Extensive Excerpts: GNSS Aided Navigation and Tracking

Two conditions affecting GNSS provided motivation for this release. First there is reduced SV availability (*GPS World*, June 2009, *p*.6 and *p*.8). Although intermittent and temporary, the prospective duration is appreciable – and inevitably accompanied by uncertainty. Subsequently there will be a need to combine observations from different constellations.

The first condition will stress existing operations, and differencing used with the second will introduce correlations. This book presents original practical solutions to both – not presented in any other book – rigorously validated analytically and in flight test with a low-cost IMU —

"*leaves no stone unturned when it comes to optimizing performance*" - Prof. Frank van Graas

"*teeming with insights that are hard to find or unavailable elsewhere*" - Dr. Chris Hegarty

"*Unique ••• complete and rigorous ••• fully exploits the use of GPS carrier phase ••• which is not fully exploited in previous work*" - G. Jeffrey Geier

The unique treatment of carrier phase just noted is usage as sequential changes over 1-sec, which can be everywhere ambiguous and repeatedly interrupted (*i.e.*, there is no need for unambiguous cycle counts ***nor***, with FFT-based deep integration cited in the references, for track loop continuity). This makes full use of partial – even fragmented – data that current operational systems either discard or cannot reliably access. Any and all 1-sec phase changes surviving rigorous integrity criteria are fed forward, to provide streaming velocity. Integration of this precise (1-cm/sec RMS) velocity vector is then simply adjusted by the available pseudoranges (which can likewise be highly intermittent). ***A subtle but important geometry benefit***: 1-sec phase changes need no mask; with no scintillation or multipath – which would trigger data edit – maximum 1-sec changes in propagation are shown to be of order 1-cm. All of this was verified with data from an hour of flight with severe vibration.

Exploitation of all partial data provides self-evident advantages when availability is not dependable. As for integrating measurements from different constellations, carrier phase differences (over time and/or SVs) are far easier to mix than the phases themselves. Correlations thus introduced are rigorously addressed, resulting in original practical solutions. It is readily acknowledged that, while velocity is precise, position accuracy is not sub-wavelength (determined instead by pseudoranges). That performance is fully adequate for a host of applications – *e.g.*, 400-knot aircraft with 20-meter wing spans don't require sub-wavelength instantaneous location; dependability is far more important.

In a world with multiple GNSS constellations, or with an insufficient number of healthy satellites, then, methods offered in this book provide needed solutions. They are based on thousands of hours working with real GPS and inertial data, presenting fully disclosed results from tests and also the formulations used in processing all data. The formulations include those originating with this author {*e.g.*, wander azimuth with latitude sensitivity demonstrably and intrinsically negligible, closed form expressions for process noise spectral densities to be prescribed, rigorous determination of correlation effects (with distinct ramifications for each differencing operation), integrity validation extended in several ways, all estimator operations needed for separate processing of each individual 1-sec carrier phase change, and flight-validated block processing for operation without the IMU}.

With in-depth scope that addresses myriad tracking applications as well as navigation, unprecedented robustness and situation awareness are offered, with flexibility and operational versatility.

For those unfamiliar with inertial navigation and/or Kalman filtering, a pre-GPS book by this author presents the "*a*•*b*•*c*'s" with the same perspective and notation: *Integrated Aircraft Navigation* (1976), after five hardback printings, is now available from NavtechGPS in paperback at a much lower price.

To my siblings,
sisters **Corinne**, **Audrey**, and **Dolores**
and in memory of our beloved brother,
USMC Cpl. **Robert J. Farrell**

Over a hundred pages of book material follow, with occasional brief interruptions for expanded descriptions under a heading **Additional Discussion**.

Table of Contents

PREFACE . . . ix
ACKNOWLEDGMENTS . . . x
FOREWORD . . . xii

CHAPTER 1 INTRODUCTION . . . 1
1.1 ESTIMATION APPLIED TO NAVIGATION AND TRACKING . . . 3
1.2 GNSS MEASUREMENTS . . . 4
1.3 INTEGRITY OF INCOMING DATA . . . 5
1.4 DESCRIPTION OF REMAINING CHAPTERS . . . 5

CHAPTER 2 REVIEW OF FUNDAMENTALS . . . 7
2.1 BASIC TRANSLATIONAL MOTION . . . 7
2.2 NOTATIONAL CONVENTIONS AND EXTENSIONS . . . 8
2.2.1 Motion Relative To A Moving Point . . . 9
2.3 ELEMENTARY ESTIMATION: FIXED POSITION . . . 10
2.4 ESTIMATING A MOVING OBJECT'S PATH . . . 11
2.4.1 *Errors In* Position And Speed Along A Track . . . 12
2.5 RECURSIVE ESTIMATION IN MULTIPLE DIMENSIONS . . . 13
2.6 BLOCK ESTIMATION . . . 16
2.6.1 Block Estimation For Position And Speed Along A Track . . . 16
2.6.2 Position, Speed, And Acceleration Along A Line . . . 18
2.7 INERTIAL MEASUREMENTS . . . 19
2.8 GNSS MEASUREMENTS . . . 21
2.8.1 SV Range And Pseudorange . . . 21
2.8.2 Carrier Phase . . . 23
2.9 ESTIMATION AS A DATA FITTING OPERATION . . . 24

CHAPTER 3 INERTIAL INSTRUMENT PROCESSING . . . 27
3.1 BASIS OF PROSPECTS FOR CHANGE . . . 28
3.2 COORDINATE FRAMES . . . 30
3.3 ROTATIONS INVOLVING ANGLES FOR POSITION . . . 30
3.4 TRANSLATIONAL MOTION IN ROTATING FRAMES . . . 31
3.4.1 Position Incrementing . . . 32
3.4.2 Velocity Incrementing . . . 33
3.4.3 Attitude Incrementing . . . 34
3.5 DIRECT PROCESSING OF INERTIAL DATA . . . 36
3.5.1 Attitude Incrementing Task List . . . 38
3.5.2 Velocity Incrementing Task List . . . 39
3.5.3 Position Incrementing Task List . . . 39
3.6 INTERPRETATION . . . 40
3.6.1 Departure From Conventional Methods . . . 41
3.7 SUMMARY . . . 42
3.A SCULLING RECTIFICATION . . . 43

CHAPTER 4 SHORT-TERM ERROR PROPAGATION WITH TRANSLATION AND ROTATION . . . 45
4.1 FUNDAMENTAL INS ERROR PROPAGATION . . . 46
4.2 SEGMENTED APPROACH TO ESTIMATION . . . 48
4.3 OMISSION OR RETENTION OF SELECTED STATES . . . 49
4.3.1 Velocity And Fixed Misorientation (Tilt) . . . 49
4.3.2 Velocity And Misorientation With Fixed Drift Rate . . . 50
4.3.3 Alternative Formulation With Added Incremental Position States . . . 51
4.3.4 Alternate Formulations With Missing States . . . 51
4.4 DECOUPLED POSITION ESTIMATOR . . . 52
4.5 COVARIANCE PROPAGATION : DATA - AVERAGING PERIOD . . . 52
4.5.1 Covariance Adjustment With Rotations . . . 53
4.6 ATTITUDE PSEUDOMEASUREMENTS . . . 54
4.7 SUMMARY . . . 56
4.A UNSTABLE VERTICAL CHANNEL . . . 57
4.B CHARACTERIZATION OF INERTIAL INSTRUMENT ERRORS . . . 57
4.B.1 Background . . . 58
4.B.2 Approach . . . 61
4.B.3 Valuation – Inertial Instrument Degradations . . . 66

CHAPTER 5 CODE AND CARRIER DIFFERENCING . . . 67
5.1 POSITION DETERMINATION WITHOUT DIFFERENCING . . . 68
5.1.1 Mechanization Considerations . . . 69
5.2 ACROSS-SV DIFFERENCING . . . 70
5.3 DIFFERENCING ACROSS RECEIVERS . . . 72
5.4 DIFFERENCING ACROSS TIME . . . 73
5.4.1 Formation Of Residuals . . . 73
5.4.2 Formation Of Sensitivities . . . 74
5.5 INTEGRATION OF DIFFERENCE OBSERVABLES WITH IMU . . . 76
5.5.1 Implications Of Bierman's Factorized Approach . . . 78
5.5.2 Position State Estimator . . . 79
5.5.3 Broader Impact Of Correlations . . . 79
5.6 SEQUENTIAL CORRELATIONS . . . 80
5.6.1 Performance Of Alternate Configurations . . . 81
5.6.2 Block Processing . . . 84
5.6.3 Phase Continuity And State-of-the-Art . . . 85
5.7 SUMMARY OF SYSTEM APPROACH . . . 86
5.A ADJUSTMENT – SEQUENTIAL CHANGE IN PROPAGATION . . . 88
5.A.1 Ionospheric Effect . . . 88
5.A.2 Tropospheric Effect . . . 89
5.B SEQUENTIALLY CORRELATED MEASUREMENTS . . . 90
5.B.1 Classical Approaches . . . 90
5.B.2 Block Processing – WLS Formulation . . . 92
5.C SHORT-TERM INS ERROR DYNAMICS . . . 98
5.D SV MOTION WITHIN THE SUBTRACTION INTERVAL . . . 101
5.E GEODETIC SOLUTION FROM GEOCENTRIC VECTORS . . . 105

CHAPTER 6 INTEGRITY TESTS FOR DATA EDIT . . . 107
6.1 RECEIVER AUTONOMOUS INTEGRITY MONITORING (RAIM) . . . 109
6.1.1 RAIM Detection By Parity Without Differencing . . . 109
6.1.2 Extension To Include SV Offset . . . 111
6.1.3 Exclusion With SV Offset, No Differencing . . . 113

6.2 RAIM WITH DIFFERENCE MEASUREMENTS . . . 115
6.2.1 Detection By Parity With Differencing . . . 115
6.2.2 Extension To Include SV Offset With Difference Measurements . . . 116
6.2.3 Exclusion Of Difference Measurements With SV Offset States . . . 116
6.3 EXTENSION TO ASYNCHRONOUS DATA . . . 121
6.3.1 Approach . . . 121
6.3.2 Mathematical Development . . . 122
6.3.3 Recovery From Invalid Data . . . 125
6.3.4 Parameter Setting And Resetting . . . 126
6.4 WIDER OVERALL PERSPECTIVE FOR INTEGRITY . . . 127
6.A χ^2 DISTRIBUTIONS FROM ESTIMATOR RESIDUALS . . . 128
6.B **QR** FACTORIZATION . . . 129
6.C DERIVATION: DIFFERENCE EXCLUSION WITH OFFSET STATE . . . 131
6.D VALIDATION OF SINGLE-MEASUREMENT RAIM . . . 133

CHAPTER 7 PREPROCESSING AND REPROCESSING . . . 135
7.1 IMU PREPROCESSING . . . 135
7.1.1 Data Acquisition . . . 135
7.1.2 Computational Synchronization . . . 137
7.2 GNSS REPROCESSING . . . 142
7.2.1 Relocation In Time . . . 143
7.2.2 Lever Arm Adjustments . . . 144
7.2.3 Multi-Constellation Observations . . . 144
7.3 DEGREE OF INTEGRATION . . . 146
7.A CONING . . . 147

CHAPTER 8 EXPERIMENTAL RESULTS . . . 151
8.1 PERFORMANCE FROM GNSS DATA ALONE . . . 152
8.1.1 Simulation Under Simplified Conditions . . . 152
8.1.2 Flight Test Results From GPS With No IMU . . . 154
8.2 GPS / INS VAN TEST . . . 162
8.3 GPS / INS FLIGHT TEST . . . 171
8.3.1 Time History Plots . . . 172
8.3.2 Takeoff . . . 178
8.3.3 Data Edit Performance . . . 180
8.3.4 Performance Summary . . . 180
8.4 RELEVANT MECHANIZATION CONCEPTS . . . 182
8.4.1 Additional Future Applications . . . 182
8.A SIMPLIFIED FLIGHT DYNAMICS . . . 183

CHAPTER 9 AIDED TRACKING . . . 185
9.1 MUTUAL SURVEILLANCE . . . 186
9.1.1 Concept And Formulation . . . 187
9.1.2 Operation And Achievable Performance . . . 188
9.2 COOPERATIVE ENGAGEMENT . . . 191
9.2.1 Concept And Formulation . . . 192
9.2.2 Operation And Achievable Performance . . . 193
9.3 FURTHER PROCESSING OF TRACKER OUTPUTS . . . 205
9.A BISTATIC/MULTISTATIC TRACKING . . . 206
9.B LINEARIZED ORBITAL DYNAMICS . . . 207
9.C BALLISTIC OBJECT TRACKING . . . 209

9.C.1 Reentry Vehicles 209
9.C.2 Projectiles 214
9.D LITTORAL ENVIRONMENT 214
9.E SUPPORTING OPERATIONS AND RELATED ISSUES 218
9.E.1 Collision Avoidance 218
9.E.2 Transfer Alignment 220
9.E.3 SAR Motion Compensation 222
9.E.4 Retention Probability With Time-shared Sensing 222
9.E.5 Usage Of Navigation Data In Tracking – Closing Comments 224

CHAPTER 10 SUMMARY AND FUTURE PROSPECTS 225
10.1 INTEGRATED SYSTEM GOALS 226
10.2 INTEGRATED SYSTEM USES 227
10.3 CHANGE WITH GROWTH 229

APPENDIX I NOMENCLATURE – MAJOR NOTATION 233
I.1 SCALARS 234
I.2 VECTORS 235
I.3 MATRICES 237
I.4 SUBSCRIPTS AND SUPERSCRIPTS 238

APPENDIX II FREE-INERTIAL COAST 239
II.1 ANALYSIS 241
II.2 RESULTS 242
II.3 POSSIBLE FUTURE EXTENSIONS 246
II.A COAST PROGRAM 247
II.A.1 Main Program 247
II.A.2 Initialization 248
II.A.3 Scenario Generation 250
II.A.4 Error Propagation 251
II.A.5 Random Error Generation 251
II.A.6 Bookkeeping 251
II.A.7 Plotting 252
II.B PERFORMANCE VALIDATION 253
II.C COAST INITIALIZATION 256

APPENDIX III IMPLEMENTATION ISSUES 259
III.1 INSTITUTIONAL CONSTRAINTS 259
III.1.1 Air Traffic Control (ATC) Applications 260
III.2 DATA FLOW 261
III.2.1 Internal I/O 261
III.2.2 External I/O 263
III.2.3 Timing: Further Details 264
III.3 SOFTWARE CONSIDERATIONS 266
III.3.1 Validation (General) 269
III.3.2 Validation (A Specific Example) 270
III.4 COORDINATION 271

SUBJECT INDEX 273
AUTHOR INDEX 284

Preface

In the mid 1990s a GPS/Inertial project was undertaken at Ohio University, with FAA support, under the direction of Prof. Frank van Graas. This writer was tasked with formulation and software programming to process raw data from gyros (angle increments), accelerometers (velocity increments), and GPS (ambiguous carrier phase and pseudorange with nav messages). Experience gained from the project was combined with other previous activities, to form a book with perspective differing considerably from earlier approaches.

In assembling the material it was deemed unnecessary to cover every facet from the beginning. Widely available other sources, for example, allowed assumption of readers' familiarity with various features of estimation. In addition, expressions are succinctly given herein for strapdown and GNSS operation, without repeating detailed derivations from other sources cited. Material in the chapters will identify and explain areas of common concern elsewhere that can be de-emphasized or even omitted here. In many chapters there is an attempt to avoid detailed expositions, except where either
1) concentrating on details can provide much insight to the uninitiated, or
2) this author's approach differs markedly from commonly used methods.

The scope of this writing was intentionally limited by emphasizing applications requiring high accuracy in dynamics (*e.g.*, velocity to within a few centimeters – *not* meters – per second) ***without*** requiring extreme precision (sub-wavelength at L-band) in position. Most operations can succeed under those conditions, though a few notable exceptions exist. A major reduction in length is permitted by de-emphasizing operations requiring free-inertial coast for extended durations. Again, that affects only a very small percentage of applications (*e.g.*, submarines without access to updating information). Frequent updates continue to be available in most operations, whether from satellites or Earth-stationary transmitters (radio, TV). The latter (terrestrial) sources provide phase – and in some cases also pseudorange data which, though differing in many ways from satellite observations, have comparable functional form. For purposes of this book, emphasis is on usage of satellite observations for updating – other references provide the mechanization specifics for any specialized system or equipment. Thus while discussion herein emphasizes GNSS, much of the methodology could be modified for adaptation to terrestrial sources as well.

Goals just described are motivated by multiple reasons. The most obvious benefit is simplification of both concept and implementation (and therefore a less arduous journey from formulation to operational readiness, checkout, etc.). An equally important consideration, however, is robustness. By using carrier phases as sequential changes only, the entire realm of integer ambiguity resolution is sidestepped. That enhances availability and also removes risk of an occasional misleading integer solution (an issue among many sophisticated operational systems).

Eliminating any possibility of a misleading integer solution is combined with other means to enhance overall robustness. Obviating the requirement for cycle integer determination and for favorable geometry as prerequisite to integrity testing provides unprecedented protection under challenging conditions. These benefits are achieved without sacrificing the primary aim of accurate and dependable estimates for dynamics.

The age-old observation about authors "standing on the shoulders" of countless predecessors clearly applies here. Formal recognition of all would be a daunting task in itself. For bibliographies, an expedient approach is taken instead. Considerable reliance is placed on familiar references and sources cited therein. Occasionally, however, an old or obscure manuscript is highlighted because it precisely captures some key issue in a way not found elsewhere.

Work herein of course reflects the author's personal views, and does not purport to express the views or position of FAA nor any other government institution.

Acknowledgments

The GPS/inertial project team I joined in the 1990s offers living proof that teamwork with all of the right ingredients is quite realistic. The Avionics Engineering Center at Ohio University has a track record of accomplishments earning the highest respect from the very best experts in navigation. Already a recognized Center of Excellence for GPS before the project started, OU was tasked by FAA to develop a capability that can take maximum advantage of low-cost inertial sensors.

Selection of OU was wise, not only due to unparalleled competence in flight test and data collection, but also due to unerring direction consistently provided by the designated principal investigator: Professor Frank van Graas has either written, co-authored, or directed much of the GPS work that established and continues to maintain OU's leadership status, performing the technical and administrative functions with equal precision at the same time. Without his expertise and vision this project could not have succeeded.

A special note of recognition is also due to Dr. Maarten Uijt de Haag, for building and installing the first hardware interface used for this project – while still a graduate student, He has since become a recognized authority in his own right. Another major contributor is Dr. Sai Kiran who, also during graduate student days, performed several "nitty-gritty" tasks toward preparing user-friendly program input data files from decidedly *un*friendly A/D converter outputs. Those efforts plus the work described in his M.S. thesis (cited as a reference in Chapter 7) were vital to our success. Prof. van Graas directed all this work, as well as the algorithm development documented here. Overall project experience, culminating in successful test results (van testing with ground stations pre-2000; stand-alone flight tests after SA removal), can be summarized thus: If avionics were baseball, Ohio University would be World Series champs.

This book contains much material from previously published manuscripts -- more tightly integrated (with correction of an occasional typo) here. I am grateful also to the staff at Institute of Navigation (ION) headquarters for granting permission to reuse much published ION journal and conference material – in fact, for providing advice and in some cases formatting assistance involved in preparation of those original publications.

One more major contribution came from outside the realm of navigation: for computer operations: it was necessary to make these algorithms available in a form useful for a UNIX environment. My son, Mike Farrell, set up a thorough up-to-date LINUX operation on my desktop, networked with his. At age 19 he then wrote *C++* code enabling me to plot all flight test results – in the same LINUX environment used by my *C* program for generating results. Unable to perform all this computing and communications work (LINUX setup, networking, *C++* coding) myself, I'm still flabbergasted by what he was (and is) able to do.

My most profound gratitude is reserved for the Ultimate Source of Unlimited Generosity: Only God in Heaven could have enabled my mind to comprehend this material and focus it into some (hopefully coherent) form.

J.L.F.

Foreword

Inertial measurement sensors are increasingly benefiting from reduced cost through micro-mechanical and impending nanotechnology fabrication techniques. It is therefore reasonable to forecast that future navigation systems will consist of inertial sensors integrated with "something else." Options for "something else" include satellite navigation, Loran, signals-of-opportunity, Electro-Optical, Synthetic Aperture Radar, Radar, Lidar, frequency references, and combinations of all these. Most of these integrations will primarily require short-term integration on the order of several seconds to several minutes to achieve the desired navigation performance.

This book fills a gap in the literature in that it provides a complete framework for short-term inertial integration, written by an internationally recognized expert. Dr. Jim Farrell's unsurpassed insights into classical mechanics and estimation theory form the basis for elegant analytical developments of all necessary equations. He leaves no stone unturned when it comes to optimizing performance, while at the same time, no unnecessary complications are carried along. The resulting framework applies equally to mass-production as well as niche applications.

The reader will find the book to be rich in subtleties of inertial measurement processing with numerous hard to find references on key concepts. All the details are provided to enable a designer to take raw satellite/inertial flight test data to state-of-the-art performance, not only in terms of accuracy and dynamics, but also in the crucial area of integrity.

Another major advantage of this book is its focus on commonalities between navigation and tracking. The joint treatment of navigation and tracking is paramount in a world that is quickly changing from not only wanting to know one's own navigation state, but also wanting to know and track everything else in one's surroundings.

Finally, this book is more than just theory. Dr. Farrell implemented the short-term GNSS/inertial framework and applied it to real flight test data. Up to the time of submitting this book for publication, he never had access to reference data from a post-processed kinematic GPS solution and a navigation grade Inertial Navigation System. Only after completion of the integration were the final results compared with the reference data, confirming all benefits provided in this book with respect to modernized short-term GNSS/inertial integration.

