles and occurred on the lowest elevation (i.e., lowest signal strength) satellite at randomly selected times during pseudolite overflight. Two of the remaining four approaches were done with the reference station moved to an unsurveyed site two meters away from nominal. The last two approaches were left as controls; no failures were induced.

The results of these 24 approaches are summarized in Figure 6.5 as a plot of vertical position error vs. the normalized residual. In these experiments, the vertical error was obtained to within approximately 5 cm (1σ) by comparison of the IBLS positioning output with the runway surface fit [Cohen, e] at aircraft touchdown. The magnitude of the least squares residual for the cycle ambiguity estimation process was normalized by the threshold R. This threshold was selected by table look-up in flight based on the number of redundant measurements collected. The probability of false alarm under

normal error conditions was set at 10^{-6} per approach. A protection radius of one meter—roughly consistent with the ILS vertical protection limit at the runway threshold—is also shown on the plot. The results show a relatively strong correlation between vertical position error and the residual in the presence of the failure modes induced. Using the assumed one meter protection limit, no missed detections were registered. The two real-time data points nearest the missed detection region are associated with vertical pseudolite movement. This result is not unexpected, since vertical and crosstrack pseudolite movements are not nominally detectable using a two pseudolite configuration (see Section 5.2.4.1).

A larger number of failures were also induced in the post-processing of data previously collected at Palo Alto Airport during Piper Dakota flight tests under normal error conditions. Post-processing was performed using the same measurement processing software used in the real-time airborne navigation system. A total of 116 post-processed approaches were executed. Forty approaches were run with cycle slips induced during pseudolite overflight; cycle slip magnitudes again ranged from 1 to 20 cycles. The remaining approaches were run with arbitrary pseudolite survey database errors (equivalent to pseudolite movement on the ground) of up to 2 m. These results are also included in Figure 6.5. In this case, the vertical error was quantified by comparison of the normal-error and failure-state post processed results. The larger sample of failures reinforces the observation that the residual and position error are correlated for the failure modes considered.

Although it is clear from Figure 6.5 that no missed detections were present, it is also evident that a significant number of false alarms were registered in both the real-time and post-processed results. It should be recalled, however, that the resulting continuity risk (as expressed in equation (6.8)) is the *product* of the probability of false alarm and the prior probability of occurrence of the failures that were deliberately induced.

(Continuity risk is assessed quantitatively in this manner in the next section.) Furthermore, it should also be noted that the ILS continuity and integrity requirements are consistent with false alarms occurring on the order of 10^3 times more frequently than missed detections.

The statistical significance necessary to demonstrate integrity risk on the order of 10^{-9} cannot, of course, be achieved through flight test or post-processing of existing flight test data. A much larger total number of approaches and wider range of system failures is necessary and can only be achieved through computer simulation.

6.3.2 Monte Carlo Simulation

A Monte Carlo simulation of aircraft approaches at SFO Runway 28 (see description in Section 6.2.2) was performed incorporating the failure models for cycle slips, spacecraft ephemeris errors, and pseudolite antenna mis-siting described in Sections 4.2.1 through 4.2.3. The goal of these simulated approaches was a quantitative assessment of the performance of RAIM fault detection during IBLS cycle ambiguity resolution with respect to three widely dissimilar failure modes.

The failure-state contribution to integrity risk in equation (6.7) can be mathematically expressed as

$$\sum_{i=1}^{3} P(\|\delta x\| > a, \|r\| < R \,|\, SF_i) P(SF_i) \tag{6.13}$$

where $P(SF_i)$ is the probability of system failure mode i, and the various failure modes have been assumed to be mutually exclusive. For the fault models under consideration, this assumption is valid because the likelihood of two independent failure types occurring on the same approach passing undetected is always less than 10^{-9}. Similarly, failure state continuity risk may be expressed as

$$\sum_{i=1}^{3} P(\|r\| > R \mid SF_i) P(SF_i).$$ (6.14)

Because the per-approach prior probability of each failure event $P(SF_i)$ is pre-specified in the fault models and does not exceed 10^{-5}, significantly fewer than 10^9 simulated aircraft approaches are needed to quantify integrity risks on the order of 10^{-9}.

A total of 2.5 million simulated aircraft approaches were executed; during each simulated approach one of the three failure types was induced. The navigation system response to these failures was assessed at cycle ambiguity estimation in terms of bubble-exit position error and residual magnitude. A representative output histogram matrix of these parameters is shown in Figure 6.6 as a contour plot of constant probability (including the prior probability of failure assumed in the models). For clarity of illustration, this figure summarizes the simulation results for only those approaches with
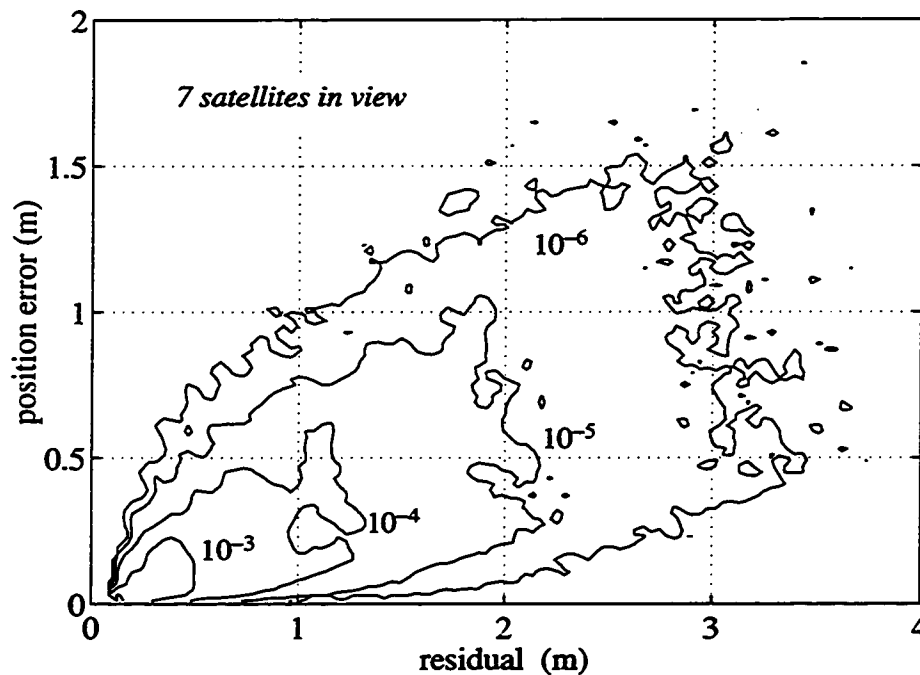
Figure 6.6: IBLS Autonomous Fault Detection Monte Carlo Simulation Results (Contours of Constant Probability)

seven satellites in view (although *all* constellation geometries were accounted for in the simulation). The contours shown in the figure are to be interpreted in the following manner: The indicated probability associated with each contour is simply the likelihood of an event lying outside that contour. It is clear from the figure, that the strong correlation between the position error and residual previously exhibited in the flight trials is also plainly evident in these simulation results.

The residual detection threshold R can, of course, be set using equation (5.14) to produce any desired continuity risk under normal error conditions. However, R must be chosen somewhat conservatively to accommodate for the additional continuity risk due to failures (see equations (6.8) and (6.14)). Figure 6.7, for example, illustrates that while the total continuity risk is dominated by fault-free conditions for small values of R, for large values of R, only failure states contribute. It is important to recall, however, that the potentially significant effects of fault isolation and removal provided by ground monitoring, redundant airborne sensor tracks, and RAIM have been deliberately excluded from consideration. These additional means of failure mitigation will obviously reduce the contribution of navigation failures to continuity risk (as suggested in equation (5.2)) and, thus, also reduce the need to relax the detection thresholds.

The complete results of the Monte Carlo simulation are summarized in Figure 6.8. This figure explicitly shows the inter-
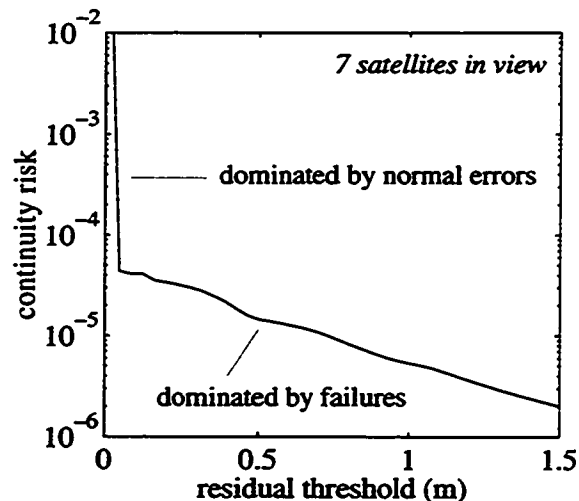
Figure 6.7: Continuity Risk vs. Detection Threshold
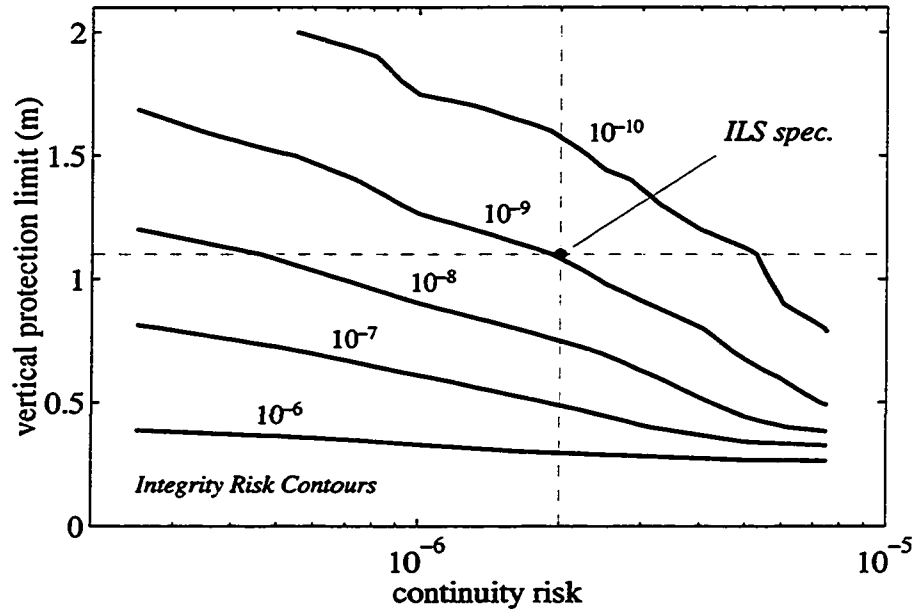
Figure 6.8: Navigation Performance Contours

relationship between integrity risk (6.7), continuity risk (6.9), and vertical protection limit a, in the presence of the three failure modes considered. As intuitively expected, integrity risk decreases as either the continuity risk or the protection radius are increased. Given that the three failure models implemented are the best available at present, the performance surface can be interpreted in an absolute sense as well: For example, for specified continuity risk of 2×10^{-6} and vertical protection limit of 1.1 m, the resulting integrity risk is less than 10^{-9}. This result is roughly consistent with the ILS specification.

As noted above, RAIM detection thresholds can be lowered significantly when the means for fault identification and removal is provided. To evaluate the expected performance in this context, additional simulations (with cycle slip failures only) were performed based on the assumption that continuity risk was due only to false alarms occurring under normal error conditions (in effect, *perfect* fault isolation and removal)

141

[Pervan, *b*]. In this idealized case, the results demonstrated that the ILS continuity risk and integrity risk requirements were achieved with a protection limit of less than 35 cm [Pervan, *b*]. The results of Monte Carlo simulations performed to date thus suggest that when a practical fault isolation and removal function is present, a protection limit on the order of 1/2 meter is not unrealistic.

Although the overall navigation availability (F) is limited to approximately 97% (unless geosynchronous spacecraft of terminal area pseudolites are assumed) by the need for RAIM during kinematic positioning after bubble exit, it is noteworthy that the availability of RAIM during cycle ambiguity resolution is constrained by only the requirement that four satellites are in view. This condition was met on 99.97% of the approaches.

CHAPTER 7

Conclusion

T he Global Positioning System can now provide the means for seamless aircraft navigation from takeoff to touchdown. While this goal has been highly motivating in the past, significant technical challenges have existed, the most difficult of which have been associated with ensuring *navigation integrity* during zero-visibility (Category III) precision approach and landing. This dissertation has directly addressed the issue of navigation integrity and has explored the inter-relationship of integrity with navigation system accuracy, continuity, and availability within the context of a unique navigation architecture founded on the use of carrier phase measurements and ground-based pseudolites.

7.1 Carrier Phase

While both *differential code* and *kinematic carrier phase* navigation architectures have been proposed for the precision landing application, the use of carrier phase has been demonstrated to provide two significant advantages over code:

1. Carrier phase positioning is the only established means by which GPS can provide the capability to meet the existing ILS accuracy specification of 2.1 ft (95%) at the 50 ft decision height. In fact, this specification can be met with a large margin.

2. The high precision of carrier phase measurements ($\sigma < 1$ cm) provides the leverage for Receiver Autonomous Integrity Monitoring (RAIM) in the sense that extremely tight detection thresholds may be set without incurring unacceptably high false alarm rates, thereby ensuring both high integrity and high continuity.

7.2 Pseudolites

In order to realize the benefits of carrier phase for precision approach and landing navigation, high-integrity cycle ambiguity resolution is necessary. The application of ground-based pseudolites provides such capability in that:

1. When placed under the aircraft approach path, pseudolites provide the means for explicit real-time cycle ambiguity estimation.