Frank van Graas

CHAPTER 1

INTRODUCTION

This book will enable designers to obtain state-of-the-art dynamics (*1* cm/s velocity, a few tenths mrad leveling corrections) from a low-cost inertial measuring unit (IMU) with frequent updates, primarily from a global navigation satellite system (GNSS) – but not limited to sources in space. Data types and typical rates include the following:

- angle increments ("delta thetas" *e.g.*, at *100* Hz) may be obtained directly from rate integrating gyros, or as products (sampled rate outputs × time increment),
- velocity increments ("delta v's" *e.g.*, at *100* Hz) from accelerometers (as above),
- ambiguous carrier phase and pseudoranges (*e.g.*, at *1*-Hz, from GPS, GALILEO, GLONASS, Beidou, QZSS, etc.) are complemented by a full GNSS navigation ("nav") message (*e.g.*, once per flight or every few hours),
- various other inputs (*e.g.*, magnetic heading, airspeed, angle-of-attack) allowing update by linear estimation. In all cases, raw data inputs provide full flexibility throughout a wide variety of applications. Preprocessed forms for inertial or GNSS data, for input to loosely coupled systems, are considered only summarily here.

Major improvements in extensive tracking operations are also offered, by methods shown following presentation of GNSS/inertial formulations and test results.

One alternate data source, a transmitting station for radio or TV, is important due to availability in many instances where GNSS signals are obscure. Mechanization is de-emphasized here but, with similar functional forms for pseudorange and carrier phase from various sources, much of the methodology herein is broadly applicable. Focus on GNSS can thus be interpreted to include pseudolites, TV stations, etc.

Minimal space is devoted herein to derivations or presentations available from numerous other sources. GNSS signals, for example, are reviewed only briefly; for broad coverage the reader is referred to excellent books such as [1-3]. Similarly, the reader is assumed to be familiar with various features of estimation; Kalman's original work [4] is expanded in several classical texts such as [5-8], while [9] covers the closely related block estimation in a unique way. In addition, succinct expressions – offered in "cookbook" form for strapdown operation – are explained, but without duplicating every detail from sources cited. One source, previously written by this author [10], is cited often. The reason is convenience of expediting developments by succinctly offering – with minimal distraction – an option to scrutinize expanded investigations without changing perspective (notation, applicable conditions, etc.).

Development herein is approached from a direction that differs in many ways from earlier references. One main departure, the short-term condition, has greatly facilitated analytical developments – and thereby enabled simplification of the designer's task. Since a low-cost IMU cannot accurately support long update spans, estimates must be based on data accumulated within limited durations. For purposes here (and for the foreseeable future) "short" can be interpreted to mean periods of several seconds to a few minutes. Immediately that prevents small initial deviations from propagating into very large errors much later; perceptions of system behavior "much later" (*e.g.*, after a half-hour) will be unaffected by data more than a few minutes old. This permits omission of "correction-to-the-adjustment" terms whose effects over a few minutes are below noise levels. Another important consequence is convergence of multiple error sources into one overall resultant with an effective duration spanning the current data – always terminating at the present instant; that span can be called a "data window." The total saving in complexity from exploiting all available opportunities is quite substantial. Later sections of this chapter will discuss the resulting benefits, and subsequent chapters will demonstrate – with supporting analysis and with validation by flight test results – how they enable full performance to be realized with simple polynomials for dynamic propagation models.

Another means of keeping the task manageable is to compromise universal scope of application. With few notable exceptions (*e.g.*, surveying; landing), the vast majority of operations can allow a few meters of position error - *provided that* very low dynamic errors are achieved. Velocity errors, for example, must be held to a few centimeters (*not* meters) per second. This reasoning has roots in the frequent need to determine future, rather than instantaneous present, location. The future time of interest might be the instant of closest approach (for collision avoidance), of projectile impact, or of any event to be governed by an anticipated state. Multiplication of velocity error by the time interval (and/or acceleration error by half its square) generally produces a dominant effect that dwarfs the influence of initial position error. Even in many applications with stated requirements for highly accurate position, that demand often has an unconscious indirect origin: continuous maintenance of accurate position will be accompanied by accurate dynamics. Since the converse is not necessarily true, opportunities to exploit methods shown herein (precise dynamics coexisting with meters of position error) could be broader than initially expected.

The basis for separation as just described is segmented estimation. Conventionally an overall state estimate provides *3*-dimensional information covering position and dynamics (velocity, attitude, inertial instrument offsets), all together in one formulation. Although the pertinent chapter presenting alternative dynamic models will include that approach, the emphasis is on separation of position from dynamics. With this segmentation, two operations are conducted concurrently:

- precise dynamics can be maintained with carrier phase sequential changes – irrespective of pseudoranges
- position is driven by a forcing function provided directly by the dynamic history, and adjustments are obtained from pseudoranges – irrespective of carrier phase.

A quick partial explanation of benefits offered by this approach lies in its complete elimination of integer ambiguity resolution – immediately removing a potential source of catastrophic failure. Nevertheless some designers may hesitate to accept the segmented approach, due to two apparent drawbacks:

- added complexity of the dynamic model needed to accommodate incremental position observables without position states, and
- exploiting carrier phase accuracy for dynamics only, while failing to capitalize on the accompanying accuracy in position.

To offset the first of these, the set of alternative dynamic models will include one with incremental position states. That model can also be used in applications demanding absolute (rather than just incremental) position to be maintained. A similar answer can be applied to the second objection – although the emphasis is on carrier phase with cycle counts unknown, there is nothing to prevent usage of unambiguous phase observables. If those are available *and correct*, there is every reason to use them.

The last item opens the door to a more subtle application of the segmented approach – operations requiring small values not only for velocity error (*e.g.*, *1* cm/s RMS) and leveling adjustments (*e.g.*, tenths of a mrad RMS), but also sub-wavelength position. Even for precise location supported by unambiguous carrier phase, separation from dynamics offers a major advantage over conventional methods, in robustness. A sudden loss of cycle count will of course bring position uncertainty into either approach – but the impact on dynamics will differ. The disruptive transient will disturb the dynamics in conventional operation, but introduce only a short-lived gap in the data stream for the segmented dynamics estimator. Still, either formulation is available – the choice is up to the designer.

Previews can now be given to approaches used in this book for the major facets of short-term aided nav/track, *i.e.*, estimation methods applied to updating of inertial navigation, processing of GNSS carrier phase and pseudorange data, and integrity of the GNSS observations. Following those preliminary descriptions, this chapter ends with discussion of material covered in the remaining chapters.

1.1 ESTIMATION APPLIED TO NAVIGATION AND TRACKING

From the vast field of linear(ized) estimation, only in-scope operations will be addressed. To avoid repetition in writing, assumptions were made concerning readers' knowledge (or access to background derivations) of familiar principles. In addition to coordinate frames, translation in rotating frames, and vector dynamics, material assumed familiar included fundamentals of estimation (state updating, transition matrices, random vector properties). At the same time, it was desired to enable beginners to design a tightly integrated GNSS/inertial system. To reconcile these somewhat conflicting objectives, the next chapter provides an application-oriented focus for specific operations.

Chapter 2 scrutinizes, at length, basic estimation in one, two, and three dimensions. Reasons for dwelling on the obvious as well as the more subtle details can be explained by immediately foreshadowing subsequent material: Inertial nav aiding involves repetitive adjustments applied to estimated quantities (*e.g.*, position, velocity, small-angle departure between perceived and actual orientation with respect to the nav reference frame). Close correspondence between that information and the more rudimentary data sets considered in Chapter 2 (position/velocity; position/velocity/acceleration) will provide valuable insight to those not fully familiar with this material. Likewise, the decoupling of position covariances from separately estimated dynamics (velocity/acceleration in Chapter 2) will prepare readers for the segmented approach emphasized herein – the basis of enhanced robustness, with far-reaching implications explained in later presentations.

It will be seen that the short-term condition, which facilitated so much development herein, supports a far wider scope of application than initially evident. One reason: no such condition restricts the formulations presented for processing inertial instrument outputs (Chapter 3); only error propagation (Chapter 4) is affected – and even that limitation is minor with GNSS aiding. Validation of that last remark is in Chapter 5, but the reason can be foreshadowed by a simple mental exercise: Imagine a perfect initialization – zero error – in position, velocity, verticality, and heading, given to a state-of-the-art inertial navigator in a maneuvering aircraft. How long could it stay in free-inertial mode before velocity error would exceed 1 cm/s ? GNSS accuracy raises performance expectations, and that has guided the presentation of material here, with many departures from conventional (pre-GNSS) concepts.

1.2 GNSS MEASUREMENTS

Pseudorange and carrier phase information can come from various satellite or terrestrial signals, as already indicated. Differencing operations have blunted many – but not all – of the distinctions among these data sources. Paired GPS or GALILEO observations can be subtracted without complication, for example – but operation should also accommodate differencing of measurements using mixed constellations. For that capability – and for updating position in the segmented configuration – the need for any extensions of basic procedures must be determined.

Complexity of mixed operation can be minimized through reexpression. Each user can adopt a world view that adapts most easily to one chosen system (*e.g.*, GPS with a WGS84 reference [11] in CONUS). Observations from another system can be reexpressed as they would have appeared in the chosen world view. A modest amount of computation can account for different system characterizations, such as different departures of each system's time base from Greenwich Mean Time (GMT); adaptations are discussed briefly in Section 7.2.3. Note how the usage of raw data everywhere facilitates interoperability, data mixing, and flexibility for future adaptation to unforeseen conditions.

1.3 INTEGRITY OF INCOMING DATA

Updating is approached from an integrity-oriented view. An oft-untapped integrity resource, the covariance matrix of a once-credible nav solution in steady state, can help guard against incorrect GNSS data – even with multiple flawed SVs. The outcome is an acceptance criterion *without* demands for geometry at every updating event. Individual measurement residuals can be rigorously tested for integrity, separate and independent of all others, for pseudoranges and in concurrent but independent tests for velocity (carrier phase sequential change) data. Every decision conforms to the widely accepted parity test, but without requiring extra satellites. All tests can be performed even if there are too few satellites available to determine an instantaneous solution. Whenever instantaneous geometry permits, those who hesitate to trust covariance usage for integrity decisions (*e.g.*, shortly after initializing an estimator with conservative values) can revisit the measurement sets with snapshot algorithms; nothing is lost. In that latter operation the impact of correlations due to differencing is taken into account. Chapter 6 presents the integrity developments.

1.4 DESCRIPTION OF REMAINING CHAPTERS

Most chapters plunge directly into their topics after minimal discussion, but there are exceptions. After briefly reviewing basic concepts such as coordinate frames {*e.g.*, nav reference directions exemplified by (but not limited to) North/East/vertical}, translation, rotation, basic estimation, and random vector covariance representations, Chapter 2 provides extended scrutiny of fundamentals as already noted.

The next two chapters constitute a strapdown inertial segment. Chapter 3 provides, in the simplest permissible form, step-by-step task lists for incrementing attitude, velocity, and position – with processing equations for raw strapdown sensor data, usable in pseudocode. That is followed by explanation of the marked contrast between inertial system error characterizations relevant here versus other items (coning, sculling) that are far less influential in short term operation. Error propagation over short durations (*i.e.*, less than a tenth of the Schuler period) is then described in Chapter 4. Additions are included for insight, closely relating parts of that development to preceding translational motion analysis, but with accelerations replaced by effects of tilt components. Driving functions used as models for inertial instrument errors are then defined, with rationale for their chosen representations.

The next chapter addresses updating from satellite signals. Heavily citing excellent references readily available elsewhere, Chapter 5 characterizes GNSS data inputs in just enough detail to enable their usage here. The various subtractions among GNSS measurements are shown with emphasis on differencing across satellites (allowing removal of user clock effects although introducing measurement error correlations that necessitate added adjustments) and, for carrier phase, in time (providing the all-important robustness while imposing requirements for added sophistication, with further ramifications introduced by this author).

The next three chapters are aimed at operational considerations. A key issue, treated in Chapter 6, involves the obvious need to reject incorrect inputs. Integrity test covers detection and identification/exclusion of seriously flawed signals. Different techniques and interrelationships among them are discussed, with various ramifications (*e.g.*, multiple simultaneous flawed signals that could arise from a variety of causes). Chapter 7 covers functions (interfacing, sampling/interpolation, lever arm adjustment, synchronization) that are in some ways peripheral but nevertheless highly important for successful implementation. Chapter 8 presents test results realized by using the algorithms presented in this book. As promised, state-of-the-art dynamic performance was obtained, accompanied by position accuracies commensurate with uncorrected pseudoranges.

Rather than end the book after presentation of the GPS/inertial integration test results, a decision was made to magnify the scope through a modest percentage increase in length. With straightforward modifications much of the material can be extended to include tracking (*i.e.*, wherein most or all of the sensors are used to determine the state of remote objects not carrying them). All combinations (air-to-air, air-to-surface, surface-to-air, and surface-to-surface) are addressed in Chapter 9, followed by application-specific topics (including multistatic operation, orbit determination, reentry vehicles, projectiles, littoral environments, and several supporting functions). Addition of that chapter provides answers to a highly pertinent issue – how well a path can be determined from observables alone, without accessing the tracked vehicle's inertial data. The payoff for this extension comes in Chapter 10, which discusses practical means to exploit these capabilities in ways not being used nor planned for usage at the time of this writing. That last remark leads directly into the closing of that last chapter, which envisions a future with full usage of all available resources.

References

[1] Kaplan, E. and Hegarty, C. (Eds.), *Understanding GPS: Principles and Applications* (2 ed), Artech House, 2006.

[2] Misra, P. and Enge, P., *Global Positioning System - Signals, Measurements, and Performance*, Ganga-Jamuna Press [Navtech distributor], 2006.

[3] Parkinson, B.W., Spilker, J.J., Axelrad, P., and Enge, P. (Eds.), *Global Positioning System: Theory and Applications*, American Institute of Aeronautics and Astronautics, 1996.

[4] Kalman, R.E., "A new approach to linear filtering and prediction problems," *ASME Journal of Basic Engr.*, March 1960, *pp*. 35-45.

[5] Jazwinski, A., *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

[6] Bryson, A.E. and Ho, Y., *Applied Optimal Control*, Wiley, 1975.

[7] Brown, R.G. and Hwang, P., *Introduction to Random Signals and Applied Kalman Filtering* (2 ed), Wiley, 1992.

[8] Bierman, G.J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, 1977.

[9] Morrison, N., *Introduction to Sequential Smoothing and Prediction*, McGraw-Hill, 1969.

[10] Farrell, J.L., *Integrated Aircraft Navigation*, Academic Press, 1976. (Now in paperback)

[11] WGS-84, U.S. Defense Mapping Agency, *World Geodetic System 1984*, Washington D.C., 1991.

Additional Discussion

For those interested in tracking, applications analyzed in Chapter 9 are far more extensive than implied by the cursory descriptions given at the end of Chapter 1. The author was fortunate to be "at-the-right-places / at-the-right-times" when a need arose to address each of the topics covered. Tracking was achieved for not only aircraft but missiles – concurrently – through usage of an agile beam radar. For another example, air-to-surface operations subdivide into air-to-ground and vessel tracking from the air. That latter case constrains tracked objects' altitudes to mean sea level – a substantial benefit since it obviates the need for elevation measurements, which are subject to large errors from refraction (bearing plus range measurements, much less severely degraded, suffice). Air-to-ground tracking, by contrast, further subdivides into stationary and moving targets; the former potentially involves imaging possibilities (by real or synthetic aperture) while the latter separates its signature from clutter via doppler.

Reentry vehicles, quite different from other tracking operations, present a unique set of "do's" and "don'ts" owing to high-precision range measurements combined with much larger cross-range errors (because of proportionality to extreme distances involved). Pitfalls from uncertain axial directions of "pancake" shaped one-sigma error ellipsoids must be avoided. Orbit determination is unique in still another way, often permitting a "patched-conic" model for its dynamics. A program based on Lambert's theorem provides initial trajectories from two vectors and the time interval separating them.

Those operations and more are addressed with most observations from radar or other (e.g., infrared imaging) sensors rather than satellite measurements. That of course applies to tracked objects carrying no squitters. Friendlies tracking one another, however, open the door for using GNSS data. Those subjects plus numerous supporting functions are discussed at some length in Chapter 9. Despite very different dynamics applicable to various operations, an underlying commonality (Chapter 2) connects the error propagation traits in their estimation algorithms *and also* – though widely unrecognized – short-term INS error propagation under cruise conditions (Chapters 2 and 5).

Again various reduced forms are easily obtained by simplifying this expression to lower dimension (*e.g.*, planar motion or zero acceleration). By accounting for an arbitrary time history of $\boldsymbol{q}$ these dynamics can cover highly complex motions of the moving position reference; only the tracked object needs a restrictive dynamic representation such as constant acceleration. This enhances capability to track remote objects from observations taken on a platform experiencing extreme dynamics.

Since navigation ("nav") involves imperfect perception as well as reality, the scope must now expand to include observed and estimated quantities, denoted by a circumflex (^) above. A static case is used to introduce the topic.

2.3 ELEMENTARY ESTIMATION: FIXED POSITION

A helicopter hovers at constant unknown altitude. Alternatively, a point is at a fixed unknown location along a line (*e.g.*, the *x*-axis). To determine location a measurement with value $\hat{Y}_1$ is obtained, producing a first *a posteriori* ("after") estimate equal to

$$\hat{X}_1^{(+)} = \hat{Y}_1 \tag{2.14}$$

This value provides an *a priori* ("before") prediction $\hat{X}_2^{(-)} = \hat{X}_1^{(+)}$ for the second measurement, and that estimate is refined by a second observation

$$\hat{X}_2^{(+)} = \hat{X}_2^{(-)} + z_2/2\ , \quad z_2 \triangleq \hat{Y}_2 - \hat{X}_2^{(-)} \tag{2.15}$$

followed by a third observation with $\hat{X}_3^{(-)} = \hat{X}_2^{(+)}$,

$$\hat{X}_3^{(+)} = \hat{X}_3^{(-)} + z_3/3\ , \quad z_3 \triangleq \hat{Y}_3 - \hat{X}_3^{(-)} \tag{2.16}$$

and then a fourth with $\hat{X}_4^{(-)} = \hat{X}_3^{(+)}$,

$$\hat{X}_4^{(+)} = \hat{X}_4^{(-)} + z_4/4\ , \quad z_4 \triangleq \hat{Y}_4 - \hat{X}_4^{(-)} \tag{2.17}$$

which clarifies the general expression for the m^{th} observation with $\hat{X}_m^{(-)} = \hat{X}_{m-1}^{(+)}$,

$$\hat{X}_m^{(+)} = \hat{X}_m^{(-)} + z_m/m\ , \quad z_m \triangleq \hat{Y}_m - \hat{X}_m^{(-)} = \hat{Y}_m - \hat{X}_{m-1}^{(+)} \tag{2.18}$$

which is the same as

$$\hat{X}_m^{(+)} = \frac{m-1}{m}\hat{X}_m^{(-)} + \frac{1}{m}\hat{Y}_m\ , \quad m > 0 \tag{2.19}$$

Substituting $m = 1$ into this equation produces (2.14) – the first *a posteriori* estimate equals the first measurement; substitution of $m = 2$, combined with that condition and $\hat{X}_2^{(-)} = \hat{X}_1^{(+)}$, yields a second *a posteriori* estimate equal to the mean of the first two measurements ($\hat{X}_2^{(+)} = {}^1/_2\hat{Y}_1 + {}^1/_2\hat{Y}_2$). Continuing with higher values yields a general result that, after m measurements, the estimate is an average of all. This establishes equivalence between the *recursive* estimate formulated in the above equations and the *block* estimate that would have resulted from averaging all data together in one step.

Since the average is widely known to be optimum when all observations are equally accurate statistically, the recursion shown here must be optimum for that condition. For measurement errors that are sequentially independent random samples with zero mean and variance R it is well known that mean squared estimation error $P_m^{(+)}$ after averaging m observations is equal to R/m. That is the variance of the *a posteriori* estimate (just *after* inclusion of the last observation); for the *a priori* estimate the variance $P_m^{(-)}$ is $R/(m-1)$. It is instructive to express the last equation above as a blended sum of old and new data, weighted by factors

$$\frac{R}{P_m^{(-)} + R} \equiv \frac{R/P_m^{(-)}}{1 + R/P_m^{(-)}} = \frac{m-1}{m} \tag{2.20}$$

and

$$\frac{P_m^{(-)}}{P_m^{(-)} + R} = \frac{1}{m} \tag{2.21}$$

so that, by combining the last three expressions

$$\hat{X}_m^{(+)} = \frac{R}{P_m^{(-)} + R}\hat{X}_m^{(-)} + \frac{P_m^{(-)}}{P_m^{(-)} + R}\hat{Y}_m \quad , \quad m > 0 \tag{2.22}$$

This form expresses dependency of weights on variances, giving primary influence to information having best accuracy. It is more general than the uniform variance case considered at first; with R in (2.22) replaced by R_m, the weights will be optimum for any measurement variance sequence. The section can end with an extension; the foregoing development is obviously a restrictive case (all R_m equal to a constant R) of the well known general update for direct ("$Y=X$") observations

$$\hat{X}_m^{(+)} = \hat{X}_m^{(-)} + \frac{P_m^{(-)}}{P_m^{(-)} + R_m} z_m \quad , \quad z_m \triangleq \hat{Y}_m - \hat{X}_m^{(-)} \tag{2.23}$$

in preparation for addressing the more challenging task of estimation with dynamics.