2. Pseudolites ensure the availability of redundant measurements for RAIM.

The Integrity Beacon Landing System (IBLS) architecture (Figure 1.3) has been developed to make full use of these advantages. In this architecture, the pseudolites are nominally placed at approximately the location of the existing ILS middle marker, corresponding to an aircraft overflight altitude of roughly 200 ft. After pseudolite overflight, cycle ambiguities are resolved, and subsequent centimeter-level kinematic position fixes are possible through the remainder of the approach, touchdown, and rollout.

144

7.3 Summary of Accomplishments

The focus of this research has been the investigation of the limits of satellite-based navigation performance for airborne applications. In particular, it has been demonstrated that carrier phase measurements from spacecraft and ground-based pseudolites can provide the basis for the highest level of GPS navigation integrity, accuracy, continuity and availability. Contributions were made in the following areas:

7.3.1 Measurement Processing

Prototype algorithms for airborne kinematic carrier phase processing, including high-speed nonlinear information smoothing algorithms for the first high-integrity airborne cycle ambiguity resolution, were developed, implemented, and flight-tested. Real-time algorithm performance was verified through an extensive battery of flight tests culminating in 110 successful automatic landings of a United Airlines Boeing 737-300. Smoother execution times of 1-1.5 sec have been consistently demonstrated in flight on a Pentium-based computer.

7.3.2 Error Sources and Fault Modes

The error sources for kinematic GPS navigation including receiver noise, multipath, latency, differential troposphere and ionosphere were characterized and quantified. Equivalent $1-\sigma$ spacecraft single difference phase range error was demonstrated to conservatively amount to approximately 1-1.5 cm. In addition, the failure classes associated with the navigation system ground, air, and space system segments were identified and their effect on navigation performance was assessed. The effect of pseudolite ground configuration was assessed with respect to fault detection capability and positioning accuracy in the presence of measurement error.

7.3.3 Fault Detection and Isolation

The theoretical framework of RAIM was generalized for application to carrier phase, cycle ambiguity resolution, and the detection of the wide range of navigation system failure scenarios. The considerable margin in precision offered by carrier phase measurements when compared to the desired protection limit (on the order of 1 m) ensures that the availability of RAIM is rather closely bounded simply by the fraction of time five or more satellites are in view. During cycle ambiguity resolution, the availability of redundant measurements for RAIM is guaranteed by the presence of ground-based pseudolites. When no additional augmentation is present (e.g., additional *terminal area* pseudolites or Inmarsat/WAAS geosynchronous spacecraft), carrier phase RAIM availability after pseudolite overflight is 97% with ILS integrity and continuity specifications (Table 1.1). In addition, a new carrier-phase parity space algorithm was introduced to provide the capability for fault identification and removal with the goal of maximizing navigation continuity subject to the constraint of maintaining high integrity.

7.3.4 Navigation Performance

Navigation performance was assessed, and the parametric inter-relationship between navigation system accuracy, integrity, continuity, and availability was quantitatively defined through analysis, simulation, and flight test. Simulation-based covariance analysis has demonstrated 95% positioning accuracy of approximately 5 cm, consistent with the results of an exhaustive series of flight tests taking place from January 1993 through the present. In addition, the performance of RAIM during cycle ambiguity resolution has been directly quantified through Monte Carlo simulation. Using representative mathematical models for three failure modes of diverse origin and character, ILS integrity and continuity specifications were met with a vertical protection limit of approximately 1 m without fault isolation and approaching 1/2 m when fault

146

isolation and removal is assumed. The performance of RAIM fault detection during cycle ambiguity resolution was experimentally verified by deliberately inducing real-time navigation failures during aircraft approach.

7.5 Recommendations for Future Work

A number of recommendations for future work are given below with regard to the practical application of high-integrity kinematic GPS navigation to aircraft precision approach and landing:

Augmentation. Although the availability of RAIM after cycle ambiguity resolution is quite good (97%), the desired availability of at least 99.9% cannot be achieved with the unaugmented GPS constellation. However, with ranging augmentation (Chapter 5) from either the three planned Inmarsat/WAAS geosynchronous spacecraft or a single terminal area (pulsed) pseudolite, navigation availability of 99.9% is possible. Although the former approach appears to be the most direct, as the Inmarsat/WAAS spacecraft are to be imminently (at the time of this writing) deployed, service at high latitude airports is precluded. With respect to the latter approach, which can be used to accommodate these high-latitude cases, recent flight test results with pulsed pseudolites [Lawrence, c] have been encouraging; however, the capability for a robust continuous ranging link during final approach has yet to be demonstrated. In particular, additional research and development is required in the areas of pseudolite and aircraft GPS antenna design.

RAIM Fault Identification and Removal. The real-time performance of the high-integrity parity-space fault isolation algorithm discussed in Chapter 5 is to be assessed through extensive simulation and flight test [Pervan, a]. The practical

147

benefit of RAIM-based fault isolation for precision approach and landing applications must be evaluated within the context of an envisioned integrity monitoring architecture including both ground monitoring and redundant airborne sensors. For example, the class of failures that are most uniquely suited to RAIM-based isolation are airborne GPS channel failures such as cycle slips. In this case, isolation and removal is generally also possible by voting between redundant airborne sensor tracks, but only at the expense of removing an entire GPS *receiver* from operation. Whether the resulting reduction in continuity risk obtained by removing the faulty channel *only* using RAIM justifies the associated reduction in navigation availability (because more ranging measurements are needed when RAIM-based fault identification and removal is implemented) is yet to be determined.

Ground Integrity Monitoring. Ground-based integrity monitoring for precision approach and landing has been covered only at the conceptual level in this dissertation, and in general it has not been widely addressed in the GPS research community. Thus, there exists a need for detailed architectures and algorithms for ground monitoring of the GPS signal during precision approach and landing. Particular emphasis is to be placed on failures such as satellite ephemeris error, SV 19-type anomalies (see Chapter 4), and rare-event ground multipath which will likely be the most difficult to detect using ground monitoring.

7.6 Closing

The use of carrier phase measurements and ground-based pseudolites, as combined in the IBLS architecture and documented in this thesis, has provided the first known means by which ILS performance specifications may be attained using differential satellite-based navigation.

Appendix A

Generalized Double Difference

The single difference carrier phase (2.19) with known baseline removed is

$$\varphi_{ik}^s \equiv \phi_{ik}^s + e_{ik}^{\mathsf{T}} x_k = \left(1 - \dot{\phi}_{ik}^r\right)\tau_k + N_i^s + v_{ik}^s \tag{A.1}$$

Applying (A.1) to two arbitrary spacecraft (designated as 1 and 2) yields after some rearrangement,

$$\begin{bmatrix} \varphi_{1k}^s \\ \varphi_{2k}^s \end{bmatrix} = \begin{bmatrix} 1 - \dot{\phi}_{1k}^r & 0 \\ 1 - \dot{\phi}_{2k}^r & 1 \end{bmatrix} \begin{bmatrix} \tau_k - N_1^s \\ N_2^s - N_1^s \end{bmatrix} + \begin{bmatrix} v_{1k}^s \\ v_{2k}^s \end{bmatrix} \tag{A.2}$$

where we note that the products $\dot{\phi}_{1k}^r N_1^s$ and $\dot{\phi}_{2k}^r N_1^s$ are negligible when $N_1^s < 200$ cycles. This is easily achieved by differential code-based initialization of the carrier phase measurements (see Section 3.2.1). Multiplying (A.2) by the inverse of the observation matrix and extracting the second row of the result gives

$$\varphi_{2k}^s - \frac{1 - \dot{\phi}_{2k}^r}{1 - \dot{\phi}_{1k}^r} \varphi_{1k}^s = N_2^s - N_1^s + v_{2k}^s - \frac{1 - \dot{\phi}_{2k}^r}{1 - \dot{\phi}_{1k}^r} v_{1k}^s. \tag{A.3}$$

Because the difference $N_2^s - N_1^s$ is an integer and the measurement errors v_{1k}^s and v_{2k}^s are much smaller than one wavelength, the generalized double difference can be defined as

$$\text{frac}\left(\varphi_{2k}^s - \frac{1 - \dot{\phi}_{2k}^r}{1 - \dot{\phi}_{1k}^r} \varphi_{1k}^s\right) \approx v_{2k}^s - v_{1k}^s \tag{A.4}$$

where $\text{frac}(x) \equiv x - \text{round}(x)$.

Appendix B

Adding and Removing States with the Iterated Information Smoother

Adding a Satellite or Pseudolite. The principle behind adding a new satellite or pseudolite is exceedingly simple in the information smoother. If a new ranging source is acquired by the receivers between measurement epochs k and $k+1$, and no prior knowledge of the new cycle ambiguity is assumed, the new unknown cycle ambiguity state is appended at the end of the state vector, and the information matrix for epoch k is simply appended with zeros as follows

$$\hat{P}_k^{new} = \lim_{\mu \to \infty} \begin{bmatrix} \hat{P}_k^{old} & 0 \\ 0 & \mu \end{bmatrix} \quad \Rightarrow \quad \hat{S}_k^{new} = \begin{bmatrix} \hat{S}_k^{old} & 0 \\ 0 & 0 \end{bmatrix}. \tag{B.1}$$

Similarly, the information vector consistent with (B.1) is then

$$\hat{y}_k^{new} = \begin{bmatrix} \hat{y}_k^{old} \\ 0 \end{bmatrix}. \tag{B.2}$$

This process can be repeated as necessary to accommodate the case where more than one new ranging source is acquired between measurement epochs k and $k+1$. The processing of the measurements collected at epoch $k+1$ can now continue in the nominal manner described in Section 3.3.2.2.

Removing a Satellite or Pseudolite. The mechanics of removing a satellite or pseudolite is somewhat more involved in that matrix partitioning and block inversion is necessary. Consider first the case where the last cycle ambiguity state is to be removed from the state vector before processing of epoch $k+1$. In this case, the last column and row of the covariance matrix for epoch k must be deleted. Using the definition of the information matrix expressed in block partitioned form, we can write

$$\hat{P}_k^{old} = \begin{bmatrix} \hat{P}_k^{new} & \bullet \\ \bullet & \bullet \end{bmatrix} = \begin{bmatrix} \hat{S}_k^1 & \hat{S}_k^{12} \\ \hat{S}_k^{12\,T} & \hat{s}_k^2 \end{bmatrix}_{old}^{-1} , \qquad (B.3)$$

where $\hat{s}_k^2$ is a non-zero scalar. Applying the matrix block inversion identity, we obtain

$$\hat{S}_k^{new} = \left[\hat{P}_k^{new}\right]^{-1} = \left[\hat{S}_k^1 - \hat{S}_k^{12}\,\hat{S}_k^{12\,T}/\hat{s}_k^2\right]_{old} . \qquad (B.4)$$

An analogous procedure for the information vector produces

$$\hat{y}_k^{new} = \left[\hat{y}_k^1 - \hat{S}_k^{12}\,\hat{y}_k^2/\hat{s}_k^2\right]_{old} . \qquad (B.5)$$

The case where an arbitrary satellite i (not necessarily the one corresponding to the last cycle ambiguity state) is to be deleted is handled as follows: Remove the i-th row and column of the information matrix and i-th element of the information vector; the remaining matrix and vector are defined to be $\hat{S}_k^1$ and $\hat{y}_k^1$, respectively. Element (i,i) of the information matrix is defined to be $\hat{s}_k^2$, and element i of the information vector is defined to be $\hat{y}_k^2$. Finally, $\hat{S}_k^{12}$ is defined to be column i of the information matrix with element (i,i) removed. Using these definitions, equations (B.4-5) can be applied directly to obtain the new information matrix and vector.

Appendix C

Properties of the Least-Squares Residual

꙰

Probability Distribution. Under normal error conditions

$$r = \left(I - HH^+ \right) \delta z \tag{C.1}$$

where δz is gaussian and i.i.d. with zero mean and standard deviation σ_z. It can be verified by direct multiplication that the $n \times n$ matrix $I - HH^+$ is idempotent; thus, premultiplication of both sides of (C.1) by their respective transposes yields

$$\| r \|^2 = \delta z^\tau \left(I - HH^+ \right) \delta z. \tag{C.2}$$

The singular value decomposition (SVD) of the $n \times m$ observation matrix can be expressed as

$$H = U \begin{bmatrix} S \\ 0 \end{bmatrix} V^\tau \tag{C.3}$$

where $\begin{bmatrix} S \\ 0 \end{bmatrix}$ is $n \times m$, S is $m \times m$ and diagonal, U is $n \times n$, V is $m \times m$, and both U and V are unitary. Expanding the projection matrix $I - HH^+$ using the definition (5.6) gives

$$I - HH^+ = I - H \left(H^\tau H \right)^{-1} H^\tau. \tag{C.4}$$

Substituting (C.3) into this expression we obtain

152

$$I - HH^+ = I - U \begin{bmatrix} S \\ 0 \end{bmatrix} V^\tau \left(V [S \ 0] U^\tau U \begin{bmatrix} S \\ 0 \end{bmatrix} V^\tau \right)^{-1} V [S \ 0] U^\tau. \qquad (C.5)$$

This expression can be simplified by using the fact that $U^\tau U = I_n$. Executing the innermost matrix multiplication and performing the indicated matrix inversion produces

$$I - HH^+ = I - U \begin{bmatrix} S \\ 0 \end{bmatrix} V^\tau V S^{-2} V^\tau V [S \ 0] U^\tau, \qquad (C.6)$$

This expression can be further simplified using $V^\tau V = I_m$ to obtain

$$I - HH^+ = I - U \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} U^\tau. \qquad (C.7)$$

Finally, using $U^\tau U = I_n$ again, the projection matrix can be written as

$$I - HH^+ = U \begin{bmatrix} 0 & 0 \\ 0 & I_{n-m} \end{bmatrix} U^\tau. \qquad (C.8)$$