2.4 ESTIMATING A MOVING OBJECT'S PATH

An object with motion in conformance to (2.2) is appropriately estimated with the same transition matrix and the same dynamic behavior;

$$\begin{bmatrix} \hat{X}_m \\ \hat{V}_m \end{bmatrix} = \begin{bmatrix} 1 & t_m - t_{m-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{X}_{m-1} \\ \hat{V}_{m-1} \end{bmatrix} \tag{2.24}$$

which, by subtraction from (2.2), applies also to errors $\tilde{X} = X - \hat{X}$ in the estimates:

$$\begin{bmatrix} \tilde{X}_m \\ \tilde{V}_m \end{bmatrix} = \begin{bmatrix} 1 & t_m - t_{m-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{X}_{m-1} \\ \tilde{V}_{m-1} \end{bmatrix} \tag{2.25}$$

2.4.1 *Errors In* Position And Speed Along A Track

Next suppose that, instead of trying to deduce distance X and velocity V the goal is to ascertain smaller *error states* $x_1 = \tilde{X} = X - \hat{X}$ and $x_2 = \tilde{V} = V - \hat{V}$ that evolve when "good-but-not-perfect" speed data can provide continuous accumulation of position increments via dead reckoning ("Σ" in Figure 2.2) with every increment of time $\delta t = t_m - t_{m-1}$. Then velocity can vary during the interval $T = M\delta t$; only the *error* from the dead reckoning sensor is characterized as static, and

$$\begin{bmatrix} x_{m1} \\ x_{m2} \end{bmatrix} = \begin{bmatrix} 1 & t_m - t_{m-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{m-1,1} \\ x_{m-1.2} \end{bmatrix} \tag{2.26}$$

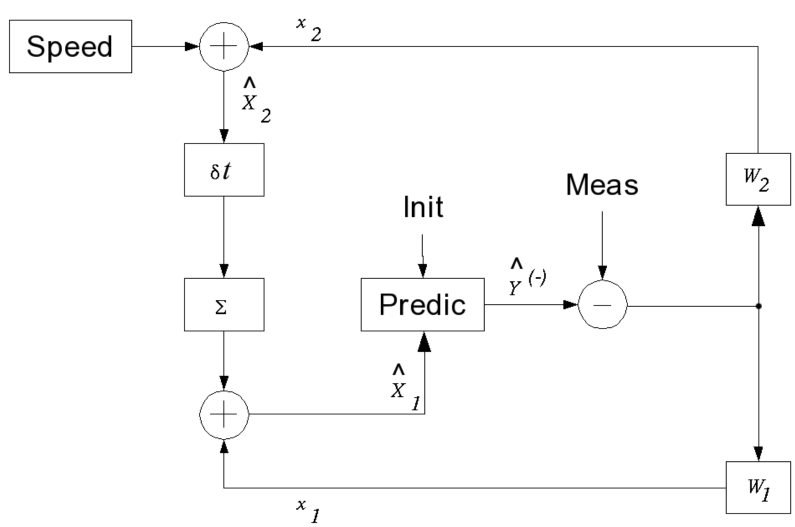

Figure. 2.2 Position and speed on a track

Error states are used in sum and reset operations at summing junctions "+" in Figure 2.2. Perceived error x_2 is added to indicated speed to form *a posteriori* estimated speed $\hat{X}_2$, and perceived error x_1 is reset to zero immediately after it is added to *a priori* estimated position to form *a posteriori* estimated position $\hat{X}_1^{(+)}$. Each *a posteriori* estimate becomes the basis for the *a priori* prediction on the next cycle, by an extension of the static scalar expression shown before (2.18):

$$\hat{\mathbf{X}}_m^{(-)} = \Phi\ \hat{\mathbf{X}}_{m-1}^{(+)} \tag{2.27}$$

with Φ in conformance to (2.24), and every update event at time t_m in Figure 2.2 produces a direct position measurement, in error by an amount ϵ

$$\hat{Y}_m = X_{m1} + \epsilon_m \tag{2.28}$$

at which time the *residual* (difference of measured minus predicted value) is formed

$$z_m \triangleq \hat{Y}_m - \hat{X}_{m1}^{(-)} \tag{2.29}$$

and weighed by a 2×1 vector $\mathbf{W}_m$ to produce the adjustments for the current cycle:

$$\hat{\mathbf{X}}_m^{(+)} = \hat{\mathbf{X}}_m^{(-)} + \mathbf{W}_m z_m \tag{2.30}$$

Be mindful that sequential position observations separated by known time intervals carry strong implications regarding velocity – that is the reason why discrete fixes in position alone can correct velocity as well as location.

All generalizations previously shown, for total distance and for dynamic variations thereof, are immediately applicable to the error states as well. These formulations thus cover all motions including, for actual, estimated, and error states,

- acceleration, and/or
- planar or 3-dimensional motion, and/or
- formation of relative states by subtracting motions of a highly dynamic reference.

The deceptively simple expressions just presented can therefore provide power and versatility to cover a wide range of applications.

2.5 RECURSIVE ESTIMATION IN MULTIPLE DIMENSIONS

Section 2.3 addressed a very restrictive case, *i.e.*, direct measurements of a scalar without dynamics. Section 2.4 introduced the application of estimation with simple (*2*-state) dynamics. As already shown, another way to have multiple states is to consider motion in two or three directions. The goal of this section is to build upon preceding examples, arriving at a general formalism covering the wide variety of cases just mentioned. This will be done by expanding (2.27)–(2.30), in sequence. First consider application of dynamics with the form of (2.27) to a more general situation wherein an accurately known forcing function is present

$$\hat{\mathbf{X}}_m^{(-)} = \Phi\ \hat{\mathbf{X}}_{m-1}^{(+)} + (\text{ forcing vector }) \tag{2.31}$$

An example was shown in (2.13). Just as (2.31) extends (2.27) – which extended the expression given before (2.18) – further generalization is sometimes needed. Rather than any restrictions in form, the *a priori* state estimate $\hat{\mathbf{X}}_m^{(-)}$ can be any vector function Φ (linear or nonlinear, with or without a forcing function) of the *a posteriori* state estimate $\hat{\mathbf{X}}_{m-1}^{(+)}$ from the preceding update cycle:

$$\hat{\mathbf{X}}_m^{(-)} = \Phi(\hat{\mathbf{X}}_{m-1}^{(+)}) \tag{2.32}$$

As (2.43) transforms the covariance matrix in accordance with continuous error state vector dynamics, the following relation expresses the discrete decrement in **P** corresponding to (2.44):

$$\mathbf{P}_m^{(+)} = (\mathbf{I} - \mathbf{W}_m \mathbf{H}_m) \mathbf{P}_m^{(-)} \tag{2.45}$$

Derivations for these "continuous-discrete" matrix expressions and for $\mathbf{W}_m$ are widespread and not repeated here. Less widely recognized, however, are means of choosing spectral densities **E** which, by (2.43), drive rates of increase for covariances – and how their propagations for different states (*e.g.*, acceleration/velocity/position) can be used to control nominal averaging durations within the estimation process. That whole issue can be more fully appreciated after examples are presented with exact values for those durations – now to be provided.

2.6 BLOCK ESTIMATION

Section 2.3 addressed the most elementary case of estimation {a static scalar) and explained that recursion with proper weighting produces the same result as processing all observations together in one block. That latter operation will now be extended, first to the two-state system of Figure 2.2, then for motion at constant acceleration, with some discussion of usage in planar (2-dimensional) and 3-dimensional. cases.

2.6.1 Block Estimation For Position And Speed Along A Track

The task now undertaken is usage of the dynamics shown in (2.2) and (2.24) to obtain estimates for both initial position X_{01} and speed X_{02}, solely from a periodic sequence of position measurements. Each measurement (with true value $Y_m = X_{m1}$) can be combined with the dynamics to produce a linear expression in terms of two unknowns:

$$\hat{Y}_m = \mathbf{J}_m \begin{bmatrix} \hat{X}_0 \\ \hat{V}_0 \end{bmatrix}, \quad \mathbf{J}_m = [\, 1 \quad m\tau \,] \tag{2.46}$$

with τ representing the time between periodic measurements. Immediately after the first observation at time t_1, the task cannot be performed; the measurement "set" provides one equation with two unknowns. At any subsequent observation, there are M equations in two unknowns; as they are compacted, vectors **J** form a matrix **J** :

$$\hat{\mathbf{Y}} = \mathbf{J}_M \begin{bmatrix} \hat{X}_0 \\ \hat{V}_0 \end{bmatrix}, \qquad \mathbf{J}_M = \begin{bmatrix} 1 & \tau \\ 1 & 2\tau \\ \vdots & \vdots \\ 1 & M\tau \end{bmatrix} \tag{2.47}$$

Although the least squares solution with uniform RMS measurement error σ is well known, ramifications are subtle enough to warrant scrutiny. $(\mathbf{J}_M{}^{\mathrm{T}}\mathbf{J}_M)$ reduces to

$$(\mathbf{J}_M{}^{\mathrm{T}}\mathbf{J}_M) = \begin{bmatrix} M & \tau M(M+1)/2 \\ \tau M(M+1)/2 & \tau^2 M(M+1)(2M+1)/6 \end{bmatrix} \tag{2.48}$$

The covariance matrix of error in the solution, under the conditions stated, is also well known to be $\sigma^2(\mathbf{J}_M{}^{\mathrm{T}}\mathbf{J}_M)^{-1}$ which, by multiplication, can be verified equal to

$$\sigma^2(\mathbf{J}_M{}^{\mathrm{T}}\mathbf{J}_M)^{-1} = \frac{2\sigma^2}{M(M-1)}\begin{bmatrix} 2M+1 & -3/\tau \\ -3/\tau & 6/[(M+1)\tau^2] \end{bmatrix} \tag{2.49}$$

Immediately there are several interpretations available that enable conclusions to be drawn from this basic example, and paving the way for more advanced applications. First, the position and velocity errors $\tilde{X}$ and $\tilde{V}$ are replaced by error state notation x_1 and x_2 respectively, components of the state vector $\mathbf{x}$ (in this example, 2×1). Then the last equation provides the initial covariance matrix $<\mathbf{x}_0\,\mathbf{x}_0{}^{\mathrm{T}}>$, related to the final covariance matrix $<\mathbf{x}_M\,\mathbf{x}_M{}^{\mathrm{T}}>$ by means of the transition matrix from initial (t_0) to final (t_M) time. This shorthand notation for error states and for transition matrices can now replace the expanded equation forms wherever convenient, *e.g.*,

$$<\mathbf{x}_M\,\mathbf{x}_M{}^{\mathrm{T}}> = \Phi(t_M, t_0) <\mathbf{x}_0\,\mathbf{x}_0{}^{\mathrm{T}}> \Phi^{\mathrm{T}}(t_M, t_0) \tag{2.50}$$

Combination with (2.49) after some manipulation yields

$$<\mathbf{x}_M\,\mathbf{x}_M{}^{\mathrm{T}}> = \frac{2\sigma^2}{M(M+1)}\begin{bmatrix} 2M-1 & 3/\tau \\ 3/\tau & 6/[(M-1)\tau^2] \end{bmatrix} \tag{2.51}$$

Appropriately, the denominator (M-1) affects only velocity variance for final error – but an attempt to deduce *initial* states from only one observation would produce singular covariances throughout. While that detail and the sign change in correlation are of some interest, the main results of this development are variances for large M. RMS errors approach $2\sigma/\sqrt{M}$ and $(2\sigma/\tau)\sqrt{3/M^3}$ in position and velocity respectively. With $T = M\tau$ the latter amount can be rewritten as $(2\sigma/T)\sqrt{3/M}$. A scanning radar with *5* seconds between successive dwells will produce RMS range rate error of *0.1* m/sec after a minute if $\sigma = 6$ meters in range. That figure assumes a straight line path for the object being illuminated by the radar. A long distance will now be considered; at long ranges the angle (cross-range) channel performance is less sensitive to the path. For a radar RMS angle measurement error of *6* mrad, RMS cross-range velocity error will be that long distance multiplied by *0.0001* reciprocal seconds. If the interval between successive dwells can vary by a factor of two (because the illuminated object is not at the center of a reciprocating scan), then the performance can be bounded between two values having a ratio of about $\sqrt{2}$: *1*.

The situation just typified gives rise to measured components of a specific force vector $\mathbf{A}_A$ or of an angular rate vector $\boldsymbol{\omega}_A$ – absolute, *i.e.*, *with respect to* an absolute reference but *resolved along* axes that change direction, often irregularly. These measured components must be converted into familiar navigation expressions (attitude, velocity, position) relative to and / or resolved along axes of a nav reference which itself changes (the vertical direction is changing while moving over a curved earth, which itself is rotating). All requisite operations involved in that conversion (integration, coordinate transformation, formation of vector sums, differences, products, etc.) must permit state-of-the-art accuracies to be maintained continuously. After performing these operations for generations, designers fortunately have access to standard means of satisfying all necessary demands. The IMU in combination with all provisions needed to perform those operations constitutes an inertial navigation system (INS). Two main strapdown functions can be summarized in terms of

- converting $\boldsymbol{\omega}_A$ into attitude – a direction cosine transformation matrix T between vehicle and nav-reference coordinates, and
- transforming $\mathbf{A}_A$ (as measured in vehicle axes) into nav-reference coordinates, adding adjustments plus a gravity vector $\mathbf{g}$, producing the velocity rate:

$$\dot{\mathbf{V}} = \mathsf{T}\mathbf{A}_A + \mathbf{g} + (\textit{adjustment terms}) \tag{2.56}$$

The adjustments enter when combining acceleration with a rotating coordinate frame. Chapter 3 offers a full set of algorithms, successfully validated in flight.

Another kind of challenge arises in preparing a dynamic error model needed for repetitive corrections from GNSS and / or other sources of information external to the INS. It was just carefully noted that the INS *permits* satisfaction of performance requirements. The INS by itself cannot *guarantee* that satisfaction; conversion of derivative data (ω, $\mathbf{A}$) into time integrals (attitude, velocity, position) carries intrinsic properties of error growth. Because ω as measured contains a nonzero drift rate $\mathbf{n}_\omega$ T as computed contains imperfections – expressed in terms of a small-angle misorientation vector $\boldsymbol{\psi}$ – having a tendency to grow with time. This attitude error interacts with specific force $\mathbf{A}$ and also sums with an error (denoted $\mathbf{n}_a$) in $\mathbf{A}_A$ as measured, producing a departure from (2.56) accurately modeled in the form

$$\textit{velocity vector } \underline{\textit{error}} \textit{ rate} = \boldsymbol{\psi} \times \mathbf{A} + \mathsf{T}\,\mathbf{n}_a \tag{2.57}$$

Chapter 4 characterizes all terms in (2.57) and discusses at length all simplifications, showing that the total modeling error is less than the noise levels included.

For a simplified preview of where this model leads, consider the case of $\mathbf{A}$ equal to *1-g* upward {counterbalancing downward $\mathbf{g}$ in (2.56), applicable to cruise flight as well as many land vehicle and maritime scenarios} with constant $\mathbf{n}_\omega$. For a short-term model let components of a scaled vector $g\mathbf{n}_\omega$ drive components of another scaled vector $g\boldsymbol{\psi}$ which drives velocity vector error which drives position error. This system has dynamics shown in (2.4–2.7) and can be analyzed as in section 2.6.2 !

2.8 GNSS MEASUREMENTS

Methods described herein exploit features common to existing and planned future satellite navigation systems. It matters little whether the pseudorange and carrier phase data used for INS updating come from GPS, GALILEO, GLONASS, Beidou, QZSS, or any other space vehicle (SV). Due to familiarity gained from a wealth of data provided thus far, a GPS time base and GPS parameters are used here in illustrative examples. Standard navigation messages are covered in references previously cited, with methods to compute

- SV clock adjustments at any time, from pertinent data in the nav message, and
- SV position and velocity, at any measurement time, from SV ephemeris data.

Ability to backtrace SV position from reception to SV transmit time is illustrated here.

2.8.1 SV Range And Pseudorange

Satellites in accurately known orbits repetitively transmit signals with specified carrier frequencies modulated by waveforms that are similar (to enable reception of many) but distinguishable (to enable extraction of separate SV information). Detections are made by recognizing modulation codes, each with a unique bit sequence (*e.g.*, a length of *1023* within a *1*-ms period has about a million time chips per second; for patterns time-aligned to within *0.01* chip, *3*-m accuracy is achievable). Ideally that signal time alignment – placed in the correct *1*-ms {*300*-km} interval – corresponds to a product, speed of light $c\times$ (interval between signal transmit epoch and matching receive epoch) – the matched epoch can be the beginning or the end of the pattern, or any chosen point between. The observation, however, contains uncorrected multipath, quantization, and noise *plus* the following effects (also in distance units; scaled by c):

- propagation timing offset from ionosphere (*Iono*) and troposphere (*Tropo*),
- delay B_u in the user's receiver (signal passage through waveguide, etc.),
- clock offsets C_u and C_j in the user's receiver and the j^{th} SV, respectively

which combine to produce the relation between *pseudorange* Y_ρ and magnitude of a vector separation between instantaneous position $\mathbf{R}_m$ of the receiver at the time of the observation and the j^{th} SV position $\mathbf{S}_j$ at the time when that signal was transmitted:

$$Y_\rho = |-\mathbf{S}_j + \mathbf{R}_m| + Iono + Tropo + B_u + C_u - C_j \tag{2.58}$$

The measurement would have the same expression with a circumflex over every term so that, as always, the *residual* (measurement - anticipation) at that time would be

$$z = \hat{Y}_\rho - \{ Y_\rho \,|\, predicted\ from\ estimates \} \tag{2.59}$$

containing a measurement error ϵ with an expanded definition as follows:

In practice the residual is formed by computing each term and subtracting the result from the observation – which contains uncorrected multipath, quantization, and noise. For residual formation these are conveniently lumped into ϵ with the effects of imperfect corrections in propagation, B_u, SV clock, and inexact knowledge of $\mathbf{S}_j$.

It is now convenient to write the distance $|-\mathbf{S}_j + \mathbf{R}_m|$ as an inner product with a unit vector $\mathbf{1}_{mj}$ from the j^{th} SV (at transmit time) to the receiver (at receive time t_m):

$$|-\mathbf{S}_j + \mathbf{R}_m| = \mathbf{1}_{mj}^{\mathrm{T}}(-\mathbf{S}_j + \mathbf{R}_m) \tag{2.60}$$

and, with the same procedure used for both true and estimated quantities – recalling that inexact SV location effects are included in ϵ – the residual becomes simply

$$z = \mathbf{1}_{mj}^{\mathrm{T}}\mathbf{r}_m + \tilde{C}_u + \epsilon \;,\; \mathbf{r}_m = \mathbf{R}_m - \hat{\mathbf{R}}_m \;,\; \tilde{C}_u = C_u - \hat{C}_u \tag{2.61}$$

A few items remain before the subject of pseudoranges can be closed. First, many operations need a precise location where measurements are applicable – rather than "at the receiver" a specific location is often the receiving antenna phase center. Secondly, the time of reception is readily available, but to form $\mathbf{S}_j$ the SV velocity must be taken into account. An accurate satellite location at the time when that received signal was transmitted from the j^{th} SV can be determined by computing

- SV location at the time of reception
- transit time, (distance from that location to $\mathbf{R}_m$) / c
- product (transit time) × (SV velocity vector)

which is subtracted from SV location at receive time. Error in this product (*e.g.*, from imprecise transit time) is minor (smaller than other effects already covered by ϵ). For 12-hour orbits higher-order SV dynamics are not needed for this adjustment because corresponding orbital radii {roughly $4\times$ (earth radius), with $1/16\ g$ inverse-square law gravitation} and transit times (up to about 0.1 s) produce acceleration effects (from $0.5 \bullet 1/16\ g \bullet 0.1^2$) of about 3 mm – not all of which is along the direction $\mathbf{1}_{mj}$.

One final issue in regard to this topic is imperfect knowledge of $\mathbf{1}_{mj}$. Inexact $\mathbf{S}_j$ has been taken into account, but inexact receiver antenna position also affects computation of $\mathbf{1}_{mj}$. Although quite manageable, the minor nonlinearity has provoked much concern. All doubts should be erased by noting widespread success of linear estimation – producing state-of-the-art results – with GPS. Even more convincing, many authors (including this one) have purposely tested extreme initial location errors – using earth-centered earth-fixed (ECEF) axes in (2.60, 2.61) until convergence to within a few meters after two or three iterations. In fact, some receivers have been mechanized with that iterative procedure starting at outlandish positions.

Most GNSS applications start with reasonably limited initial error but, if operation must be robust under extreme conditions, the last point is worth mention. Subsatellite points (on earth surface directly below each SV) can be computed with vectors $\mathbf{S}_j$ from all GNSS signal detections received. A crude initial $\mathbf{R}_m$ estimate could be on the earth's surface, at latitude midway between highest and lowest – and longitude also midway between extremities – of subsatellite points. That easily locates the proper code interval {*300* km for the previous *1*-ms example} and, once inside a linear region, estimates for $\mathbf{R}_m$ quickly converge. Usage of local reference axes in (2.60, 2.61) before convergence would of course produce poor results, but that would be a design flaw – not any intrinsic nonlinearity problem.

2.8.2 Carrier Phase

Description of GNSS observables will now be broadened, starting with similarities and differences between pseudorange and carrier phase measurements. First, both are ambiguous – but on vastly different scales, as follows: With no measurement errors and exact compensation of all timing offsets in (2.58), only $|-\mathbf{S}_j + \mathbf{R}_m|$ would remain. Often that is used to describe the ideal pseudorange measurement but, more generally and more precisely, it can relate to either carrier (at frequency f) or pseudorange {*e.g.*, from GPS coarse acquisition ("C/A") code } observations in the form

$$Y_m = integer \bullet (interval \} + (distance) \text{ modulo } (interval) ;$$
$$interval = [\, c \bullet 0.001 \text{ s } \{\text{GPS C/A pseudorange}\} \;\; | \;\; \lambda \, \{\text{carrier}\} \,] \qquad (2.62)$$

where λ is transmitter wavelength (equal to a ratio c/f) and the distance is the product $c \times$ (time from SV transmit signal epoch to the corresponding received signal epoch) at the instant (denoted t_m) chosen for measurement. Epochs are noninteger – a carrier epoch can be at any phase angle from 0 to 2π; pseudorange epochs exemplified here can correspond to chip counts (≤ 1023). For the latter the interval is much larger than almost any position uncertainty – even if not, centroiding of subsatellite points (discussed in Section 5.1.1) can easily "unfold" each modulo remainder into unambiguous pseudoranges by supplying the correct integer.