Substituting the result (C.8) into equation (C.2) and dividing both sides by σ_z^2 yields

$$\frac{\|r\|^2}{\sigma_z^2} = \delta\tilde{z}^\tau \begin{bmatrix} 0 & 0 \\ 0 & I_{n-m} \end{bmatrix} \delta\tilde{z}, \qquad (C.9)$$

where the definition $\delta\tilde{z} \equiv U^\tau \delta z / \sigma_z$ has been used. Because U is unitary (orthonormal), the random vector $\delta\tilde{z}$ is gaussian and i.i.d. with zero mean and unit variance. The normalized least squares residual can thus be expressed as the sum of squares of $n - m$ zero-mean/unit-variance gaussian random variables as follows:

$$\frac{\|r\|^2}{\sigma_z^2} = \delta\tilde{z}_{m+1}^2 + \cdots + \delta\tilde{z}_n^2. \qquad (C.10)$$

153

Expression (C.10) is the definition of a χ^2 random variable with $n-m$ degrees of freedom. [CRC]

Independence of the Residual and Estimate Error. Although the SVD can also be used to demonstrate the independence of the residual and the estimate error, in order to maximize physical insight in the discussion an alternative approach is chosen. First, the observation matrix (5.3) is written as

$$z = Hu + \Delta \tag{C.11}$$

where

$$\Delta \equiv \delta z - \delta H u. \tag{C.12}$$

The error vector Δ can be expressed as the sum

$$\Delta \equiv \Delta_H + \Delta_\perp, \tag{C.13}$$

where Δ_H is the vector component of Δ that is a linear combination of the columns of H (i.e., $\Delta_H \in Range\{H\}$) and $\Delta_\perp$ is the vector component of Δ orthogonal to the columns of H.

The least squares solution to (C.11) is

$$\hat{u} = \left(H^T H\right)^{-1} H^T z. \tag{C.14}$$

Substituting the right-hand side of (C.11) into (C.14) and using the definition (C.13) yields

$$\hat{u} = \left(H^T H\right)^{-1}\left(H^T H\right)u + \left(H^T H\right)^{-1} H^T \Delta_H + \left(H^T H\right)^{-1} H^T \Delta_\perp. \tag{C.15}$$

Because $\Delta_\perp$ is orthogonal to the columns of H the product $H^T \Delta_\perp$ is zero. The result is then

$$\delta u \equiv \hat{u} - u = \left(H^T H\right)^{-1} H^T \Delta_H . \qquad (C.16)$$

Now using the definition of the least squares residual (5.11) and equations (C.11) and (C.15) yields

$$r = Hu + \Delta_H + \Delta_\perp - Hu - H\left(H^T H\right)^{-1} H^T \Delta_H \qquad (C.17)$$

where we have used the fact that $H^T \Delta_\perp$ is zero. Since $\Delta_H \in Range\{H\}$, the substitution

$$\Delta_H = Hy, \quad y \in R^m \qquad (C.18)$$

can be made into equation (C.17) to obtain

$$r = Hy + \Delta_\perp - Hy \qquad (C.19)$$

and finally

$$r = \Delta_\perp . \qquad (C.20)$$

From equations (C.16) and (C.20), it is clear that the estimate error δu and the residual r are derived from orthogonal components of the error vector Δ.

Appendix D

Conditional Parity Vector Distribution

Given a failure on channel i with magnitude γ, the conditional density of the parity vector p is given by equation (5.27). Now if γ is uniformly distributed in the range $\left[-\frac{M}{2}, \frac{M}{2}\right]$, then the parity vector probability density given a failure on channel i can be expressed as

$$f(p|FM_i) = \int_{-\frac{M}{2}}^{\frac{M}{2}} \frac{1}{(2\pi)^{\frac{n-m}{2}}} \exp\left[-\tfrac{1}{2}(p - \gamma Q_i q_i)^{\mathsf{T}}(p - \gamma Q_i q_i)\right] \frac{1}{M} d\gamma. \qquad (D.1)$$

Let us define

$$p_\perp \equiv p - (p^{\mathsf{T}} q) q_i \qquad (D.2)$$

as the vector component of p orthogonal to the parity space failure direction q_i, and

$$p_\parallel \equiv p^{\mathsf{T}} q_i \qquad (D.3)$$

as the vector component of p in the failure direction q_i. Substituting these definitions into equation (D.1) results in

$$f(p|FM_i) = \frac{1}{(2\pi)^{\frac{n-m}{2}}} \exp\left[-\tfrac{1}{2} p_\perp^{\mathsf{T}} p_\perp\right] \cdot \int_{-\frac{M}{2}}^{\frac{M}{2}} \exp\left[-\tfrac{1}{2}\left(\|p_\parallel\| - \gamma Q_i\right)^{\mathsf{T}}\left(\|p_\parallel\| - \gamma Q_i\right)\right] \frac{1}{M} d\gamma. \qquad (D.4)$$

Making the substitution

$$w \equiv \|p_\parallel\| - \gamma Q_i \qquad (D.5)$$

into (D.4) yields

$$f(p \mid FM_i) = \frac{1}{(2\pi)^{\frac{n-m}{2}}} \frac{\exp\left[-\frac{1}{2} p_\perp^\mathsf{T} p_\perp\right]}{Q_i M} \cdot \int_{|p_{\shortparallel}|-\frac{M}{2}Q_i}^{|p_{\shortparallel}|+\frac{M}{2}Q_i} \exp\left(-\frac{1}{2} w^2\right) dw. \qquad (D.6)$$

Now if $\|p\| < B$, for any finite $B > 0$, and $M \gg 2B/Q_i$, the integral in (D.6) rapidly

approaches $\sqrt{2\pi}$, so that the conditional density function asymptotically approaches

$$f(p \mid FM_i) \rightarrow \frac{1}{(2\pi)^{\frac{n-m-1}{2}}} \frac{\exp\left[-\frac{1}{2} p_\perp^\mathsf{T} p_\perp\right]}{Q_i M} . \qquad (D.7) \cdot$$

Finally, substituting (D.7) into equation (5.25) yields the desired result (5.29).

Appendix E

Least-Squares Measurement Downdate

Suppose that after positioning has been performed using all available channels (satellites), it is desired to remove the i-th channel (perhaps after a fault has been isolated) and re-estimate the position. This procedure is called 'downdating.' It is shown below that if the original (all-channel solution) covariance matrix has been stored, downdating the all-channel solution can be performed much more efficiently than by directly repeating the entire estimation process with the i-th channel removed.

We consider first the converse situation of using the well known sequential least-squares (Kalman) update to incorporate channel i into a position solution which has already included information from all other channels. In this case, we have

$$P_i^{-1} = P^{-1} + h_i h_i^{\mathsf{T}} \tag{E.1}$$

$$P_i^{-1} x_i = P^{-1} x + h_i z_i \tag{E.2}$$

where h_i^{T} is the i-th row of the observation matrix H, z_i is the i-th measurement, P_i and x_i are the covariance matrix and state estimate when measurement i is included, and P and x are the covariance matrix and state estimate when channel i is not used. It has been implicitly assumed in (E.1) and (E.2) that measurements on all channels are i.i.d. and that covariances have been normalized by the measurement variance σ_z^2. These simplifications are not required, but are implemented for the purposes of clarity in the ensuing derivation.