Superior accuracy offered by carrier phase is accompanied by greater challenges. The correct integer for (2.62) is obviously a truncated value of the ratio $(distance)/(interval)$; applying the correct multiple of *300* km to an ambiguous code measurement is obviously easier than determining a much larger number (more than a million times greater) of wavelengths applicable to ambiguous carrier phase. Still, successful resolution of carrier ambiguities has been widely reported with a wealth of documented methods and results. Despite the ingenuity and availability of those methods, carrier ambiguity resolution is not pursued here. The purpose is not to avoid added computational requirements, but to eliminate a risk intrinsic to every means of finding the number of wavelengths: *any* departure from a correct set of multiples – complete, consistent, and continuous – places the operation in jeopardy until a correct reacquisition can be made. There are ways to reduce the risk and/or the damaging effects of incorrect integers (or near-integers in some resolution methods), *e.g.*,

- addition of a base reference receiver (Section 5.3) effectively reduces the span to be examined in each direction to values less than (*baseline distance* $\div \lambda$) ;
- conventional systems can be formulated with incremental position rather than the more demanding absolute location,

but an inescapable vulnerability, even to very infrequent disruption of operation, is a high price for sub-wavelength accuracy in most applications. That is especially true when sequential differencing offers high accuracy in dynamics without the risk. Again the overriding priority of accurate dynamics with robustness is the reason for emphasizing sequential changes in carrier phase – wherein unknown integers cancel.

While implications of the decision as just stated are fairly obvious, some additional benefits are more subtle. To illuminate those, the term *integrated doppler* is applied literally to carrier phase: Aside from modulation and amplitude, a transmit waveform is proportional to $u_T(t) = \cos 2\pi f t$. The waveform received at the later time $t + R/c$ is proportional to $u_R(t + R/c) = \cos 2\pi f t$. When distance R varies, the doppler shift resulting from reexpression of $u_R(t)$ is instantaneous,

$$u_R(t) = \cos\{2\pi f[t - (R_0 + \dot{R}t)/c]\} = \cos\{2\pi[(f - \dot{R}/\lambda)t - R_0/\lambda]\}, \; \ddot{R}=0 \quad (2.63)$$

i.e., the fixed frequency shift applies when R varies linearly with time – otherwise an instantaneous range rate gives rise to the concept of an instantaneous frequency. The concept is paradoxical in that the only way to measure it is by observing phase change over a finite interval – and using diminutive intervals for that purpose produces accuracy that is unspectacular compared with actual capability. Those limitations are avoided, with or without carrier ambiguity resolution, by working with phase – the time integral of frequency. The constant of integration, of course, is unknown; carrier observations have the form of (2.62), but generally with arbitrary integer values. Using sequential differences relieves *both* the burden of ambiguity resolution *and also* the requirement to maintain phase continuity throughout. Chapter 5 discusses implementation in greater detail.

2.9 ESTIMATION AS A DATA FITTING OPERATION

A broad theoretical basis portrays methods and concepts used here in terms of optimal performance; estimates obtained by these methods are widely depicted as having minimum variance, least squared error, conformance to conditional means, etc. While these traits are obviously desirable, they are rigorously true only under restrictive conditions (perfect linearity, exact conformance to models used, exactly known parameters, etc.) that are often unreachable in practice. Designers cannot ignore the implications, especially since the ultimate in performance can often compromise robustness. Wherever necessary from this point forward, some optimality is unhesitatingly sacrificed for the essential virtue of stability. That process begins by "overbounding" – assigning pessimistic values to parameters used – thereby settling for performance that *would* be best *if* those parameters had the larger-than-expected values assigned.

A first candidate for conservative parameter setting is measurement error variance. Adopting a larger-than-actual value causes degradations obtained in operation to be more plausible members of the ensemble being represented. A system designed to withstand larger errors will more readily survive with smaller ones. Some applications require pull-in from extreme initial errors – examples include sudden acquisition of radar track files or satellite orbit determination, with very large velocities that are initially unknown.

Propagation of initial conditions "by the book" can violate the theoretical underpinnings in some applications, unless remedial measures are taken. One simple but quite successful device is a multiplier for σ^2, initially large but subsequently stepped down (*e.g.*, by a factor of two on each step), until a steady value (*e.g.*, *2*) is reached. Detailed decision points will of course vary with the application. GNSS initialization is often less extreme and therefore less critical, but a modest multiplier could still suppress transients from satellites just coming into view over the horizon.

For many designers the process noise settings have been more challenging. Again, though, an opportunity for systematizing exists. For a scalar measurement the matrices $\mathbf{W}$ and $\mathbf{H}$ in (2.39) become the column vector $\mathbf{W}$ and the row vector $\mathbf{H}$ respectively, while the matrix inversion collapses into division by a scalar, with $\boldsymbol{R}$ replaced by a scalar σ^2;

$$\mathbf{W}_m = \mathbf{P}_m \mathbf{H}_m^{\mathrm{T}} \ / \ (\mathbf{H}_m \mathbf{P}_m \mathbf{H}_m^{\mathrm{T}} + \sigma^2) \tag{2.64}$$

After pull-in transients subside, (2.43) is effectively controlled by propagating $\mathbf{E}$ through the transition matrix. With the integral in that expression then replacing $\mathbf{P}_{\mathrm{m}}$ in $\mathbf{H}_m \mathbf{P}_m \mathbf{H}_m^{\mathrm{T}}$, a duration T can be determined for which the two terms in the above denominator are equal:

$$\mathbf{H} \int_{t-T}^{t} \Phi(t, \ell)\, \mathbf{E}(\ell)\, \Phi^{\mathrm{T}}(t, \ell)\, d\ell \ \ \mathbf{H}^{\mathrm{T}} = \sigma^2 \tag{2.65}$$

as in (5-57) of [4]. For data taken T seconds ago, the integral accumulation provides a balance – for weights given to data at that age and a new measurement. For data taken much earlier than T seconds ago, much accumulation from $\mathbf{E}$ has occurred since those observations; the denominator carries that increased accumulation, and lower weight is given to older data. For data much more recent than T seconds ago the opposite is true; newer data will receive greater weight. Usage of this expedient has consistently facilitated tuning of this author's estimation algorithms, both for tracking remote objects and for nav (*e.g.*, ownship INS updating with GNSS).

A nominal "data window" – or effective data-averaging interval – with duration T offers ways to avoid several problems. Immediately the burden of rigorous adherence to prescribed models can be relaxed; consider for example the need to track a remote aircraft that maneuvers at will through an irregular serpentine 3-dimensional path: Rather than imposing a rigid pattern that the dynamics "must" follow, the designer imposes whatever limitations physically constrain the motion (*e.g.*, the tracked aircraft's controls disallow wing loading changes at rates that can cause structural damage). Conventional flight paths thus have acceleration histories that are quasistatic for at least a few seconds. In analogy with splines, process noise levels conforming to (2.65) roughly approximate a convoluted space curve by a sequence of overlapping arcs, each resembling a low-order polynomial valid for T seconds and terminating at the most recent observation time.

With comparable values for T and the preceding section's $T = M\tau$, performance will not be identical – but with well-chosen decisions, conservative characterizations can be realized. In subsequent chapters this reasoning will be reapplied to other operations (*e.g.*, inertial navigation – with updates of position, velocity, and tilt – and then velocity, tilt, and drift – instead of position, velocity, and acceleration). That will bring a sense of "mission accomplished" to those facing demands for dependable requirements in the presence of inescapable real-world inexactness.

A similar relaxation of strict conformance to theory is permitted for noise spectra. Observation errors are nearly sequentially independent for autocorrelation times substantially lower than the intermeasurement interval (pessimistic RMS values can offset small reductions in information content from nonzero correlation). Likewise, process noise does not have to be white – *e.g.*, see *pp*. 185-186 of [4].

To reiterate, appreciable effort has been devoted at the outset here toward enabling firm commitment to specified performance levels. The world offers good – not perfect – models. Successful identification of a configuration deemed acceptable *and superior to another that was pessimistically characterized* gives credibility to that commitment. Conservatism can solve multiple modeling problems; impact of nonlinearities and of parameter inaccuracies can be effectively included in the random errors with overbounding statistics. Weights that are not precisely optimal – thus rendering (2.45) not quite valid – will not trigger computational instability for that same reason (they *would* be optimal in a system with the pessimistic model).

Some insight into that latter remark can be grasped from (2.40). With inexact weight $\mathbf{W}_m + \Delta\mathbf{W}_m$ the adjustment contains an unwanted product $\Delta\mathbf{W}_m\,\mathbf{z}_m$ – of second order (again, the overbounding statistics are chosen to dominate products of small quantities). Note how that leniency does *not* extend to the spare-no-effort formation of a residual as a small difference of large numbers in (2.35).

The foregoing material provides a reliable basis for later chapters to enable successful practical designs in a wide scope of operations. Outputs can have myriad variations, applicable to navigation or to track (cooperative or non cooperative; monostatic, bistatic, or multistatic), for absolute or relative states (*e.g.*, coordinates, velocity components), expressed as totals or as departures from specified nominal values, planar or 3-dimensional, resolved along fixed or rotating coordinate frames, with or without direct sensing of derivatives (doppler or inertial information).

References

[1] Brown, R.G. and Hwang, P., *Introduction to Random Signals and Applied Kalman Filtering* (2 ed), Wiley, 1992.
[2] Kalman, R.E., "A new approach to linear filtering and prediction problems," *ASME Journal of Basic Engr.*, March 1960, *pp*. 35-45.
[3] Morrison, N., *Introduction to Sequential Smoothing and Prediction*, McGraw-Hill, 1969.
[4] Farrell, J.L., *Integrated Aircraft Navigation*, Academic Press, 1976. (Now in paperback)

CHAPTER 3

INERTIAL INSTRUMENT PROCESSING

For motions described in the preceding chapter, all of the emphasis was on translation. Navigation or tracking developments cannot proceed much further before considering vehicles that rotate while changing position over a curved earth. That topic includes definitions of accepted conventions for relevant coordinate axis directions and relations between them. After examination of essential rotation characteristics, preparation is complete for step-by-step portrayal of inertial instrument (gyro and accelerometer) output processing. Again the pertinent issues are stated succinctly – familiarity with, or availability of, elemental reference material is assumed.

That last statement does not foreshadow complexity in what follows. In fact, the opposite is true; algorithms for processing raw strapdown inertial instrument outputs – long regarded as an exclusive province of experts – can be dramatically simplified for most applications. Reasons are largely traceable to technological advances that followed many current practices; a brief history can explain:

Strapdown inertial navigation has a close historical association with digitization. Highly dependent on computation, it became feasible only when computing speeds permitted; inevitably then, those speeds were marginally adequate in early configurations. That condition spawned a critical need for inventive minds of pioneers – and invent they did. John Bortz [1] almost singlehandedly turned an obscure study by Laning [2] into a firm foundation, blending developments in both hardware and algorithms that continue to guide inertial system designs to this day. Paul Savage [3] developed and propagated myriad techniques, providing beneficial contributions for four decades. John Mark [4] and others, too numerous to mention, perfected their algorithmic approaches until little or no further improvement in computational efficiency would be necessary (or hardly even possible). For proprietary reasons, many algorithms were not divulged until years after development. Ironically, some became widely available only after faster processing capabilities made them less critically urgent. Also, after decades of algorithmic efforts spawned by Miller [5], it was shown in [4] how the coefficients need to account for gyro response (in fact, it showed how the intended corrections, formed without regard to gyro frequency response, can produce amplification – not reduction – of error). Fortunately, as the ensuing material illustrates, all this complexity can be sidestepped in a direct formation of position/velocity/attitude time histories from raw data inputs.

3.1 BASIS OF PROSPECTS FOR CHANGE

Work performed to date, by those cited above and various others, has brought strapdown to an impressive level of capability. There is, however, urgent need for change. Immediate attention will be directed first to why, and then how.

Throughout the events just described, prospects were repeatedly raised for low-cost inertial systems. Whatever evaluation may apply to progress thus far, the field now stands at the brink of major change. Application of micro electromechanical sensor (MEMS) technology [6] to strapdown carries potential for unprecedented success on a massive scale. To obtain it, however, additional conditions must be met. First is interface modernization: Present antiquated standards call for inertial nav system (INS) outputs that follow after processing, producing velocity in hybrid units and attitude in terms of abstract "gimbal" (Euler) angles with precision and timing characteristics that are pitifully incapable of supporting modern system requirements [7] (and have thus necessitated costly workarounds in countless extensions for modified operation). To halt this massive loss to the industry, flexibility for widespread application will require inertial measuring unit (IMU) outputs: velocity increments from accelerometers and rotational increments from gyros. It cannot be emphasized enough – full realization of low-cost inertial systems will require raw data availability, plus an accepted standard for output formats. [8]

Furthermore, to make low-cost inertial navigation a reality, MEMS technology must be accompanied by another event: absolute clarity of computational algorithms. Users must understand exactly how to use the raw increments of rotational and translational motion. Initially this objective might seem unrealistically ambitious, in view of the admittedly brilliant (and in fact ingenious) sophistication implicit in computational schemes devised by early strapdown pioneers. The goal of results without complexity (and without limitations of the past) is immediately achievable, however, through exploitation of four modern opportunities:

- emergence of low-cost instruments, enabling selection of IMU components for flexible design – rather than acceptance of an inflexible system
- an accompanying increase in inertial instrument sampling rates (from tens to hundreds of Hz), enhancing accuracy of attitude and velocity increments formed by basic procedures with little or no coning and sculling compensation
- emergence of GNSS and other aiding sensors providing inputs at *0.1* Hz or higher data rates, sharply reducing reliance on free-inertial error propagation
- the vastly improved processing capability of today's computers.

Impact of processing capability is readily apparent from a highly informative two-part description of accepted processing methods. [9] Algorithms therein, undeniably brilliant, originated when achievable computing speeds were marginal, and retain (by that author's admission) heavy influence of that past circumstance. Operational systems have evolved using those algorithms, and users understandably would not retrofit them just because today's computers are faster.

For future systems, however, it does no discredit to illustrate how methods in [9] can now be replaced with simpler approaches facilitated by modern capability. Since computational efficiency is no longer a premium, processing can be more direct (and therefore more understandable). This chapter addresses that endeavor in depth, with support from discussion of successful validations using this direct approach.

Another key condition offered in most applications contributes to the effectiveness of straightforward processing for IMU outputs: availability of update information, at intervals on the order of *1*-min or less apart, from aiding devices with accuracy allowing satisfaction of overall system requirements. Since updates repeatedly interrupt the error growth sequence that would otherwise flow during free-inertial error propagation, the IMU needs to maintain velocity and attitude accuracy for just a short period (*e.g*., less than a minute). As one quantitative real-world example, [10-12] show a nominal *8*-second tuning achieving RMS accuracies better than a meter in position, *1*-cm/s in velocity, and tilt states at a few tenths mrad, using GPS inputs at *1* Hz. In this scheme the IMU does not have to maintain accuracy, for velocity nor attitude, over a long period (nowhere near any significant fraction of the *84*-min Schuler period). This consigns to total irrelevance many considerations typically associated with inertial nav performance. Expression of position error growth in nmi/h has no significance, since it represents only an average that would apply over an uninterrupted Schuler cycle. With updates only seconds apart (sometimes less), a cycle can hardly get started, let alone finished. Drift rates can be completely outside the range of the industry's "nav-quality" ratings; the example just described, with *8*-s nominal averaging duration for GPS updates, exceeded typical *0.01*-deg/h nav-quality drift rates by over five (actually closer to six) *orders of magnitude*. Moreover, turns and speed changes activate motion-sensitive inertial instrument errors tending to disrupt any semblance of a Schuler cycle, even in the absence of updates [13]. This can call for short data-averaging intervals of aiding information, even when average drift rates are lower. In all cases, short-term expressions are straightforward and simple {*e.g.*, Eq. (3-53) of [14] expresses error propagation as a truncated power series in time; a cubic (and in many cases a quadratic) series is adequate}.

Provided that sampling rates are sufficient to avoid aliasing, importance of coning and sculling will be minimized in many operations with short-term conditions applicable. Preprocessing in accordance with [4] can be prescribed if needed (*i.e.*, if short-term conditions are *not* invoked) and, with or without it, procedures described herein can be applied unconditionally.

If extended coast durations are not required (*i.e.*, repetitive updates are available), error propagation is also vastly simplified. With maximum coast duration quantified in relation to gyro and accelerometer quality, this chapter describes straightforward computational procedures capitalizing on modern processing capabilities. Validation will again be provided by results with measured data in later chapters.

3.2 COORDINATE FRAMES

This chapter uses orthogonal triads defined with directions as follows – unit vectors along $+x$, $+y$, $+z$ coordinate axes denoted **I**,**J**,**K** respectively with subscripts $_{A}$, $_{E}$, $_{L}$:
Vehicle: $\mathbf{I}_A$, $\mathbf{J}_A$, $\mathbf{K}_A$, outward along the roll, pitch, and yaw axes respectively
Earth: $\mathbf{K}_E$ upward along the geodetic pole; $\mathbf{I}_E$ outward to Greenwich meridian
Platform: $\mathbf{K}_L$ downward along the local vertical; $\mathbf{I}_L$ - $\mathbf{J}_L$ plane locally level.

Where only two axes are defined the third is derived by orthogonality, *e.g.*, $\mathbf{J} = \mathbf{K} \times \mathbf{I}$. To clarify the last triad, $\mathbf{I}_L$ is initialized along North so that $\mathbf{J}_L$ initially points East – and those axes would remain in cardinal directions if the vehicle stayed at fixed latitude and longitude. With unrestricted translation over a curved earth, the horizontal reference $\mathbf{I}_L$, $\mathbf{J}_L$ slowly changes. That variation, according to a simple scheme shown later in this chapter, maintains computational stability even if the vehicle crosses over the North or South pole (where longitude and heading are both undefined as expressions for their theoretical rates of change are singular).

Subsequent developments use another triad $\mathbf{I}_P$, $\mathbf{J}_P$, $\mathbf{K}_P$ of considerable importance in addressing *misorientation* – small-angle deviations of the *perceived* nav reference triad off $\mathbf{I}_L$, $\mathbf{J}_L$, $\mathbf{K}_L$. Only perceived quantities can be available in algorithms used to process onboard information. For immediate purposes, then, aside from momentary discrete small-angle corrections, algorithms can be presented without yet distinguishing true ($\mathbf{I}_L$, $\mathbf{J}_L$, $\mathbf{K}_L$) from perceived ($\mathbf{I}_P$, $\mathbf{J}_P$, $\mathbf{K}_P$) nav reference frames.

3.3 ROTATIONS INVOLVING ANGLES FOR POSITION

Transformations between coordinate frames are expressed here in terms of 3×3 direction cosine matrices (denoted T with subscripts for axes) and – later, for attitude increments – in terms of a 4×1 array $\boldsymbol{q}$. Immediately it is emphasized that the latter is easily represented without quaternion algebra or any other specialized methodology. The corresponding quaternion would have the same elements as $\boldsymbol{q}$ but only the array – treated here as a 4×1 vector – will be needed.

The direction cosine matrix transformation *from* nav *to* vehicle axes has notation $\mathsf{T}_{A/L}$, with subscript sequence in conformance to the rudimentary property

$$[\mathbf{I}_A \ \ \mathbf{J}_A \ \ \mathbf{K}_A]^T = \mathsf{T}_{A/L}\,[\mathbf{I}_L \ \ \mathbf{J}_L \ \ \mathbf{K}_L]^T \tag{3.1}$$

The other direction cosine matrix of primary interest in this chapter is the transformation from earth to nav coordinates. With position expressed in terms of geodetic latitude ("Lat") and longitude ("Lon") that transformation is

$$\mathsf{T}_{L/E} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\sin(\mathrm{Lat})\cos(\mathrm{Lon}) & -\sin(\mathrm{Lat})\sin(\mathrm{Lon}) & \cos(\mathrm{Lat}) \\ -\sin(\mathrm{Lon}) & \cos(\mathrm{Lon}) & 0 \\ -\cos(\mathrm{Lat})\cos(\mathrm{Lon}) & -\cos(\mathrm{Lat})\sin(\mathrm{Lon}) & -\sin(\mathrm{Lat}) \end{bmatrix} \tag{3.2}$$

where α, the *wander angle*, has meaning only at nonpolar locations ($|\mathrm{Lat}| \neq \pi/2$); with double-subscripted T used here for elements of $\mathbf{T}_{L/E}$,

$$\alpha = \arccos\{ T_{13} / \cos(\mathrm{Lat}) \} = \arcsin\{ T_{23} / \cos(\mathrm{Lat}) \} \tag{3.3}$$

while, also at nonpolar latitudes,

$$\mathrm{Lon} = \arccos\{ -T_{31} / \cos(\mathrm{Lat}) \} = \arcsin\{ -T_{32} / \cos(\mathrm{Lat}) \} \tag{3.4}$$

and, at any latitude,

$$\mathrm{Lat} = -\mathrm{Arcsin}(T_{33}) \tag{3.5}$$

As appropriate, latitude is a principal value, always computed; longitude and wander angle are four-quadrant expressions evaluated only at nonpolar locations.

The last three equations constitute inversion of (3.2); nothing has yet been shown about the generation of $\mathbf{T}_{L/E}$ – *i.e.*, the fact that, aside from initialization, it provides the angles, not vice-versa. That explanation is involved with motion over the curved earth, which immediately draws attention to the next section's topic.

3.4 TRANSLATIONAL MOTION IN ROTATING FRAMES

Even with all the earth's irregularities, geographic locations can be accurately expressed in terms of departures from a reference ellipsoid of revolution. Since 1984 a "World Geodetic System" (WGS84) has been widely accepted, with parameter values including semimajor axis a_E and flattening f from [15]

$$a_E = 6378137 \text{ m}; \quad f = 1/298.257223563; \quad e_E^2 = (2 - f) f \tag{3.6}$$

The radius of curvature R_M in a meridian, and the radius of curvature R_P for planes parallel to the equator, both vary with latitude:

$$R_M = \frac{a_E (1 - e_E^2)}{[1 - e_E^2 \sin^2(\mathrm{Lat})]^{3/2}} \tag{3.7}$$

and

$$R_P = a_E / \sqrt{1 - e_E^2 \sin^2(\mathrm{Lat})} \tag{3.8}$$

At moderate latitudes the curvature radii R_M and R_P are commonly used to advance latitude and longitude, respectively, using North and East components of geographic velocity. To maintain operation even at the poles, however, a more versatile formulation is required. All situations are covered by forming the angular rate $\boldsymbol{\omega}_R$ of the nav reference ($\mathbf{I}_L$, $\mathbf{J}_L$, $\mathbf{K}_L$) *relative to* the earth frame ($\mathbf{I}_E$, $\mathbf{J}_E$, $\mathbf{K}_E$) and expressing it in ($\mathbf{I}_L$, $\mathbf{J}_L$, $\mathbf{K}_L$) – which was just shown to be locally level, offset in azimuth from North by the angle α. The angular rates, simply the ratios of those North and East velocity components to their appropriate curvature radii, are thus rotated through that wander angle,

$$\omega_R = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_{East} / (R_P + h) \\ -V_{North} / (R_M + h) \\ 0 \end{bmatrix} \tag{3.9}$$

If the denominators were equal this expression would reduce to the desired form, with components of the velocity vector $\mathbf{V}_L$ resolved along the nav frame. It is then quite helpful to construct the difference between radii, which reduces exactly to

$$r \triangleq R_P - R_M \equiv R_P e_E^2 \cos^2(\text{Lat}) / \left[1 - e_E^2 \sin^2(\text{Lat})\right] \tag{3.10}$$

Usage of this relation in the denominators produces, after some simplification, a highly accurate general expression for ω_R – without cardinal directions:

$$\omega_R = \begin{bmatrix} V_{L2} / (R_P + h) \\ -V_{L1} / (R_M + h) \\ 0 \end{bmatrix} + \frac{r \sin\alpha}{(a_E + h)^2} \begin{bmatrix} V_{L1} \cos\alpha + V_{L2} \sin\alpha \\ V_{L1} \sin\alpha - V_{L2} \cos\alpha \\ 0 \end{bmatrix} \tag{3.11}$$

Alternatively a_E in (3.11) can be replaced with an amount that varies by $\pm 0.34\,\%$:

$$\sqrt{R_M R_P} = a_E \sqrt{(1 - e_E^2)} \; / \; [1 - e_E^2 \sin^2(\text{Lat})] \tag{3.12}$$

but, in any case, the objective has been met. The relative angular rate ω_R is
- quite accurate (a few cm per km worst nav error), and in fact –
- exact in some cases (at the poles or at $\alpha = 0$ anywhere).
- only weakly dependent on α

For that last trait, any imperfection in computing ω_R (caused by ill-defined wander angle near the poles) is blunted by the factor $\cos^2(\text{Lat})$ in r. That is an essential characteristic of algorithm robustness; α can be accurately determined at all times except when its value doesn't matter.

Additional Discussion

The above material is followed by "cookbook" formulations that show incrementing of position, velocity, and attitude at each small time step. Readers unfamiliar with the background derivations for those expressions can find them in other texts. One possible reference for that purpose, *Integrated Aircraft Navigation* – written by this same author, gives the background developments with similar notation and perspective. Making the background presentations optional in that way enables readers to proceed with minimal distraction

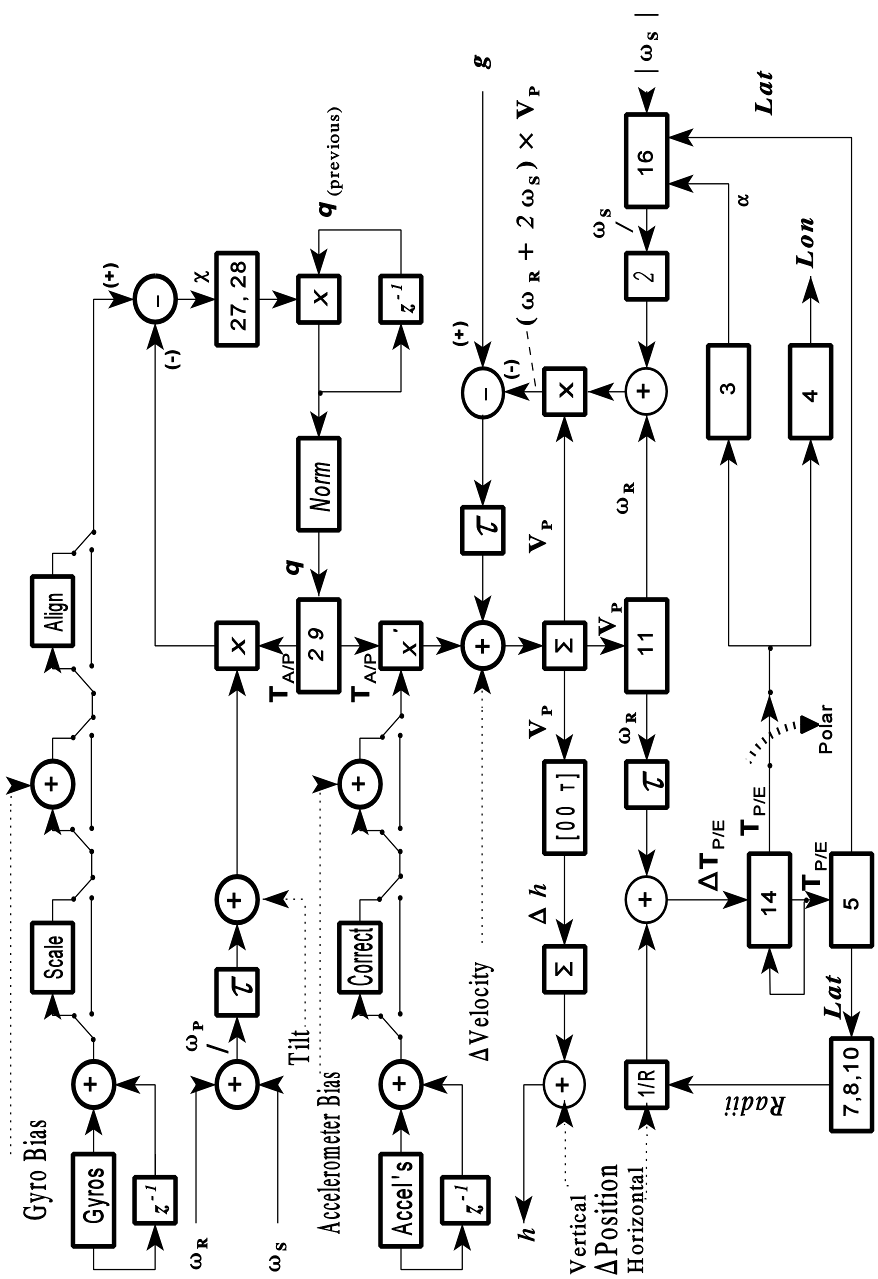

Figure 3.1 From raw inertial instrument data to final nav outputs

3.5.1 Attitude Incrementing Task List

This function, entered on alternating IMU data vector readings, performs a running account of all influences acting on the *4*x*1* matrix $\boldsymbol{q}$ and its corresponding *3*×*3* direction cosine matrix $\mathbf{T}_{A/P}$ expressing the instantaneous platform(nav)-to-vehicle frame transformation. To maintain clarity of code structuring, operations are partitioned into separate functions. Rotational increments accumulated during each brief interval are represented by small angle approximations throughout. After initialization logic (which detects first entry into the function and controls a modulo-2 count, while imposing negligible demands on computing load) the sequence of computations follows a straightforward task list:

- Sum rotations from two consecutive intervals (add the current and preceding intervals' "delta-theta" vectors)
- Apply gyro scale factor corrections from lab calibration data if present (this and other optional functions are shown in Figure 3.1 by "switch up" positions when active, or "switch down" positions if bypassed)
- Conditional drift adjustment (*i.e.*, if drift states are included in an active updating loop)
- Collect all gyro-related terms and transform from gyro-based axes to accelerometer-based axes, using mechanical mounting alignment adjustment data if present
- Combination of nav frame adjustments (its own rotation $\omega_L \tau$ plus any occasional Kalman filter misorientation state input)
- Compute net (vehicle / nav frame) relative rotation χ from (3.26)
- From (3.27,28) form the *4*×*4* coefficient matrix used to premultiply the previous *4*×*1* array $\boldsymbol{q}$, thus forming the current value for $\boldsymbol{q}$
- Occasionally normalize $\boldsymbol{q}$ (*i.e.*, divide each element by the root-sum-square of all four elements), every hundredth iteration
- Compute the *3*×*3* direction cosine matrix $\mathbf{T}_{A/P}$ from $\boldsymbol{q}$ by (3.29).

For those wanting detailed explanations of these steps [14] is cited: Eq. (2-19) for relative rotation rates; Eqs. (2-38–43) and Eqs. (2-74–77) for expressions involving $\boldsymbol{q}$. It is acknowledged that $\mathbf{T}_{A/P}$ could have been developed directly without $\boldsymbol{q}$ but that would have required a more complex normalization ([14], Section 2.B). Clarity is highly valued in this direct approach.

Note that $\mathbf{T}_{A/P}$ as just computed is applicable to the time at the end of two rotation-summing intervals – which (by reason of alternating cycles for rotation and translation incrementing) is just the time instant between two accelerometer -summing intervals, thus preventing what would otherwise be a dominant sculling effect. Note also that, by using relative (vehicle/nav) rotations in the above procedure, both the direction cosine matrix and $\boldsymbol{q}$ have a slowly rotating (rather than an inertial) reference. This corresponds to the common practice of separate row/column adjustments for vehicle (gyro) and nav (reference frame) rotations – but is slightly superior to row/column separations that use reduced reference adjustment rates.

3.5.2 Velocity Incrementing Task List

The list for translational incrementing is even simpler than the list given earlier:

- Sum outputs from two consecutive intervals (add the current and preceding intervals' time integral of $\mathbf{A}_A$, "delta-v" vectors from integrating accelerometers)
- Apply corrections for accelerometer scale factor and nonorthogonality, from lab calibration data if present
- Conditional offset adjustment (*i.e.*, for any accelerometer states included in a Kalman filter dynamic model if actively updating)
- Transform from accelerometer-based axes to nav coordinates, using the direction cosine matrix $\mathbf{T}_{A/P}$ obtained from attitude updating. Add any nonzero velocity error states occasionally available from Kalman filter updating.
- Include the gravity vector (in many short-term applications, a nominal z-component suffices)
- Add the product (*time step*)×kinematical adjustment vector { first term on the right of (3.22) } as shown in Figure 3.1.

3.5.3 Position Incrementing Task List

The following task list was also successfully used for development cited in [10-12]:

- Initialize position to approximate estimated longitude, geodetic latitude and altitude above the ellipsoid, with zero wander angle. Use (3.2) to form a corresponding direction cosine matrix $\mathbf{T}_{P/E}$ from earth to locally level, wander-azimuth (nav-reference) coordinates.

On each position computing event, perform all of the following steps:

- Evaluate the radius of curvature R_M in the meridian, curvature radius R_P for planes normal to meridians, and their difference r {see (3.7), (3.8), and (3.10) }. Use (3.11) to compute the angular rate of the slowly rotating nav reference frame relative to earth. Multiply that angular rate by the time step and combine with any nonzero x- and y- Kalman filter input position corrections, divided by their curvature radii. Recognize the overall result as a vector (having zero z-component) representing a small-angle rotation for the nav platform reference frame – which in this scheme has azimuth rotation limited to the component of earth sidereal rate projected about the vertical.
- Use the small-angle transformation (3.14), to rotate the platform reference axes relative to earth (the test described in [10-12] used double precision for this operation). This maintains precise position in terms of the 3×3 matrix $\mathbf{T}_{P/E}$, stable at all times.
- Use (3.5) to compute the new latitude and, if nonpolar, also form the new wander angle and longitude from (3.3,4). Optionally use trigonometry to express position with Cartesian representation, in conformance to any desired convention.
- Multiply the vertical velocity component by the time step and combine with any estimated z- component of position correction, producing the new altitude.

3.6 INTERPRETATION

It was already noted that direct processing methods offered here are valid for most – but not all – applications. Criteria for applicability of "short-term" conditions can be based on a nominal figure for maximum allowable blackout interval, T_B seconds without updates (during which the system "coasts" in free-inertial mode). Position error grows by amounts described in the next chapter and in Eq. (5-88) of [14], with conversion factors added for units used here. Immediate focus will be on biases from inertial instruments with their sensitive axes level. From accelerometer bias of $n\,g$ and a drift rate of $d\,°$/h (denoting, as always, amounts remaining after imperfect compensation), with $g = 9.8$ m/s^2, error in coast grows to

$$\tfrac{1}{2}\, n\, g\, T_B^{\,2} + \tfrac{1}{6}\, g\, d\, T_B^{\,3}\ (1/3600 \cdot 57.3) \doteq 5\, n\, T_B^{\,2} + 8 \bullet 10^{-6}\, d\, T_B^{\,3} \ \text{ meters} \qquad (3.30)$$

Various examples can use radically different parameter values for different conditions and applications, as follows:

- $n = 10^{-4}$, $d = 3$, $T_B = 100$ $\Rightarrow$ error growth < 30 m (LORAN aiding)
- $n = 0.01$, $d = 100$, $T_B = 25$ $\Rightarrow$ error growth < 45 m (LORAN aiding)
- $n = 0.001$, $d = 50$, $T_B = 2$ $\Rightarrow$ error growth < 2.5 cm (GNSS carrier phase)
- $n = 0.5 \times 10^{-4}$, $d = 1$, $T_B = 8$ $\Rightarrow$ error growth $= 2$ cm (GNSS carrier phase)
- $n = 10^{-4}$, $d = 1000$, $T_B = 1$ $\Rightarrow$ error growth < 1 cm (GNSS carrier phase)

The drift rate in the first example exceeds "nav-quality" value (commonly depicted at $0.01\,°$/h) by a large factor, and the last exceeds it by more than the square of that factor. In recent literature, both levels of performance have been described with the same catch-all phrase ("low-cost IMU"). The industry would benefit from an expanded vocabulary, providing different categories for very different levels of performance. For example, $\log_{10}(d)$ for gyros would assign "class 2" for $100\,°$/h. Because different T_B values will dictate which error dominates, accelerometers should be categorized separately (this would be preferable to a broad "IMU class").

These examples are oversimplified; a thorough evaluation would include additional motion-sensitive effects (*e.g.*, cross-axis degradations analyzed in [13], combined with vibration-induced rectifications [17] – see Addendum 4.B), while also accounting for random components of drifts and accelerometer errors. Still, (3.30) can provide a quick assessment. Significantly, the 1-cm/s performance described in [10-12] was obtained with an IMU having drift ratings more than five times higher than the largest value (in the last example) above – which was already 10^5 times "nav-quality" level.

Preliminary evaluations can also prescribe or exclude coning and sculling preprocessing. In any case, one concession is granted: gyro and accelerometer outputs are processed on alternating rotational and translational increment cycles. That was the *only* concession made throughout the successful tests just mentioned. Because that experience differs so conspicuously from conventional practice, an expanded discussion is now presented for comparison *vs*. customary processing approaches as defined in [9].

3.6.1 Departure From Conventional Methods

In discussing IMU data handling here, both rotational and translational incrementing are addressed with some comparison *vs*. methods used in [9]. To aid the comparison, two general contrasts will enhance clarity:

- Savage [9] provides insight into numerous individual contributors, propagating them through the full process separately. The direct approach collects all terms as early as possible and puts only the accumulated result (all of it; not piece-by-piece) into the requisite expressions.
- Coning and sculling are major issues throughout [9]. With preprocessing by [4] and/or repetitive updates, the direct approach suppresses their accumulations.

Conventional strapdown attitude algorithms were developed to counter a fictitious coning that would result from faithful adherence to sensed (*i.e.*, "reported" but physically improbable) multiaxis rotation sequences. The effect would propagate if unaided, but is repeatedly interrupted in a system with updating of IMU states. Even with low Kalman update rates, the attitude reset would nullify the growth of computational error. Thus in many applications not designed to accommodate long periods without external observations, computational steps used for counteracting pseudoconing can be omitted with no discernible performance penalty.

Most of [9] is quite accurately expressed, but one area needs revisiting. The first paragraph of the last page of Part 1 advocates avoidance of higher-than-necessary computation rates to prevent error buildup from certain effects including roundoff (which used to be a bigger factor than it is today) but also including "rectification of high-frequency multi-axis sensor errors ⋯ denoted as pseudoconing error ⋯."

In actuality, higher-than-necessary computation rates reduce pseudoconing error. Aside from instrumenting error (Section 4.3.1 of [14]), high-frequency multi-axis rates could be accommodated without rectification except for finiteness of sampling rates and of rotation increments that don't commute. The higher the sampling rate, the less pseudoconing error – until in the limit, infinitesimal rotation angles do commute, exactly, with no pseudoconing – no matter what motions occur at what frequencies with or without cross-axis correlations.

This author has concentrated much effort on sensor responses, motion spectra, higher-than-half-sampling-frequency content, cross-axis correlations, and rectification. Chapter 4 of [14] goes deeper into that than most people take time for; a brief synopsis is available from Eq. (3-61) plus Section 4.3.4 of [14]. Insight given therein is reduced due to 1960's/70's mechanization concepts but, of all the equations in Section 4.3.4, only the last needs any reinterpretation: replace the gyro quantization by an effective value based on resolution capability with whatever algorithm is being employed. The analytical expression for commutation error buildup was confirmed by rigorous simulation in [16]. In that simulation, a very broad range of rotation frequencies, cross-axis effects, sensor response characteristics and errors were considered; with up to a million computational iterations per run, the pseudoconing (referred to therein as "fictitious coning") error conformed to the analytical model.

Finally it remains to justify, in the successful development of [10-12], absence of detailed sculling compensation (*i.e.*, the previously described odd-even incrementing sequence was the only accommodation). Sculling is one of many rectifications arising from rotational / translational interaction; it could be ineffective to correct one without addressing the others – see Tables 4-2 (*p*. 123) and 4-3 (*p*. 128) plus Ex. 4-9, *p*. 140 of [14]. When sequentially correlated errors are unmodeled, an estimation algorithm attributes their effects to states that *are* included; here that means accelerometer bias in the vertical channel and, for horizontal directions, misorientation $\boldsymbol{\psi}$ (and, if included, drift rates). That type of modeling imperfection was already accepted in the attitude computation – because the coning correction, also prescribed at the high frequency (rather than the "moderate" computation rate, to use Savage's two-speed parlance), was also omitted. That term is likewise one of many (Table 4-1, *p*. 122 of [14]); in any system it could be

- neglected (as it was in the successful development of [10-12]),
- inserted at the high rate by preprocessing [4], or
- replaced by mechanized adjustment, as suggested near the bottom of *p*. 117 of [14] and in Addendum 7.A, if correlations are known.

3.7 SUMMARY

Complexity of customary processing methods for raw strapdown inertial instrument outputs can be supplanted by intuitive tasks shown here in Sections 3.4 and 3.5. All requisite operations can be carried out by elementary functions without esoteric techniques (*i.e.*, neither quaternion algebra nor slow/moderate/fast-computational cycle separation is needed), and only minimal extension is needed to reexpress the commonly used geographic reference in a stable wander azimuth form. The simplifications are permitted for reasons involving capabilities in computing and instrument sampling that were nonexistent when algorithms now used were developed. A versatile definition is provided for short-term conditions which in many applications eliminates the need to compensate coning and sculling (which "reincarnate" as noise plus contributions to slowly varying Kalman filter update states). All methods proposed herein were verified by a GPS / IMU integration which, with drift ratings above 5000 ^{0}/ h, produced RMS accuracies of *1* m in position, *1* cm / s in velocity, and a few tenths mrad tilt states; further flight validation is described in Chapter 8. Even when long-term free-inertial coast must be performed and / or IMU mounting structures endure extreme vibration environments, methods presented herein can be used, preceded by separate coning and sculling compensations based on [4].

Validation by successful tests cited herein has secured the principal goal of this development – delivery of high accuracy with robustness and versatility for widespread application.

Procedures herein, with an updated interface { [7,8] }, will prepare the industry for MEMS – finally making low-cost inertial navigation a reality.

CHAPTER 4

SHORT-TERM ERROR PROPAGATION WITH TRANSLATION AND ROTATION

Inertial instruments inherently sense motion derivatives {absolute angular rates for gyros; specific force (total nongravitational acceleration) for accelerometers}. To produce velocity and position their outputs must therefore be summed (integrated). Signals are thus formed as sequences of small angular increments ("delta theta") from gyros and velocity increments ("delta v") from accelerometers,.either mechanized as time integrals (as in rate integrating gyros and integrating accelerometers) or sampled at a high rate and multiplied by the intersampling interval. Whether those increments are processed conventionally or by methods shown in Chapter 3, the results contain *cumulative* effects of sensing imperfections (*e.g.*, a biased angular rate measurement causes a growing error in perceived attitude). Before completion of any detailed plan to counteract accumulation, behavior and interactions must be clearly understood. Fortunately a straightforward characterization is available and, once the ramifications are clear, applicable even to operations involving complex motion patterns.