Solving (E.2) for x and using (E.1) to eliminate P_i from the result yields

$$x = x_i - Ph_i\left(z_i - h_i^\mathsf{T} x_i\right).$$

(E.3)

Noting that

$$P_i \equiv \left(H^\mathsf{T} H\right)^{-1}$$

(E.4)

and substituting into (E.1) produces the following expression for P:

$$P = \left[H^\mathsf{T} H - h_i h_i^\mathsf{T}\right]^{-1}.$$

(E.5)

Using a variant of the Matrix Inversion Lemma [Kailath], the right-hand side of (E.5) can be expanded to produce

$$P = \left(H^\mathsf{T} H\right)^{-1} + \frac{\left(H^\mathsf{T} H\right)^{-1} h_i h_i^\mathsf{T} \left(H^\mathsf{T} H\right)^{-1}}{1 - h_i^\mathsf{T} \left(H^\mathsf{T} H\right)^{-1} h_i}.$$

(E.6)

Substituting this expression into (E.3) and simplifying yields

$$x = x_i - \frac{\left(H^\mathsf{T} H\right)^{-1} h_i}{1 - h_i^\mathsf{T} \left(H^\mathsf{T} H\right)^{-1} h_i}\left(z_i - h_i^\mathsf{T} x_i\right)$$

(E.7)

which is identical to equation (5.33). Note that the most demanding computation involved in the numerical execution of (E.7) is merely the multiplication of the pre-assembled covariance matrix $\left(H^\mathsf{T} H\right)^{-1}$ and the vector h_i.

References

[Abramowitz]   M. Abramowitz and I. A. Stegun, ed., *Handbook of Mathematical Functions*, Dover Publications, New York, 1972.

[Ashley]   H. Ashley, Department of Aeronautics and Astronautics, Stanford University, *personal communication*, January 1994.

[AC120-28C]   Federal Aviation Administration, *Advisory Circular 120-28C:* "Criteria for Approval of Category III Landing Weather Minima," Washington, DC, 9 March 1984.

[Beser]   J. Beser and B. W. Parkinson, "The Application of NAVSTAR Differential GPS in the Civilian Community," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 29, No. 2, Summer 1982.

[Blomenhofer]   H. Blomenhofer and G. W. Hein, "Tropospheric Effects in DGPS and their Removal During Precision Approach and Automatic Landing," International Symposium on Precision Approach and Landing, Braunschweig, Germany, February 1995.

[Braasch]   M. S. Braasch, A. Fink, K. Duffus, "Improved Modeling of GPS Selective Availability," ION National Technical Meeting, San Francisco, CA, January 1993.

[Braff]   R. Braff, "Alarm Limits for Local-DGPS Integrity Monitoring," ION National Technical Meeting, Anaheim, CA, January 1995.

[Brown, *a*] R. G. Brown and P. W. McBurney, "Self-Contained GPS Integrity Check Using Maximum Solution Separation," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 35, No. 1, Spring 1988.

[Brown, *b*] R. G. Brown, "A Baseline RAIM Scheme and a Note on the Equivalence of Three RAIM Methods," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 39, No. 3, Fall 1992.

[A. Brown] A. Brown, F. Van Diggelen, and C. LaBerge, "Test Results of a GPS/Pseudolite Precision Approach and Landing System," ION GPS-93, Salt Lake City, Utah, September 1993.

[Bryson] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*, Hemisphere, New York, 1975.

[Carlson] A. B. Carlson, *Communication Systems*, McGraw Hill, Inc., New York, NY, 1986.

[Chao] Y. C. Chao and B. W. Parkinson, "The Statistics of Selective Availability and Its Effect on Differential GPS," ION GPS-93, Salt Lake City, Utah, September 1993.

[Chin] G. Y. Chin, J. H. Kraemer, and R. G. Brown, "GPS RAIM: Screening Out Bad Geometries Under Worst-Case Bias Conditions," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 39, No. 4, Winter 1992-93.

[Chou] H.-T. Chou, "An Adaptive Correction Technique for Differential Global Positioning System," Stanford University Ph.D. Dissertation, Department of Aeronautics and Astronautics, June 1991.

[Cobb, *a*] H. S. Cobb, C. E. Cohen, and B. W. Parkinson, "Theory and Design of Pseudolites," ION National Technical Meeting, San Diego, CA, January 1994.

[Cobb, *b*] H. S. Cobb, Stanford University Ph.D. Dissertation, Department of Aeronautics and Astronautics, 1996.

[Cohen, *a*] C. E. Cohen, "Attitude Determination Using GPS," Stanford University Ph.D. Dissertation, Department of Aeronautics and Astronautics, December 1992.

[Cohen, *b*] C. E. Cohen, B. S. Pervan, H. S. Cobb, D. G. Lawrence, J. D. Powell, and B. W. Parkinson, "Real-Time Cycle Ambiguity Resolution using a Pseudolite for Precision Landing of Aircraft with GPS," Second International Conference on Differential Satellite Navigation Systems (DSNS-93), March 30 - April 2, 1993, Amsterdam, The Netherlands.

[Cohen, *c*] C. E. Cohen, B. W. Parkinson, and B. D. McNally, "Flight Tests of Attitude Determination Using GPS Compared Against an Inertial Navigation Unit," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 41, No. 1, Spring 1994.

[Cohen, *d*] C. E. Cohen, B. S. Pervan, H. S. Cobb, D. G. Lawrence, J. D. Powell, and B. W. Parkinson, "Achieving Required Navigation Performance using GNSS for Category III Precision Landing," DSNS-94, London, UK, April 1994.

[Cohen, *e*] C. E. Cohen, B. S. Pervan, D. G. Lawrence, H. S. Cobb, J. D. Powell, and B. W. Parkinson, "Real-Time Flight Testing Using Integrity Beacons for GPS Category III Precision Landing," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 41, No. 2, Summer 1994.

[Cohen, f] C. E. Cohen et al., "Flight Test Results of Autocoupled Approaches Using GPS and Integrity Beacons," ION GPS-94, Salt Lake City, Utah, September 1994.

[Cohen, g] C. E. Cohen et al., "Automatic Landing of a 737 using GNSS Integrity Beacons," International Symposium on Precision Approach and Landing (ISPA 95), Braunschweig, Germany, February 1995.

[Counselman] C. C. Counselman and I. I. Shapiro, "Miniature Interferometer Terminals for Earth Surveying," Bulletin Geodesique, Vol. 53, No. 2, pp. 139-163, 1979.

[CRC] W. H. Beyer, ed., *CRC Standard Mathematical Tables*, 27th Edition, CRC Press, Boca Raton, FL, 1984.