For reasons more fully clarified in Chapter 5, priority is given here to estimation of dynamics apart from location. Chapter 2 scrutinized dynamics for various translational motion models, including error state histories – and also including coupled velocity-cum-acceleration dynamics – independent of position. This chapter defines dynamic models for errors in velocity and tilt (misorientation about level axes), with generalization to three dimensions, and expands the scope to add inertial instrument biases. It is then shown how state formulations with different dimensionality can easily adapt to presence or absence of various conditions or features. Matrices are given for formulations, in preparation for the next chapter (wherein matrices formed over each short IMU sampling interval are cascaded, to span intervals always reinitialized to the time of the most recent update observation).

In dynamic estimation without location states, position-dependent measurements are related to *time integrals of lowest-order states*. To accommodate this situation, computational procedures go beyond the ordinary. Requisite operations are introduced here and, for those preferring a more customary approach, incremental position states are subsequently added for an alternative conventional formulation.

The chapter ends with expanded explanations for remaining pertinent topics, followed by appendices containing additional error characteristics.

4.7 SUMMARY

A key step in aided inertial system design is realistic dynamic modeling, covering both state and covariance propagation. This chapter addresses both fidelity and practicality of the overall modeling process. Departures from convention, adopted here, include

- restriction to short-term applications, so that only drift rates (and not inexact velocity) contribute to the misorientation dynamics
- separate estimation of dynamics (with velocity-related observables) from location (with position-related observables)
- optional exclusion of error state components (azimuth, drift, accelerometer offset)
- determination of all process noise spectral densities from effective estimation memory spans based on conservatively chosen data-averaging durations
- attitude transient suppression via thorough usage of translational and rotational data
- detailed models of inertial instrument errors, including motion-sensitive effects

The first three items, first fully developed in [1]-[3], were influenced by the advent of MEMS (microelectromechanical sensors [4]). They carry ramifications both obvious and subtle. Influence on formulation is shown in this chapter; a thorough explanation must await the next. The fourth item has roots over a generation old, used in tracking applications of Chapter 9 and documented in [6]. The fifth is covered in Sections 4.5.1 and 4.6.

Proposed handling of motion-sensitive gyro and accelerometer degradations, spanning a period of decades [7-11], is quite far-reaching. The scope of instrument degradation is multifaceted; covering the variety of effects has necessitated addition of appreciable length to this chapter. Viability of (4.10) is a welcome benefit but, with insufficient scrutiny, has led to oversimplification and complacency. Familiar effects (*e.g.*, turn-on to turn-on variations, thermal sensitivities) are accompanied by motion-dependent errors which, though widely known, have not yet produced firm commitments in IMU specifications. Implications are confronted at length in Appendices of this chapter, prompting ***recommendations for major changes in the way specifications are defined for an IMU with extended free-inertial coast requirements***. Conservative values should be required, tied to specific probabilistic levels, for coefficients identified in Addendum 4.B – expressed generically in units of quantities measured (angular rate or specific force, not the physics of any particular sensor) – so that decisions are independent of sensing technologies.

Time histories of those angular rates and forces offer standardized benchmark case scenarios for comparing alternative IMU mechanizations. Combined with mission profiles and vibration properties at IMU mounting sites, coefficient values just described can enable definite accept/reject decisions of IMUs for specific operations. When all those steps are taken, everything humanly possible is done to ensure model fidelity – a vital factor for success in operation. Without that (*e.g.*, by incomplete coefficient sets and/or uncommitted descriptions for "typical" performance), success or failure in a given application is inescapably – and often highly – uncertain.

4.B.2 Approach

Evaluation of navigation performance as a function of inertial instrument errors, covered in other sections and in numerous publications elsewhere, need not be repeated here. As already explained the immediate focus is determination of gyro and accelerometer error sources, with emphasis on contributions not scrutinized customarily. This goal is approached at an intermediate level of detail, to provide enough insight enabling informed decisions – but without digression for an unduly long and distracting analytical development. To make that insight a reality, while still observing the need for comprehensibility, certain elements of common knowledge are taken into account – but without rigorous insistence on exactness.

Discussion begins with a configuration using triads of integrating accelerometers and rate-integrating gyros, assembled at mutually orthogonal nominal orientation. Regardless of mechanization, each instrument has an input (sensitive) axis x, an output axis y, and another axis z perpendicular to both x and y. The input axis of each gyro is intended to be parallel to an accelerometer input axis and perpendicular to all other instrument input axes. Imperfect construction raises the prospect for any instrument (*e.g.*, a roll gyro) to include a small fraction of the motion intended for another (*e.g.*, a yaw gyro; for a *0.1*-mrad misalignment the fraction is *0.0001*). Imperfect scaling raises another prospect for an instrument to accumulate slightly incorrect amounts (*e.g.*, 100.1^0 instead of 100^0 for *0.001* scale factor error). The seemingly small amounts used in these examples are quite significant when verticality to within tenths of a milliradian is desired.

This is an appropriate point to raise another common practice affecting accelerometer and gyro errors in general: instrument sensitivities to various motions are observed by test, and firmware provisions enable compensations to occur in operation, to deactivate much degradation at the source. All further performance analysis can then be interpreted as applicable to the uncompensated amount that either (1) remains after imperfect tests, or (2) changes due to various factors including aging and thermal effects. This issue also clarifies a fivefold purpose for directing IMU suppliers to define sets of error coefficients:

- attention drawn to the full set of possible instrument degradations
- motivation to perform thorough lab calibrations
- dependable assessment of amounts that can be left uncompensated
- more thoroughness and commitment than what is prevalent now in specifications
- attention drawn to dominant effects that limit achievable performance – therefore lighting the way toward the most effective improvement strategy.

Vibrations generating bias-like effects can have complex waveforms, but immersion into complexity can be avoided. First consider an oscillatory angular rate $\omega \sin(2\pi ft)$; both the corresponding angular acceleration (derivative) and rotational excursion (integral) are proportional to $\cos(2\pi ft)$. Then replace the sine wave by a band of sinusoids. That narrowband Fourier spectrum, a standard model for complex random waveforms, preserves the way toward powerful insights, as follows.

The basic angular rate { $\sin(2\pi ft)$ } is in phase quadrature with both the corresponding angular acceleration and rotational excursion { $\cos(2\pi ft)$ }. Also the average product <sin • cos> is zero because its time integral (a double-frequency sinusoid) tends to zero; thus the rate is uncorrelated with both the rotational excursion and angular acceleration. *The same is largely true of narrowband waveforms*; for derivatives or integrals, all cosines in the Fourier expansion are replaced by sines and vice-versa – with algebraic sign changes in one direction but not the other).[1] Thus components of angular acceleration and excursion have essentially

- zero correlation with the vibratory angular rate (so that the mean product of a rate component multiplied by its corresponding excursion is zero)
- full negative correlation with each other (so that the mean product of any acceleration component multiplied by its corresponding excursion is equal to the negative product of their RMS values).

These concepts, standard in communications analysis, are key to the next section wherein random IMU error models are characterized. Another enabler, likewise fundamental, is the obvious intrinsic coupling between a vibratory angular rate $\boldsymbol{\omega}$ and the vibratory translational velocity $\omega \times$ (**lever arm**) it produces.

With a clear understanding of fundamentals just presented, preparation is complete for IMU evaluation – avoiding unnecessary complexity but still enabling meaningful quantitative inference. While that cannot be achieved by simply ignoring inherent complication, opportunities have been exploited to provide insights. While not necessarily "easy reading" for all, the level of detail reasonably steers between extremes of undue complexity vs. oversimplification.

4.B.2.1 Motion-dependent Errors In An IMU

Each gyro and each accelerometer can be adversely affected by every component of rotation and specific force it experiences – and by *products* of those components. An obscure publication [7] advocated a convenient way to systematize a wide collection of motion-sensitive degradations. A generalization of the approach advocated therein will be adopted here. First define a column array $\boldsymbol{\chi}$ with twelve components:

- three components (ω_1, ω_2, ω_3) of a vibratory angular rate $\boldsymbol{\omega}$
- three components (ω_1*, ω_2*, ω_3*) in phase quadrature to that angular rate
- three components (A_1, A_2, A_3) of specific force **A**
- three components (A_1*, A_2*, A_3*) in phase quadrature to that force

[1] The above description neglects the change in spectral shape due to nonuniform frequencies that are multiplicative factors in differentiation and divisors in integration. Narrowband waveforms, however, have small ratios of bandwidth to center frequency; scaling with a constant center frequency throughout is a permissible approximation. Another assumption in the above description, again standard, is equivalence of time averages to ensemble averages. Finally the same phase quadrature relation for derivatives and integrals in rotation (excursion, rate, and angular acceleration) will apply also to translational motion (incremental position/velocity/acceleration).

For quadrature components, visualize every cosine in a spectral representation replaced by a sine, and vice-versa (therein lies the reason for generalizing the approach in [7], which omitted the important quadrature waveforms). Again, algebraic sign changes accompany one (not both) replacement operations; recall the association with derivatives and integrals (for sines and cosines) previously discussed.

The transposed array χ^{T} arranges the components into a row,

$$\chi^{\mathrm{T}} = [\ \omega_1\ \omega_2\ \omega_3\ \omega_1{*}\ \omega_2{*}\ \omega_3{*}\ \ A_1\ A_2\ A_3\ A_1{*}\ A_2{*}\ A_3{*}\ ] \tag{4.42}$$

An extensive set of motion-sensitive degradations is covered by an array pair for gyros and another pair for accelerometers. Each pair includes a 12×1 vector **b** or **c** and a 12×12 matrix **B** or **C**. The effects are contained in simple but versatile expressions

$$gyro\ error = \chi^{\mathrm{T}}\,\mathbf{c} + \chi^{\mathrm{T}}\,\mathbf{C}\,\chi \tag{4.43}$$

and

$$accelerometer\ error = \chi^{\mathrm{T}}\,\mathbf{b} + \chi^{\mathrm{T}}\,\mathbf{B}\,\chi \tag{4.44}$$

Gyro response to motions will first be illustrated for familiar instrument types. Let subscripts $1/2/3$ represent roll/pitch/yaw axes, respectively. The first component of **c** is then a roll gyro scale factor offset. The second (third) component is the roll gyro's unit input axis projection along pitch (yaw). Either c_8 or c_9 is the cross-axis g-sensitive coefficient (depending on how the gyro is mounted; for the old spinning gyros that would be mass unbalance along the spin axis; c_7 would be input axis mass unbalance). If that mounting makes the cross-axis g-sensitive coefficient of a spinning gyro the eighth (ninth) element of **c** then the sixth (fifth) component is the sensitivity to output-axis angular acceleration ("J/H" in a spinning gyro).

For a roll accelerometer the seventh element of **b** is the scale factor offset, while the eighth and ninth elements are that accelerometer's sensitive axis components along pitch and yaw, respectively. Sensitivity to cross-angular acceleration can be either the fifth or sixth element (mounting-dependent).

Motion products are also readily represented by (4.43,44). Again for a spinning-gyro example, two familiar products are anisoelasticity and anisoinertia. The former involves either C_{78} or C_{79}; the latter, either C_{12} or C_{13}. Roll accelerometer vibropendulous coefficient is B_{78} or B_{79} (again dependent on instrument mounting). These examples are for so-called "g-squared" and "angular rate squared" sensitivities, both involving rectification. The phenomenon is easy to understand in this way: Subjecting an inertial instrument to g-forces causes its sensitive element to move off its reference null position. The momentary offset, proportional to the g-force, creates a temporary sensitivity to another g-force instantaneously along an orthogonal axis – so that the overall effect is proportional to the product of two perpendicular components of specific force. If those components are correlated (analogous to two sine waves in phase with each other; easily realized in practice by a vibration axis in a skewed direction), the average value of the product is not zero (*e.g.*, consider squaring a cosine or sine wave - either wave has zero mean but nonzero mean square).

Another characteristic of the examples just given for rectification products is generation wholly within an instrument. A more subtle rectification phenomenon involves interaction between

- an error component proportional to a zero-mean vibratory motion and
- a direction cosine matrix element

As one example, consider the accelerometer error caused by cross-axis angular acceleration. Since the total observed specific force (which includes that contribution) is repetitively multiplied by the direction cosine matrix (for transformation into nav reference axes), there are products involving time derivatives and time integrals of angular rate. It has already been explained how that produces rectification with narrowband vibration waveforms; for those wanting to scrutinize the issue, Ex. 4-9 of [6] shows all details. In fact the entire fourth chapter (the "forgotten chapter") of [6] derives and tabulates a host of rectification sources. Most serious among them are the ones contributing to rotational drift rates – the example just given is readily applicable to drift, as will now be shown (again using a spinning gyro):

Total drift includes a term proportional to output-axis angular acceleration. The time integral of a gyro's drift rate (a component of a vector denoted ξ) thus includes a term proportional to output-axis angular *rate* – immediately producing rectification in the cross product $\omega \times \xi$ (again this invokes the phase quadrature issue; this and the cross product, as well as the reason for generating ξ itself, appeared with verification by rigorous simulation in [8], reiterated on *pp*. 116-117 of [6]).

Expansion from the examples just given to a broad class of motion-sensitive errors can now follow from Table 4.2, depicting the types of vibratory motion under consideration:

Table 4.2

Products of vibratory motion components

	ω_x	ω_y	ω_z	ω_x^*	ω_y^*	ω_z^*	A_x	A_y	A_z	A_x^*	A_y^*	A_z^*
ω_x	a	b	b		c	c		•	•		g	g
ω_y		a	b			c	•		•	g		g
ω_z			a				•	•		g	g	
ω_x^*												
ω_y^*												
ω_z^*												
A_x							d	e	e		f	f
A_y								d	e			f
A_z									d			
A_x^*												
A_y^*												
A_z^*												

CHAPTER 5

CODE AND CARRIER DIFFERENCING

The Global Navigation Satellite System (GNSS) includes the Global Positioning System (GPS) plus all other satellite systems in operation. Examples used to illustrate principles in formulation and development here are strongly related to GPS, since it has thus far provided the broadest base of experience. However, decisions affecting development herein would be largely the same for any code on any carrier from any space (*e.g.*, GPS, GALILEO, GLONASS, Beidou) – or terrestrial (radio, TV) station.

In any discussion involving GPS, a few acronyms arise repeatedly, *e.g.*,

- SV – space vehicle; synonymous with "satellite"
- C/A – coarse acquisition – 1023-bit Gold codes with *1*-ms period, chosen for multiple access performance (low cross-correlation between separate SV signals)
- SA – selective availability – degradations formerly inserted into GPS signals, intended to deny full capability to users lacking privileged information.

Except to clarify terminology or issues central to the approach being presented, little attempt is made here to describe satellite navigation systems; excellent books such as [1–3] can readily provide the needed material. After a very brief discussion of measurement data, attention is drawn immediately to measurement differencing, fully used herein to suppress all major error sources, particularly biases.

As with all other facets of this approach, robustness has top priority. The need to achieve robustness in the presence of degradations encountered with low-cost equipment has led to several departures from customary approaches. For the choice of segmented estimation (see discussion preceding Section 1.1), these include

- Exploitation of carrier phase without integer ambiguity resolution. Discussion of ramifications is deferred until after all differencing operations have been defined.
- Velocity history feedforward to a separately adjusted *3*-state estimator for position.
- Full compatibility with dynamic models presented in Chapter 4.
- Extensive usage of triangular matrix forms [4] – not only upper but also lower.
- An original lower triangular matrix form for correlated across-SV differences.
- Standard lever arm adjustment accompanied by more thorough impact on $\mathbf{H}$ [5].

All of these features were present while processing inputs – exclusively raw data:

- small-angle increments from gyros and velocity increments from accelerometers
- ambiguous carrier phase and pseudorange, plus one nav message data set

– to generate the test results presented in Chapter 8.

5.2 ACROSS-SV DIFFERENCING

With (2.58) applied to both SV $\#j$ and SV $\#k$ at the same time, the result is

$$\begin{aligned} Y_{\rho j} &= |-\mathbf{S}_j + \mathbf{R}_m| + Iono_j + Tropo_j + B_u + C_u - C_j \\ Y_{\rho k} &= |-\mathbf{S}_k + \mathbf{R}_m| + Iono_k + Tropo_k + B_u + C_u - C_k \end{aligned} \tag{5.7}$$

so that when they are subtracted, both B_u and C_u cancel. The difference between two simultaneously observed pseudoranges can use the notation

$$\nabla Y_\rho = Y_{\rho j} - Y_{\rho k} = |-\mathbf{S}_j + \mathbf{R}_m| - |-\mathbf{S}_k + \mathbf{R}_m| + \nabla Iono_{jk} + \nabla Tropo_{jk} - C_j + C_k \tag{5.8}$$

so that, when the same reasoning used in forming (2.59) – (2.61) is reapplied, the following notation for the residual is useful:

$$z_\nabla = \mathbf{h}_m^{\mathrm{T}} \mathbf{r}_m + \epsilon_j - \epsilon_k \ , \ \mathbf{r}_m = \mathbf{R}_m - \hat{\mathbf{R}}_m \ , \mathbf{h}_m = (\mathbf{1}_{mj} - \mathbf{1}_{mk})^{\mathrm{T}} \tag{5.9}$$

Appropriately, no fourth (user clock) state is present – but this measurement error contains a reminder of how it was formed. That is useful when the process is repeated with another satellite (*e.g.*, SV $\#n$) replacing SV $\#j$ in (5.9); the result would have an error $\epsilon_n - \epsilon_k$ correlated with the original difference. If the observation error from each individual SV can be represented by

- independence from all other SV observation errors,
- zero mean, and
- RMS value σ_1

then the variance of each difference is $\sigma_\nabla^2 = 2\sigma_1^2$ and the correlation coefficient is ½:

$$< (\epsilon_j - \epsilon_k)(\epsilon_n - \epsilon_k) > \ = \ < \epsilon_k^2 > \ = \ \sigma_1^2 \ = \ \tfrac{1}{2}\sigma_\nabla^2 \tag{5.10}$$

It is now instructive to revisit (5.1) in an alternate but equivalent form. Instead of four individual satellite observations there are three differences with a common SV used for subtraction – eliminating the clock state:

$$\mathbf{z} = \mathbf{H}\,\mathbf{r} + \epsilon, \quad \mathbf{H} = \begin{bmatrix} \mathbf{1}_{m1}^{\mathrm{T}} - \mathbf{1}_{m4}^{\mathrm{T}} \\ \mathbf{1}_{m2}^{\mathrm{T}} - \mathbf{1}_{m4}^{\mathrm{T}} \\ \mathbf{1}_{m3}^{\mathrm{T}} - \mathbf{1}_{m4}^{\mathrm{T}} \end{bmatrix} \tag{5.11}$$

The solution formed by inversion at time t_m (now denoted $\hat{\mathbf{r}}_m$, again with nonzero determinant $|\mathbf{H}_m|$) of course conforms to (5.2) and produces a new error state (now expressed as $\mathbf{r}_m = \mathbf{H}_m^{-1}\,\epsilon_m$) similar to (5.4) – but now with covariances from

$$< \epsilon_m \epsilon_m^{\mathrm{T}} > \ = \ \sigma_\nabla^2 \begin{bmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix} \tag{5.12}$$

A convenient reexpression for $< \epsilon_m \, \epsilon_m^{\mathrm{T}} >$ in this context is

$$< \epsilon_m \, \epsilon_m^{\mathrm{T}} > \; = \sigma_{\nabla}^2 \, \gamma^2 \; , \quad \gamma \; = \; \frac{1}{3\sqrt{2}} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix} \tag{5.13}$$

so that, as is easily verified,

$$\gamma \; = \; \kappa^{-1} \; , \quad \kappa \; = \; \frac{1}{3\sqrt{2}} \begin{bmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{bmatrix} \tag{5.14}$$

Because κ is an inverse square root of the normalized matrix in (5.12), these definitions enable formation of a DOP matrix similar to – and consistent with – (5.5),

$$< \tilde{\mathbf{r}}_m \, \tilde{\mathbf{r}}_m^{\mathrm{T}} > \; = \; < \mathbf{r}_m \, \mathbf{r}_m^{\mathrm{T}} > \; = \; \sigma_{\nabla}^2 \; \boldsymbol{H}_m^{-1} \, \boldsymbol{H}_m^{-\mathrm{T}} \; , \quad \boldsymbol{H}_m \; = \; \kappa \, \mathbf{H}_m \tag{5.15}$$

As discussed in [7], linear transformation of (5.1) to (5.11) must produce consistent results. Appendix A of [8] presents an example illustrating

- consistency between (5.5) and (5.15) – mindful that $\sigma_{\nabla}^2 = 2\sigma_1^2$ – and
- necessity of including correlations to avoid inconsistent DOP values.

Expressions involving the error covariance matrix for measurements at t_m are typified for a set of three in (5.12). RMS error in a component of ϵ is denoted by σ – whether a single or double difference (with or without a ground station, so that it includes any factors of *1.414*). When more measurements are included, the form of (5.12) remains applicable (unity on the diagonal, ½ elsewhere), but a different type of square root matrix is needed – a lower triangular matrix $\mathbf{L}$ with a property noted on *p*. 47 of [4]

$$\mathbf{L} \, \mathbf{L}^{\mathrm{T}} \; = \; < \epsilon \, \epsilon^{\mathrm{T}} > \, / \, \sigma^2 \tag{5.16}$$

The reason is that, although the correlations being addressed are for multiple measurements at the same time, they are processed in sequence (an issue not to be confused with sequential correlation effects, addressed in Section 5.6).