[Daly] K. C. Daly, E. Gai, and J. V. Harrison, "Generalized Likelihood Test for FDI in Redundant Sensor Configurations," *Journal of Guidance and Control*, Vol. 2, No. 1, Jan.-Feb. 1979.

[Davis] J. M. Davis and R. J. Kelly, "RNP Tunnel Concept for Precision Approach with GNSS Application," Institute of Navigation 49th Annual Meeting, Cambridge, MA, June 1993.

[DO-178B] RTCA Special Committee 167, *Software Considerations in Airborne Systems and Equipment Certification*, Document No. RTCA/DO-178B, December 1, 1992.

[Durand] J-M. Durand and A. Caseau, "GPS Availability, Part II: Evaluation of State Probabilities for 21 and 24 Satellite Constellations," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 37, No. 3, Fall 1990.

[Enge] P. Enge, T. Walter, S. Pullen, Y.-C. Chao, and Y.-J. Tsai, "Wide Area Augmentation of GPS," *Proceedings of the IEEE*, July 1996.

163

[Euler]   R. R. Hatch and H.-J. Euler, "Comparison of Several AROF Kinematic Techniques," Proceedings of ION GPS-94, Salt Lake City, Utah, September 1994.

[Fernow]   J. P. Fernow, "FAA SOIT Consensus on Requirements for the Local Area Augmentation System (LAAS) ORD," Presentation to RTCA Special Committee 159 Working Group 4, Washington DC, April 19, 1995.

[Francisco]   S. G. Francisco, "GPS Operational Control Segment," contribution to *The Global Positioning System: Theory and Applications*, B. W. Parkinson, J. J. Spilker, Jr., P. Axelrad, P. Enge, editors, AIAA Progress in Aeronautics and Astronautics Volume 163-4, Washington, DC, 1996.

[Frank]   P. M. Frank, "Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy–A Survey and Some New Results," *Automatica*, Vol. 26, No. 3, 1990.

[Frodge]   S. L. Frodge, S. R. Deloach, B. Remondi, D. Lapucha, and R. A. Barker, "Real-Time On-the-Fly Kinematic GPS System Results," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 41, No. 2, Summer 1994.

[Gelb]   A. Gelb, ed., *Applied Optimal Estimation*, M.I.T. Press, Cambridge, MA, 1974.

[Girts]   D. Girts, Honeywell Corp., Presentation to Boeing GPS Landing System (GLS) Requirements Group, Seattle, WA, 14 February 1995.

[Goad]   C. C. Goad, "Optimal Filtering of Pseudoranges and Phases from Single-Frequency GPS Receivers," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 37, No. 3, Fall 1990.

[Gower] A. G. Gower, "Putting a Number on GPS Integrity: The IBM GPS Integrity Study for the DoD," ION GPS-91, Albuquerque, NM, September 1991.

[Green] G. B. Green, P. D. Massatt, and N. W. Rhodus, "The GPS 21 Primary Satellite Constellation," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 36, No. 1, Spring 1989.

[Hatch, *a*] R. R. Hatch, "The Synergism of Code and Carrier Measurements," Proceedings of the Third International Geodetic Symposium on Satellite Doppler Positioning, Las Cruces, New Mexico, February 1982.

[Hatch, *b*] R. R. Hatch, "Instantaneous Ambiguity Resolution," KIS Symposium, Banff, Canada, September 1990.

[Haupt] G. Haupt, N. Kasdin, G. Keiser, and B. Parkinson, "An Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation," AIAA Guidance, Navigation, and Control Conference, Baltimore, MD, August 1995.

[Hundley] W. Hundley, S. Rowson, G. Courtney, V. Wullschleger, R. Velez, R. Benoist, P. O'Donnell, "FAA-Wilcox Electric Category IIIB Fesability Demonstration Program Flight Test Results," ION GPS-95, Palm Springs, CA, September 1995.

[Hwang] P. Y. C. Hwang and R. G. Brown, "GPS Navigation: Combining Pseudorange with Continuous Carrier Phase Using a Kalman Filter," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 37, No. 2, Summer 1990.

[ICAO] International Civil Aviation Organization (ICAO), *International Standards, Recommended Practices and Procedures for Air Navigation Services -- Annex 10*, April 1985.

[ICD-200]   Interface Control Document (ICD) GPS-200, Revision B, Rockwell International, 1987.

[Jazwinski] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[Kailath] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[Kalafus] R. M. Kalafus, A. J. Van Dierendonck, and N. A. Pealer, "Special Committee 104 Recommendations for Differential Service," *Global Positioning System, Vol. III*, Papers Published in NAVIGATION, Reprinted by the ION, Washington D. C., 1986.

[Kayton, *a*] M. Kayton and W. R. Fried, ed., *Avionics Navigation Systems*, John Wiley & Sons, New York, 1969.

[Kayton, *b*] M. Kayton, ed., *Navigation: Land, Sea, Air, and Space*, IEEE Press, New York, 1990.

[Kee] C. Kee "Wide Area Differential GPS (WADGPS)," Stanford University Ph.D. Dissertation, Department of Aeronautics and Astronautics, December 1993.

[Kelly] R. J. Kelly and J. M. Davis, "Required Navigation Performance (RNP) for Precision Approach and Landing with GNSS Application," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 41, No. 1, Spring 1994.

[Klass] P. J. Klass, "Inmarsat Plan Spurs GPS Debate," *Aviation Week and Space Technology*, July 26, 1993.

[Klein] D. Klein and B. W. Parkinson, "The Use of Pseudo-Satellites for Improving GPS Performance," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 31, No. 4, Winter 1984-85.

[Klobuchar] J. A. Klobuchar, "Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-23, No. 3, May 1987.

[Lawrence, *a*] D. G. Lawrence, B. S. Pervan, C. E. Cohen, H. S. Cobb, J. D. Powell, and B. W. Parkinson, "Real-Time Architecture for Kinematic GPS Applied to the Integrity Beacon Landing System," ION 51st Annual Meeting, Colorado Springs, Colorado, June 1995.

[Lawrence, *b*] D. G. Lawrence, B. S. Pervan, H. S. Cobb, C. E. Cohen, J. D. Powell, and B. W. Parkinson, "Augmenting Kinematic GPS With a Pulsed Pseudolite to Improve Navigation Performance," ION National Technical Meeting, January 1996, Santa Monica, CA.

[Lawrence, *c*] D. G. Lawrence, Stanford University Ph.D. Dissertation, Department of Aeronautics and Astronautics, 1996.

[Lear] W. M. Lear, M. N. Montez, L. M. Rater, "The Effect of Selective Availability on Orbit Space Vehicles Equipped with SPS GPS Receivers," ION GPS-92, Albuquerque, NM, September 1992.

[LIP] *Stanford University Report to the LAAS Integrity Panel (LIP)*, January 1996, Washington, DC

[MacDoran] P. MacDoran, "Satellite Emission Radio Interferometric Earth Surveying (SERIES)–GPS Geodetic System," Bulletin Geodesique, Vol. 53, No. 2, pp. 117-138, 1979.