Premultiplication of (2.36) by $\mathbf{C} = \mathbf{L}^{-1}$ produces a modified expression,

$$\mathbf{C}\mathbf{z} \; = \; \mathbf{C}\mathbf{H}\mathbf{x} \; + \; \mathbf{C}\epsilon \tag{5.17}$$

enabling estimation of $\mathbf{x}$ from a weighted residual vector $\mathbf{C}\mathbf{z}$ with weighted sensitivity $\mathbf{C}\mathbf{H}$ and, from (5.16), with uncorrelated measurement error;

$$\mathbf{C} \; < \epsilon \, \epsilon^{\mathrm{T}} > \; \mathbf{C}^{\mathrm{T}} \; = \; \sigma^2 \; \mathbf{C} \, \mathbf{L} \, \mathbf{L}^{\mathrm{T}} \, \mathbf{C}^{\mathrm{T}} \tag{5.18}$$

which reduces to the identity matrix scaled by measurement variance σ^2. $\mathbf{L}$ and $\mathbf{C}$ – chosen dimensionless here – otherwise match the characteristics used in [4]. With both $\mathbf{H}$ and the residuals at time t_m weighted by these dimensionless elements, the result is independence with full flexibility applicable to measurement sets at any time.

The lower triangular form of **C** facilitates one-at-a-time residual processing thus: Measurements at any time t_m comprise a set. The first measurement of each set has error ϵ independent of preceding measurement errors (this simplification – sequential differences involve a common component – is resolved in Section 5.6). The second residual of each set is recomputed as a linear combination of itself and the first in the set. The third residual of each set is recomputed from a linear combination of itself and the first two residuals in the set. The pattern continues, with each new addition involving previous – but not subsequent – data. For error covariance matrices in the form of (5.12) – unity on the diagonal, ½ elsewhere – an extremely useful recursion found for **C** is exemplified by this MATLAB®[§] program for seven difference measurements (eight SVs):

```
C = zeros(7,7);
C(1,1) = 1;   K = 1;   frac = 1;
for m=2:7
  K = K + m ;
  for j=1:m-1, C(m,j)=-1/sqrt(K); end
  frac = frac + 1 / K ;    C(m,m) = sqrt (frac) ;
end
```

Conformance of $\mathbf{C}^{-1}$ to the property of (5.12) is easily verified.

5.3 DIFFERENCING ACROSS RECEIVERS

The opportunity to use a reference receiver for canceling major error sources was recognized long ago. For local area differential GPS the basic principle is subtraction of simultaneous (or computationally synchronized) measurements from two different receivers in proximity, having similar effects of propagation, ephemeris, and inexact SV data (plus SA when it was active). Combined with across-SV subtraction the resulting *double differencing* became widely accepted as a powerful means of minimizing systematic error sources. Consider two receivers employing (5.8) with the same SV pair while experiencing essentially the same ionospheric and tropospheric effects – so that they, as well as SV clock offsets, largely cancel in subtraction. With the reference receiver at a location designated as $\mathbf{R}_0$ – and all cancellations successful enough to leave only minor contributions to ϵ – this double difference has the form

$$\nabla\Delta Y_\rho = \Delta Y_{\rho j} - \Delta Y_{\rho k} = |-\mathbf{S}_j + \mathbf{R}_m| - |-\mathbf{S}_k + \mathbf{R}_m| - |-\mathbf{S}_j + \mathbf{R}_0| + |-\mathbf{S}_k + \mathbf{R}_0| + \epsilon \quad (5.19)$$

For application of (5.9 – 15) $\mathbf{r}_m$ adjusts the *baseline* from $\mathbf{R}_0$ to $\mathbf{R}_m$, $\mathbf{1}_{m\,j}$ and $\mathbf{1}_{m\,k}$ point to mid-baseline, and the variance of ϵ is increased again – even if all propagation errors cancel exactly, at least the noise is now $4\sigma_1^2$.

Ground receivers are impractical for many applications. The next section facilitates operation without them.

§ MATLAB is a registered tredemark of The MathWorks, Inc.

Performance at these levels will not be seen in specifications any time soon. Inertial instruments today – and for some time in the future – can maintain accuracies commensurate with carrier phase precision for very limited durations only. While integrated systems do sustain cm/s velocity accuracy, it is clearly the repetitive updating – not today's IMU – providing the capability. Free-inertial coast over longer durations is of course achievable, but with far higher error levels (*i.e.*, recapture of sub-wavelength accuracies requires reacquisition). Obvious as that may seem, implications are revisited to ensure clarification:

It has just been shown that, in attempts to maintain cm/s velocity accuracy – even for only a few minutes, practical experience does not conform to theoretical propagation behavior. The same source [18] that supplied data for Table 5.1 clearly explains: T must not exceed model fidelity limits, and dimensionality is also restricted thereby (adding states or extending T can be not only futile but highly detrimental). Balance must be struck between random and systematic errors. Since the full set of gyro and accelerometer degradations defies precise characterization (Addendum 4.B), experience provides the most dependable indication. For a few tenths mrad leveling, corresponding to a few hundred μg, a comparison of models enables determination of durations for which theory *does* match practice. With $T = 12$ replacing 180 in the last full (*4*-state) model example shown, RMS error in x_3 is about $330\ \mu g$ (3σ just under 1 mg); for $T=16$ the RMS is $160\ \mu g$. Actual improvements would be smaller, due to positive sequential correlation effects of plant noise noted in Section 5.5.3. Furthermore, with spans this brief, the conservative (*3*-state) model provides comparable success for x_2: $480\ \mu g$ RMS at $T = 12$ or $312\ \mu g$ for $T = 16$.

Higher-order models, containing larger numerators, are effective only when they maintain fidelity over sufficient durations. That has always been true of data fitting in general and, at *1* Hz ($\tau = 1$ s) for the two models under scrutiny here, the theoretical break-even point for tilt-bias ($72\sqrt{5}/T^2 = \sqrt{2} \bullet 8\sqrt{3}/T$) is approximately 8.2 s. With shorter T the conservative model gives superior performance; longer T favors the full model. Significance of this development can now be brought to focus:

- At $T=12$ s (producing 3σ leveling error at 1 mg), the full-model advantage is not dramatic. Improvement via longer T risks model mismatch – practical experience seldom produces RMS errors below $100\ \mu g$, even for state-of-the-art systems.
- Efforts to reduce error by faster sampling (same T with lower τ) could be futile – a drift state constant throughout T is a useful (not rigorous) model expedient.
- *Potential* advantage of the full model occurs only with long T. Conservative model performance can mimic that of the full model dynamic states at any T/τ with sequential correlations taken into account – *and* at practical T/τ values with those correlations neglected. That omission, then, sacrifices little achievable accuracy.

State-of-the-art GPS/INS performance (a few tenths mrad leveling, producing a few hundred μg in coast) can therefore be obtained with consecutive GPS carrier phase differences at *1* Hz treated as if they were sequentially uncorrelated. This analytical result, substantiated by test results in Chapter 8 and also in [15], is due to limited durations T – not restricted to low-cost systems.

5.6.2 Block Processing

Although validated both analytically and by test results, a sequentially uncorrelated error model for carrier phase differences offers further insight from another view. A *7*-minute van test run from [13] was repeated with the dynamics segment in block form. Again correlations from across-SV differencing were included in the program, but separate runs were made – first with, then without, sequential correlation effects. The latter did not use any of the classical methods reviewed for this application in Addendum 5.B.1; instead, sequential correlations were derived from off-diagonal elements inserted into the measurement error covariance matrix. The method, shown in Addendum 5.B.2, is not of primary interest; even the unsurprising results are not in themselves important. From that endeavor only the conclusions matter: accuracy did not improve when sequential correlation effects were added – but the exercise highlighted block / recursion similarities and differences, now to be described.

The nominal correspondence of averaging durations in Sections 4.5 and 5.6.1 produced similar results for velocity accuracy by block and recursion methods. With process noise only in the latter, however, the agreement is not exact. The last remark is more true of higher order states; verticality showed substantially less agreement between block and recursion. The reason is clear from action of process noise variance. Emphasis of that action has thus far focused on limiting effective estimation memory to short durations. With recursion, however, its presence *within* those durations depreciates *all* data – including even the data within – according to age via increased variance for all states at every update. Forfeiting the modest relative emphasis of new data reduces responsiveness of block estimators without plant noise.

Memory fading [18] can remedy the shortcoming just identified, whether using unambiguous phase or accounting for sequential correlations in a dynamics segment. Fading heeds dynamic uncertainty at realistic levels throughout a data sequence. Accelerometer noise at *1*-mg, for example, can add a half-cm phase error between *1*-Hz updates. Even if hovering at just $10\ \mu g$, an error beyond a quarter-wavelength ($\lambda/4$) for GPS C/A accumulates (via $½ \bullet 10^{-5} \bullet g \bullet t^2$) in about a half-minute. For $100\ \mu g$ the duration shrinks to about *10* s ; for a milli-g, only *3* s. With no acceleration error and gyro drift rates of 0.01^{o} / h , growth via $^1/_6 \bullet drift \bullet g \bullet t^3$ (integral of the expression just used with drift in rad/s) would still exceed $\lambda/4$ before *1.5* minutes. At 0.001^{o} / h $\lambda/4$ is reached before *4* % of a Schuler period – even "tomorrow’s" INS will not maintain *1* cm/s velocity without frequent updating. *1*-Hz is practical; higher rates sacrifice coherence in the segmented estimator, while lower rates compromise robustness – in any case Morrison [18] gives the needed guidance.

The examples are not precise; replacing the biases by spectral densities would reduce the exponents of time by ½ – but the conclusion would not change: ***Estimators continuously maintaining GNSS-era dynamic accuracies (cm/s velocity) have short spans for observables. In view of Table 5.1, then, carrier phase changes updating IMUs at 1 Hz can be processed as sequentially independent data without significantly degrading dynamic estimation***.

5.6.3 Phase Continuity And State-of-the-Art

This section has highlighted some ironic performance features. Effort was directed toward accounting for sequential correlations in observation errors, with the outcome that they can be ignored. Also, state-of-the-art carrier phase information from all SVs does not by itself continuously provide precise instantaneous velocity. For example, the total change in phase over a *1*-s interval obviously does not define detailed history within that interval. There are operations needing those details {*e.g.*, synthetic aperture radar (SAR) }, but they are not obtainable by drastically reducing the interval {recall the discussion following (1.8) and see the discussion of receiver track loop response in Chapter 7}. Whether to obtain detailed phase history over diminutive subintervals or simply to maintain phase precision over the full interval – in the presence of high dynamics but without overtaxing the receiver response – inertial information is used. IMU behavior, however, presents another irony:

Continuous monitoring of GNSS phase provides highly accurate history of excursions, raising expectations for refinement of other sensors simultaneously monitoring that same motion. Those expectations are dependably realized *provided that* realistic characterization is given to behavior of the other sensors (*e.g.*, gyros and accelerometers). IMU "calibration" is clearly made perishable by sensitivity to conditions that subsequently change. Scale factor and cross-axis effects in the presence of climbs, turns, or speed changes only begin to list the changes; compounding factors include absence of a data base for vertical deflections {recall the discussion involving $\mathbf{n}_a$ and (4.3) } and a host of vibration-sensitive degradations (Addendum 4.B). Total variation not only imposes demands to limit the duration for data averaging but, as seen from Section 5.6.2, even the variation *within* it is significant. That behavior is not restricted to low-cost equipment; the need just mentioned for realistic characterization, applied with typical values for uncorrected vertical deflections and IMU error coefficients, has produced a landmark conclusion: ***At accuracy levels offered by GNSS, short-term operation is the only option***.

Immediately a long list of past successful long-term applications could be raised – where's the contradiction? Actually there is none; the explanation clearly lies in the scale of accuracy. The systems of yesteryear were adequate for levels of performance prevalent a few decades ago. Classical INS performance is commonly expressed in terms of nautical miles, and more than a few ft/s would typify current specifications for velocity accuracy. Much classical inertial methodology was developed with error volumes far greater than what GNSS offers and, with updates from sensors of pre-GNSS vintage, instantaneous velocity was not nearly as accurate as *1* cm/s. Phase coherence over at least several seconds is the *sine qua non* for that capability – and with inertial instrument errors constantly active, repetitive carrier phase updates are essential for continuous maintenance of accuracies at that level.

To reiterate, it is widely known that ***inertial instruments can maintain carrier-enabled accuracies over short durations only. That recognition can be accepted as a fundamental influence to guide methodology for years to come***.

5.7 SUMMARY OF SYSTEM APPROACH

Except for details involved in preparation and validation of incoming data – to be discussed in the next two chapters – the GNSS/INS methodology and rationale have now been covered. What has unfolded incrementally from material presented can now be revisited, for a unified description addressing various interrelationships among different elements of the general plan.

The overall process includes several departures from customary GNSS/INS mechanizations. Many of these changes stem from algorithmic reductions, but the most important departures were driven by realism (presence of challenging gyro and accelerometer degradations with motion sensitivities described in Addendum 4.B) and motivated by performance. The approach, to be sure, contains elements that are well-known – tightly coupled integration; differencing (across SVs, time, and – optionally – receivers); velocity via integrated doppler. Even the occasional previous usage of observables proportional to time integrals of the state is noted – but not the time integral of the lowest-order state. Many facets of the approach herein differ from custom in ways not always obvious without the detailed descriptions given in the preceding material.

Deviation from usual methodology began with adaptation of adjusted sequential differences from [9,10] to another purpose – with elimination of integer ambiguity resolution – to preclude all of the following potential problems:

- the burden of resolving and maintaining precise cycle counts (without which position estimates are ambiguous; in error by an unknown integral number – potentially a very large number – of wavelengths)
- risk of temporary instability via false indication of successful ambiguity resolution
- development delays arising from added software provisions that could prolong system evolution and validation
- delays that can disrupt real-time operation (*e.g.*, waiting for geometry changes needed in acquisition, plus shorter pauses for mini-search in reacquisition)

The system *allows* operation with unambiguous carrier phase but does not *require* ambiguities to be resolved. The resulting configuration provides separate estimation of position from dynamics (Section 5.5), enabling update for the latter whenever any SV is observed twice in close succession. That circumvents the last delay just itemized and also contrasts markedly *vs*. the delayed pull-in experienced with CSC. In addition to removing both those delays and all problems associated with integer ambiguity resolution, immediate successive-pair-update capability eases requirements for phase continuity throughout. Formulations are provided for operation with and without ground stations.

Accurate performance demands rigorous accounting for sequential phase change sensitivity to velocity error states that evolve continuously throughout a differencing interval. That demand is met by the following segmentation features (most of which are quite easy to accept, since they enhance both performance and ease of implementation):

- All sequential difference (Section 5.4) adjustments are made with high accuracy. This includes SV motion within the subtraction interval (Addendum 5.D), lever arm dependency represented with unusual thoroughness, and rotation of the nav frame where applicable (*i.e.*, where a ground station was employed).
- Instantaneous velocity error state at the time of each measurement is premultiplied by the transition matrix during the current differencing interval, thus allowing the state to be taken out of the integrand.
- A time integral of the transition matrix over the differencing interval, with a *future* reference time (*i.e.*, the next update) – not yet transpired and thus containing elements thus far unknown – is required. The next item satisfies that need.
- The transition matrix is factored into two components. The premultiplier allows numerical integration to proceed forward in time; the postmultiplier, taken out of the integrand, is formed by easy inversion at the end of each differencing interval.
- Prescribed computations are carried out analytically to the maximum extent possible, followed by programming the final result.
- Options are made available to omit marginally observable states without affecting any other modeling properties, and information can be gleaned from misorientation pseudomeasurements (especially in azimuth; Section 4.6).
- Adaptive variance increments (Section 4.5.1) counteract model misrepresentation that could otherwise be introduced from gyro cross-axis and scaling errors.
- Propagation timing offsets are minimized by differencing across time, and optional corrections are supplied (Addendum 5.A).
- Programming is facilitated considerably by usage of triangular matrices – made possible by decisions affecting the formulation. This includes Bierman's upper triangular forms [4] – enabled by short-term characterization (Chapters 3, 4) – and also a lower triangular whitening matrix with direct algorithmic formation of its inverse square root, originating with this author.
- User clock error effects are canceled by across-SV differencing. Resulting correlations are taken into account via whitening as just mentioned (Section 5.2).
- Sequentially correlated measurement errors, introduced by sequential differencing in the dynamics segment, are shown to be inconsequential with present (and even foreseeable future) state-of-the-art for inertial instruments. Section 5.6 presents supporting analysis, with extensive confirmation from test results in Chapter 8 and from other references (many of which were examined for relevance to this operation in Addendum 5.B – containing a parallel evaluation with block estimation).

The last item, successful test, offers powerful validation of the overall approach.

Cascading transition matrices for precise integration over intermeasurement intervals imposed bookkeeping chores. Any inconvenience from those chores was far outweighed by benefits in performance, such as suppression of noise from specific force components – which appear in transition matrix elements. Other features listed simplified the computational burden as already noted.

There is no better way to end this summary than to recall the form for data inputs defined at the end of the first page of this chapter.

Additional Discussion

The last five pages (83-87) just shown are preceded, in the book, by full exposition of the segmented scheme wherein position (fed by streaming velocity with adjustment from pseudoranges as already described) is separated from dynamics estimation. The detailed formulation needed for independent processing of each individual 1-sec carrier phase change is first given. That is followed by rigorous analysis clearly identified as optional. Readers just wanting to process what works, irrespective of why, can omit the deep topic of investigating sequential correlation effects. Even for them, a simple mental exercise offers explanation for what follows: Envision a free-inertial coast starting from perfect initialization (*i.e.*, with zero error in every component of position, velocity, leveling, drift, etc.) – how long would it take for velocity error to reach 1-cm/sec? *Answer*: appreciably less than a minute; clearly we are in the short-term realm *for estimation models*; classical nmi/hr concepts are irrelevant for carrier phase processing.

Those wishing to explore correlations in carrier phase sequential changes will find full analytical support for the major issue: At 1-sec, the *sequential* correlations (not to be confused with correlations due to differencing across SVs or receivers) are ineffective; reasons stem directly from the short-term condition noted in the last paragraph.

To illustrate performance, the table below was taken from the vast data block of flight history (about 1-hr). Columns left-to-right expressed in meters represent sequential changes in across-SV difference phase, along-range position for each satellite and the subtraction-reference satellite, the effect of transforming asynchronous (1-sec apart) data into a unified coordinate frame, the effect of vehicle (in this case, aircraft) motion over the 1-sec interval, and finally carrier phase residuals:

$\delta\nabla Y$	a	b	$\mathbf{h} \bullet \omega_s \times \mathbf{R}_{m-1}$	$\mathbf{h} \int \mathbf{V}\, d\tau$	Residual
-359.71	818.26	-245.14	-174.79	-38.63	-0.01
-169.81	57.75	-245.14	303.22	53.97	-0.01
-31.75	402.64	-245.14	- 110.76	-14.99	0.00
416.93	-309.48	-245.14	120.14	17.55	-0.01
-271.26	651.70	-245.14	-116.03	-19.27	0.00
74.17	357.41	-245.14	-160.37	-26.07	-0.01

CHAPTER 6

INTEGRITY TESTS FOR DATA EDIT

Perspective for this chapter is slanted heavily toward usage of observable data to update either a track file or an inertial system. The make-or-break role of observables, self-evident in the former, applies to the latter as well; the INS performs maintenance – not original generation – of position and velocity. INS outputs provide velocity history only after the constants of integration are initialized or, after an extended data lapse, reinitialized. Stressing the obvious in that way calls attention to the critical importance of rejecting flawed observations, and also clarifies the focus here. Whereas GNSS performance assessments have customarily centered on signal veracity, experience shows that the user segment warrants far greater scrutiny [1].

Failures attributed in [1] to *previously validated* receivers vastly outpaced allowable space and control segment flaws. A one-to-one extrapolation from those results (beyond allowable levels by *four orders of magnitude*) to everyday operation was not claimed but, by any reasoning, the lesson is inescapable. Designers wishing to provide GNSS-rated reliability will need precisely time-stamped carrier and pseudorange data to perform integrity tests outside the receiver. That conclusion does not imply inability of receivers to perform their required function; it merely reflects what that required function should (and should not) be – according to authoritative users [2,3]. Sentiments expressed in [2],

> "⋯ *A proliferation of GPS receivers* ⋯ *has resulted in a plethora of data formats and interfaces that have caused confusion not only to the non-technical user but to the sophisticated user* ⋯ *OEM vendors do not provide all the information*"

followed by a recommendation proposing that

> "*OEMs focus on receiver design and not burden themselves with the ever-burgeoning variety of applications*."

are echoed in [3],

> "⋯ *products are not functional* ⋯ *Interfacing considerations are extremely important* ⋯ *Customers even in the same industry have different needs* ⋯ *solid courses of action slow in coming* ⋯ *Obviously this causes incompatibility of systems* ⋯ *The problem is echoed throughout* ⋯ *an interfacing specification called the Universal Common Data Points* [is] *neither universal nor common* ⋯ *ability to make these rapid adjustments has its roots in an open architecture philosophy. This appears to be an obviously simple concept but one that has eluded designers of other GPS-based systems*."

with sharp criticism for lack of user-friendly features voiced in [4],

"*If you're making maps, traveling, geocaching or otherwise using a GPS* ··· *don't let the crazy array of GPS data formats get you lost* ··· *command-line tool* ··· *converts what you have to what you need* ··· "

Issues just raised were expressed for inertial as well as receiver data in [5], a decade and a half old at the time of this writing:

" *information available from extant sources is deficient in content, form, timeliness, or precision. Unfortunately this is a common occurrence, not an occasional oversight; information is typically conveyed in ways that became standard long before modernization* ··· *Prime examples are attitude (expressed in terms of the familiar roll-pitch-heading convention) and velocity components in single precision* ··· *These and other instances of accepted procedures* ··· *impose fundamental but completely unnecessary limitations on achievable integration performance.* ··· *available means of correcting all deficiencies, which have been widely known for years, straightforward measures are proposed whereby standards can be updated*."