[Masson] A. Masson, AA 290 Graduate Research Final Report, Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, April 1994.

[Markin] K. Markin and C. Shively, "A Position-Domain Method for Ensuring Integrity of Local Area Differential GPS (LDGPS)," ION 51st Annual Meeting, Colorado Springs, CO, June 1995.

[Miller] D. C. Miller, Boeing Commercial Airplane Group, Presentation to Boeing GPS Landing System (GLS) Requirements Group, Seattle, WA, 26 June 1995.

[Nordwall] B. D. Nordwall, "Filter Center" Column, Aviation Week and Space Technology, August 30, 1993.

[ORD] Operational Requirements Document: Local Area Augmentation System, Satellite Navigation Program Office, Federal Aviation Administration, February 28, 1995.

[Paielli] R. A. Paielli, B. D. McNally, R. E. Bach, and D. N. Warner, "Carrier Phase Differential GPS for Approach and Landing: Algorithms and Preliminary Results," ION GPS-93, Salt Lake City, Utah, September 1993.

[Parkinson, a] B. W. Parkinson and P. Axelrad, "Autonomous Integrity Monitoring Using the Pseudorange Residual," NAVIGATION: Journal of the Institute of Navigation, Vol. 35, No. 2, Summer 1988.

[Parkinson, *b*] B. W. Parkinson and K. T. Fitzgibbon, "Aircraft Automatic Landing Systems using GPS," *Journal of Navigation*, Vol. 42, No. 1, January 1989.

[Parkinson, *c*] B. W. Parkinson, J. J. Spilker, P. Axelrad, and P. Enge (editors), *The Global Positioning System: Theory and Applications*, AIAA Progress in Aeronautics and Astronautics Volume 163-4, Washington, DC, 1996.

[Pervan, *a*] B. S. Pervan, C. E. Cohen, B. W. Parkinson, "Parity Space Methods for Autonomous Fault Detection and Exclusion using GPS Carrier Phase," IEEE Position, Location, and Navigation Symposium (PLANS '96), Atlanta, GA, April 1996.

[Pervan, *b*] B. S. Pervan, C. E. Cohen, and B. W. Parkinson, "Integrity in Cycle Ambiguity Resolution for GPS-Based Precision Landing," Third International Conference on Differential Satellite Navigation Systems (DSNS-94), London, UK, April 1994.

[Phlong] W. S. Phlong and B. D. Elrod, "Availability Characteristics of GPS and Augmentation Alternatives," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 40, No. 4, Winter 1993-94.

[Remondi] B. W. Remondi, "Performing Centimeter-level Surveys in Seconds with GPS Carrier Phase: Initial Results," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 32, No. 4, Winter 1985-86.

[Romrell] G. Romrell, R. Brown, G. Johnson, and D. Kaufman, "FAA/FEDSIM – E-Systems Cat IIIB Feasibility Demonstration Flight Test Preliminary Results," ION GPS-95, Palm Springs, CA, September 1995.

[Rowson]   S. Rowson, G. Courtney, and R. Hueschen, "Performance of Category IIIB Automatic Landings Using C/A Code Tracking Differential GPS," ION National Technical Meeting, January 1994, San Diego, CA.

[SCAT-I]   Minimum Operations Performance Standards, *DGNSS Instrument Approach System: Special Category I (SCAT-I)*, RTCA Draft Paper No. 116-93/SC-159-421, March 5, 1993.

[Skillicorn]   G. N. Skillicorn, "The FAA's Program to Sustain and Improve ILS Capability In a Rapidly Changing Environment," International Symposium on Precision Approach and Landing (ISPA-95), Braunschweig, Germany, February 21-24, 1995.

[Spilker, a]   J.J. Spilker, *Digital Communications by Satellite*, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[Spilker, b]   J. J. Spilker, Jr., "GPS Signal Structure and Performance Characteristics," Global Positioning System papers published in *NAVIGATION*, Vol. I, The Institute of Navigation, Washington DC, 1980.

[Spilker, c]   J. J. Spilker, Jr., "Tropospheric Effects on GPS," contribution to *The Global Positioning System: Theory and Applications*, B. W. Parkinson, J. J. Spilker, Jr., P. Axelrad, P. Enge, editors, AIAA Progress in Aeronautics and Astronautics Volume 163-4, Washington, DC, 1996.

[SPS]   "Global Positioning System Standard Positioning Service Signal Specification (Second Edition)," June 2, 1995.

[Sturza]   M. A. Sturza, "Navigation System Integrity Monitoring Using Redundant Measurements," NAVIGATION: *Journal of the Institute of Navigation*, Vol. 35, No. 4, Winter 1988-89.

[Teasley]   S. P. Teasley, W. M. Hoover, and C. R. Johnson, "Differential GPS Navigation," Position, Navigation, and Location Symposium (PLANS 80), Atlantic City, New Jersey, December 1980.

[Van Graas, *a*]   F. van Graas, D. Diggle, R. Hueschen, "Interferometric GPS Flight Reference/Autoland System," ION GPS-93, Salt Lake City, Utah, September 1993.

[Van Graas, *b*]   F. Van Graas and S.-W. Lee, "High-Accuracy Differential Positioning for Satellite-Based Systems Without Using Code-Phase Measurements," Proceedings of the ION National Technical Meeting, Anaheim, CA, January 1995.

[Van Graas, *c*] F. van Graas, D. Diggle, M. Uijt, V. Wullschleger, R. Velez, D. Lamb, M. Dimeo, G. Kuehl, and R. Hilb, "FAA/Ohio University/United Parcel Service DGPS Autoland Flight Test Demonstration," ION GPS-95, Palm Springs, CA, September 1995.

[Van Trees]   H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, John Wiley & Sons, New York, 1968.

[Walter]   T. Walter and P. Enge, "Weighted RAIM for Precision Approach," ION GPS-95, Palm Springs, CA, September 1995.

[Willsky, *a*]   A. S. Willsky, "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica*, Vol. 12, pp. 601-611, 1976.

[Willsky, *b*] A. S. Willsky and H. L. Jones, "A Generalized Likelihood Ratio Approach to State Estimation in Linear Systems Subject to Abrupt Changes," *IEEE Transactions on Automatic Control*, AC-21, February 1976.

[WMOPS] RTCA Special Committee 159, *Minimum Operation Performance Standards for Airborne Equipment using GPS/WAAS*, Document No. RTCA/DO-229, January 16, 1996.

[Zumberge] J. F. Zumberge and W. I. Bertiger, "Ephemeris and Clock Navigation Message Accuracy," contribution to *The Global Positioning System: Theory and Applications*, B. W. Parkinson, J. J. Spilker, Jr., P. Axelrad, P. Enge, editors, AIAA Progress in Aeronautics and Astronautics Volume 163-4, Washington, DC, 1996.