A "raw-data-across-the-board" approach, (first page of the preceding chapter) is unhesitatingly advocated here. Changes identified in [2-5] are slowly emerging, due to growing realization that the high rated performance can otherwise be unfulfilled. While no system can be perfect, the GPS performance track record is encouraging [6]; GPS is performing as advertised. To capitalize fully on that capability, designers can take charge by applying methods presented in this chapter to raw GNSS data.

What is needed from integrity testing is the most intelligent data editing possible for incoming observations. Recognizing the inevitability of two error types (rejection of some valid data, acceptance of some invalid data), the designer can err on the safe side, sacrificing some availability to minimize missed detections. Decisions can be quantified for given a probability distribution – but the distribution is not precisely known. Conservative characterization provides adequate protection with overbounding. As with many real-world designs, performance is not quite optimum but *lagom* (fully acceptable) – and superior to that of a pessimistic model.

Priorities just expressed place emphasis on technical rather than institutional considerations. The classical definition of integrity, involving requirements for timely alerts and alarms when specified error limits are exceeded, is obviously important for air traffic control [7]. Those requirements have changed over time (and could change again); furthermore, methods now to be described are intended for general application (*e.g.*, including military systems with different institutional needs). Material that follows will therefore steer primarily toward *means* of satisfying requirements, and allude in tentative ways to specific numerical values.

Alternative interpretation of integrity is not unprecedented; even a modified *definition* was offered in [8] (which noted that a logical performance metric is not just detection probability but that value multiplied by likelihood of occurrence for failures that necessitate detection). This chapter thus begins by introducing familiar analytical concepts to be used and – for multiple reasons (inexorable advances in state of the art, SA removal, integrated system needs) – promptly proceeds to less familiar methods.

6.3 EXTENSION TO ASYNCHRONOUS DATA

A prescient observation in [19] noted that, despite valid reasons for widespread usage of snapshot methods (*i.e.*, with focus on synchronized data), integrity decisions can benefit substantially from history of parity information. About a decade later a related concept – usage of covariances and past as well as present observations for current integrity evaluation – experienced renewed interest within the industry. Full adoption of those practices here can include histories of bias estimates from Sections 6.1/6.2 and expansion of the latter concept (exploitation of covariances and past observations) to the limit.

A most thorough integrity probe could call into question not only the state but every parameter or condition used in its formation. Rather than attempting that it has (understandably) become customary to believe validity of the following items:

- initial error state $\mathbf{x}_0$ and covariance matrix $\mathbf{P}_0$
- dynamic model process noise spectral densities
- the time tag, variance, and sensitivity applied to each measurement.

These conditions will be adopted for this development.

6.3.1 Approach

Data validity in the present context implies not perfection but reasonableness within statistical bounds. Actual initial error is thus regarded as a plausible member of an ensemble characterized with zero mean and values inserted into the covariance matrix $\mathbf{P}_0$ used to start the process. Likewise, any imperfections in parameter values used to characterize measurements and dynamic modeling may have ramifications not to exceed known effects of noise statistics already taken into account. Often this latter requirement is satisfied by using conservative values, again sacrificing optimality for the essential virtue of stability.

Validity of initialization by these concepts is quite common in estimation – and is not too much to ask for in a GNSS/INS integrity scheme. Inertial navigation, as a dead reckoning process, requires "constants of integration" for position and velocity. Few would oppose the need for an accessible beginning (*e.g.*, nominally stationary at takeoff, with location, heading, and attitude approximated to within accuracies again in compliance with $\mathbf{P}_0$). Moreover, conservative modeling and successful integrity will *maintain* conformance of estimates to $\mathbf{P}$ throughout – except *very* occasionally, upon discovery that corrupt data had degraded the estimation process beyond limits set by the statistics being used. For *re*initialization, some combination of the following steps can be prescribed:

- increment selected diagonal elements of the current covariance $\mathbf{P}$ or, in the limiting case, reset it to a diagonal matrix with large elements. The diagonal condition is known to give conservative properties [21];
- return to a location and/or a set of conditions conducive to valid initialization;
- use backup information to help reinitialize, often at lower performance capability.

6.3.4 Parameter Setting And Resetting

This overall section (6.3) aims at a conservative modeling strategy to protect against invalid data. At the same time, excessive conservatism would waste availability. The fundamental purpose – preventing undue error levels with low probabilities of alarm P_A and missed detection P_{MD} – is appraised on the basis of minimum detectable bias. Example values for the ratio $b_D/\sqrt{<z^2>}$ are tabulated below:

Table 6.1
Normalized detectable bias levels

$P_A \backslash P_{MD}$ →	0.001	0.0001	0.00001	0.000001
0.001	6.3808	7.0095	7.5554	8.0440
0.0001	6.9808	7.6096	8.1555	8.6440
0.00001	7.5074	8.1362	8.6821	9.1706
0.000001	7.9819	8.6107	9.1565	9.6451

Tabulated values conform to (6.14) but with no dependence on geometry of other measurements – "geometry" of the identity partition in (6.35) produces no amplification $1/|q_j|$ for individual observables separately processed. Overbounding in (6.45) can be conservative without being excessively so – that is known from experience with real data [23] and from analytical verification: As a design condition, (2.65) sets a balance between

- measurement variance and
- cumulative impact on the residual from process noise *amassed over a Kalman filter data window*, *many times greater than* the intermeasurement interval.

Dynamic process noise effect accumulated between measurement events therefore cannot overshadow $<z^2>$.

Because estimation schemes rely heavily on model adherence, vigilant residual monitoring is essential. A residual at *1.5* times its RMS value, for example, would not be rejected – but a persistent pattern at that level or a significant nonzero residual average would undermine validity. Fortunately there are applicable statistical tests (Addendum 6.A provides an introductory description, citing references for further tests prescribing added computations in parallel with estimation). Outcomes of tests can promptly signal a need for steps such as

- amplifying the RMS observation error setting
- adopting a higher elevation mask
- reconsidering propagation compensation (Addendum 5.A) if omitted
- soft reset: enlarging the diagonal elements of **P**
- hard reset; replacing **P** by a diagonal matrix with large element values.

In addition to residual patterns, there are of course the usual clues from SV health, signal quality (spectral purity), etc. Also, as advocated in [15,19], histories of integrity test results offer a wealth of decision aids; Sections 6.2.2 and 6.2.3, in fact, include direct estimates of each measurement's offset. As applied here, that offers independent estimation of each pseudorange's bias and each carrier phase drift rate.

6.4 WIDER OVERALL PERSPECTIVE FOR INTEGRITY

Material in this chapter, as noted at the beginning, is presented from a viewpoint driven by specific aims – aided navigation and tracking. The full realm of integrity development – its complete relation to accuracy, availability, and continuity – is much broader. Following is a partial listing of added considerations: Institutional needs; precise definitions of terms (*e.g.*, isolation *vs* exclusion) with all nuances resolved (*e.g.*, distinction between incorrect exclusion and wrong exclusion); specific numerical values {*e.g.*, probabilities, horizontal and vertical alert limits, protection levels (with reduction by biased estimation [24]) } formally stated in specifications; Markov state diagrams; failure modes in risk analysis – with ramifications involving control and space segments as well as implications involving the user segment; demands imposed by confidence coefficients at high levels; gaussian distribution traits in low-probability regions; test requirements. A vast array of literature is published on these issues; adequate coverage of all would require a separate book. Many are discussed in [19] and [25] – though clearly the last word has not yet been spoken. As one example of that assessment, consider the last item (test requirements), and how they are affected – in some cases adversely and in other cases positively – by the items preceding it (gaussian distribution traits and confidence coefficients):

- In operation the ratio $|\mathbf{q}^T\mathbf{z}/\sigma|$ tested *vs* T_D is evaluated on the basis of a *perceived* value for σ. When a true RMS error exceeds it by just *10%*, the actual alarm rate can grow by an inordinate amount (*e.g.*, almost an order of magnitude [26]). On the other hand, the next two items identify benefits of "gaussian tails" –
- The characteristic just mentioned facilitates probability scaling, wherein a modest increase of input noise enables observation of multiple events – affecting test confidence – from practical sizes for run trial count [27].
- In another example from [26] a bias enlargement of less than *50%* reduces the missed detection probability by six orders of magnitude.
- Without the probability scaling via raised noise level, confidence would impose demands for excessive test trial runs. The requirements are not at all intuitive [28].

Another option that could enhance conclusiveness of integrity testing would be to abandon a GO/NO-GO characterization of results [29]. Consider a test with a maximum allowable number N of missed detections – irrespective of whether each may be a near-miss or a *blunder* – with the following hypothetical outcome from a large number of trial runs for two receivers:

- RCVR #1 produces N missed detections, each occurring with errors exceeding allowable levels by orders of magnitude. Decision : *Accept*
- RCVR #2 produces $N+1$ missed detections, each occurring with errors exceeding allowable levels only slightly. Decision : *Reject*

This simple example illustrates the point: conclusive test requirements for general application need added development. Further considerations are described in [30].

Unfinished issues just identified do not affect the narrower scope of data editing for aided navigation and tracking.

References

[1] Nisner, P.D. and Johannessen, R., "Ten million data points from TSO-approved GPS receivers: results of analysis and applications to design and use in aviation," *ION Journal*, Spring 2000, *pp*. 43-50.
[2] Babu, K., Brooks, W.R., Kovach, J.P. and Robbins, G.M., "Recommendations for a generic receiver interface," Best-Paper-Award-winning manuscript, ION GPS-93.
[3] Stoddart, L., "Open architecture, a new concept in vehicle location systems," ION GPS-94.
[4] Marti, D., "GPS babel," *Linux Journal*, June 2002, *p*.22.
[5] Farrell, J.L. and van Graas, F., "That all-important interface," ION GPS-90.
[6] van Dyke, K., et.al, "Status update on GPS integrity failure modes and effects analysis," ION Nat'l Tech Meeting, 2004.
[7] *U.S.DOD/DOT Federal Radionavigation Plan*, Washington D.C., 2005.
[8] Kovach, K. and Conley, R., "SATZAP: a novel approach to GPS integrity," *ION Journal*, Summer 1991, *pp*. 163-190.
[9] Parkinson, B. and Axelrad, P., "Autonomous GPS integrity monitoring using the pseudorange residual," *ION Journal*, Summer 1988, *pp* 255-271.
[10] Kalafus, R. and Chin, G.Y., "Performance measures of receiver-autonomous GPS integrity monitoring," ION Nat'l Tech Meeting, 1988.
[11] Lee, Y., "Analysis of range and positon comparison methods as a means to provide integrity in the user receiver," ION 42nd Annual Meeting, 1986.
[12] Brown, R.G., "GPS RAIM: calculation of thresholds and protection radius using chi-square methods – a geometric approach ," RTCA paper 491-94/SC159-584, RTCA, Washington D.C., 1994.
[13] Brenner, M., "Implementation of a RAIM monitor in a GPS receiver and an integrated GPS/IRS," ION GPS-90.
[14] van Graas, F., "In-flight demonstration of hybrid GPS / Loran," ION Nat'l Tech Meeting, 1991.
[15] Farrell, J.L., "Extended RAIM (ERAIM): estimation of SV offset," ION-GPS 92.
[16] Farrell, J.L., "GDOP and RAIM in differential operation," *ION Journal*, Fall 2001, *pp* 195-203.
[17] Farrell, J.L., "Full integrity testing for GPS/INS," *ION Journal*, Spring 2006, *pp* 33-40.
[18] Farrell, J.L., *Integrated Aircraft Navigation*, Academic Press, 1976. (Now in paperback)
[19] Kline, P., "Fault detection and isolation for the integrated navigation systems using the Global Positioning System " (Master's Thesis), Ohio University, Nov. 1991.
[20] Young, R.S and McGraw, G.A., "Fault detection and exclusion using normalized solution separation and residual monitoring methods," *ION Journal*, Fall 2003. *pp* 151-169.
[21] Schlegel, L.B., "Covariance matrix approximation," *AIAA Journal*, Nov. 1963., *pp* 2672-2673.
[22] Bierman, G.J., `*Factorization Methods for Discrete Sequential Estimation*, Academic Press, 1977.
[23] van Graas, F. and Farrell, J.L., "GPS/INS – a very different way," ION 57th Annual Meeting, 2001.
[24] Ober, P.B., "Towards High Integrity Positioning," ION-GPS 99.
[25] Lee, Y., et. al., "Summary of RTCA SC-159 GPS integrity working group activities," *ION Journal*, Fall 1996, *pp* 307-338.
[26] van Graas, F. and Farrell, J.L., "Receiver autonomous integrity monitoring (RAIM): techniques, performance, and potential," ION *47th* Annual Meeting, 1991.
[27] Farrell, J.L. and van Graas, F., "Integrity testing for GNSS sole means," ION Nat'l Tech Meeting, 1994.
[28] Farrell, J.L., and van Graas, F., "Statistical validation for GPS integrity test," *ION Redbook vol V*, 1998. *pp* 89-100.
[29] Farrell, J.L., "A quantitative measure of integrity error," IEEE PLANS, 1994.
[30] Farrell, J.L., "Unfinished business: glaring absences from the achievement list," IEEE PLANS, 2004.
[31] Sturza, M. A., "Navigation system integrity monitoring using redundant measurements," *ION Journal*, Winter 1988-89, *pp* 483-501.
[32] Teunissen, P. J., "An integrity and quality control procedure for use in multi sensor integration," ION GPS-90.

CHAPTER 7

PREPROCESSING AND REPROCESSING

This chapter is concerned with modification of data as obtained from available equipment. Some limited attention is given also to processing within the equipment but, as in the other chapters, heavy reliance is placed on external references for highly specialized mechanization concepts. Since the entire focus has been on short-term applications throughout, emphasis here falls primarily on operations with limited durations for data-averaging. As a result, a preprocessing task of major importance elsewhere – coning – is addressed here only briefly, only for completeness, and only in an Addendum to the chapter. Successful test results obtained with the algorithms in this book did not use any coning corrections; designers with updates consistently available might skip Addendum 7.A.

7.1 IMU PREPROCESSING

Preferably the gyro and accelerometer outputs are digitized as received from the IMU. If not, data acquisition must be performed by the user. An excellent case study, used for the generation of successful test results, is given in [1]. Section 7.1.1 gives a brief summary of that approach.

The preferred IMU sampling control employs an external trigger, based on a GNSS 1-pulse-per-second (1pps) output. Without that capability the digitized IMU outputs as received are not in step with a GNSS time base. The 1pps output is then used for computational synchronization of IMU outputs with the GNSS data. An easy method is given in Section 7.1.2.

A system with analog outputs needs digitization (Section 7.1.1); a system with digital outputs synchronized without GNSS needs to be resynchronized via Section 7.1.2; no system should need both.

7.1.1 Data Acquisition

If IMU outputs appear in analog form, digitization is required. The procedure described next is to be performed for both gyros and accelerometers. The latter will be used for illustration.

7.1.2.6 Results

IMU measurements at *100* Hz from van tests described in [3] were used as inputs for testing the approach just described. The input data stream for one typical time trace appears in the center plot of Figure 7.2. Outputs are shown above and below that input trace, for higher and lower data rates, respectively. To show the full capability, an abscissa calibrated in time would be necessary. Instead, for expediency these outputs were plotted *vs* integer sample counts. As a result, the plots contain small timing misalignments (output data rates tested were chosen as prime numbers, not commensurable with *100* Hz). Performance is therefore actually better than the appearance indicates. In any case the interpolation clearly functions as intended; the approach has straightforward mechanization with –

- normalization by the input sample period producing a pair of arrays (integer and fraction) having the requisite number of elements, controlling the generation of all outputs within the current 1-second period
- an outer loop controlling all 1-second intervals,

providing all characteristics desired –

- GNSS synchronization reference
- output timing always kept current
- minimal latency; outputs lag real time by no more than two input sample periods.

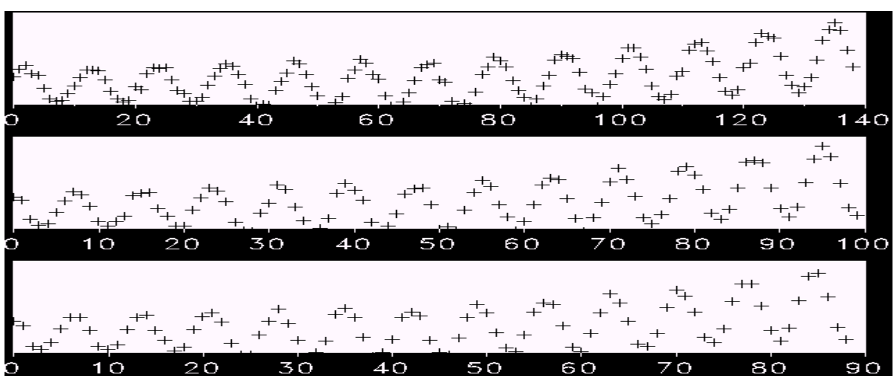

Figure. 7.2 139-Hz output, 100-Hz input, and 89-Hz output

7.2 GNSS REPROCESSING

Reexpression of GNSS data can involve

- relocation in time,
- relocation in space (lever arm adjustment), and
- observations using data from different constellations (*e.g.*, GPS + GALILEO)

All of these functions will now be described, in the order just listed.

CHAPTER 8

EXPERIMENTAL RESULTS

Material and methods from earlier chapters were used to obtain successful results, from raw GPS data alone and also from raw GPS data integrated with *100*-Hz data records containing outputs from raw gyro and accelerometer triads. The first case (*sans*-IMU) subdivides into two instances, *i.e.*, simulation and flight test. The simulated flight is described briefly, for the purpose of highlighting

- success achievable with sparse GNSS data availability, and
- specific benefits given by GNSS-independent dynamic data when maneuvers occur.

Still without IMU data, an actual 7-min flight leg with a nominal 90^0 turn and *50 %* speed change is described, wherein standalone *1*-Hz GPS updates produced *1*-m RMS position errors (from pseudoranges) and RMS velocity errors of order *1* decimeter/s (from sequential changes in carrier phase). The latter, crude by GNSS standards, came from constant-acceleration modeling in the presence of severe DC-3 vibrations.

Following those results presented in Section 8.1, Sections 8.2 (van) and 8.3 (flight) show extensive test data obtained from integration with the *100*-Hz IMU data records. Time histories are given for

- North-*vs*-East plan view,
- speed and altitude,
- attitude (roll, pitch, and sub-mrad leveling corrections) with azimuth shown in terms of the drift angle (deviation of velocity direction from heading, mainly due to wind),
- pseudorange residuals indicative of *1* to *2* m RMS position accuracies, and
- carrier phase residuals indicative of cm/s RMS velocity performance.

Plots involving dynamics and carrier phase residuals are the most revealing; position accuracy alone is not a reliable indicator of high achievement. That point is discussed further in Section 8.3.4, following presentation of flight test plots.

Test results shown were obtained with tight coupling; the receiver interface provided access to raw GPS measurement information – but not to the track loops. Neither ultratight coupling nor deep integration could be performed with the equipment used. The limitation was not serious for these tests, which produced enough SV data throughout. When a track loop was unsteady, the data from that SV could simply be omitted. If unsteadiness had existed in enough cases to hinder observability, however, a higher degree of integration would have been necessary.

8.1.2.2 Dynamic Sequence

Flight conditions are described by:

- high-speed motion
- presence of irregular motions containing high-frequency spectral components, and
- acceleration estimates representing averages over duration denoted as T.

The last condition produces the simple transition matrix used in (2.12), which merely formalizes mathematically the simple dynamics used in extrapolation. Results to be presented here were generated for *1*-second steps ($\tau = 1$, producing a sequence of overlapping blocks of size T / τ), with terminal conditions of each *1*-s interval initializing the next.

With a geographic (North/East/down; NED) frame used for navigation, the following quantities are formed by methods widely known (therefore needing only cursory descriptions here):

- average velocity during the interval, simply by summing its initial value with half of (acceleration vector) × (time increment)
- standard ellipsoid curvature radii, as weak functions of geodetic latitude (3.7,8)
- angular rate of NED frame relative to Earth {ratios of velocity components to appropriate curvature radii summed with altitude, with applicable trigonometric scaling; Eq. (3-24) of [4], excluding the sidereal rate terms}
- step increments in latitude and longitude from those relative angular rate components {Eq. (3-24) of [4]; this and the preceding step, though less general than the wander azimuth approach in Chapter 3, are stable at moderate latitudes}
- step altitude increment from (vertical velocity component) × (time increment)
- new ECEF position vector, using (5.6) with values just formed for curvature radii, latitude, longitude, and altitude
- transformation between Earth and NED frames, from sines and cosines of latitude and longitude { postmultiplying matrix on the right of (3.2) }
- optional rotation of the acceleration vector, through an angular rate [5] equal to the vector cross product,

$$\text{angular rate of acceleration vector} = \text{velocity} \times \text{acceleration} \div |\,\text{velocity}\,|^2 \qquad (8.1)$$

This last option, potentially beneficial for longer data window durations, is made available to vector extrapolation only, and omitted from extrapolation of covariance matrices. The omission is consistent with the common practice of allowing modeling imperfections in covariance dynamics. That is justified in two different ways, *i.e.*,

- <u>by results</u> – successful performance presented herein
- <u>by analytical reasoning</u> – simplified covariances produce a small gain variation which then becomes multiplied by residuals – producing a second-order effect, smaller than measurement noise (recall the closing discussion of Chapter 2 and Addendum 5.C). Even that small effect propagates into subsequent residuals, which are then taken into account by normal operation of the estimation process.

8.1.2.3 Flight Segment

Figures 8.4 and 8.5 show a *7*-min path containing a nominal right-angle level turn. Altitude variation was modest, but speed increased by about *30%* .

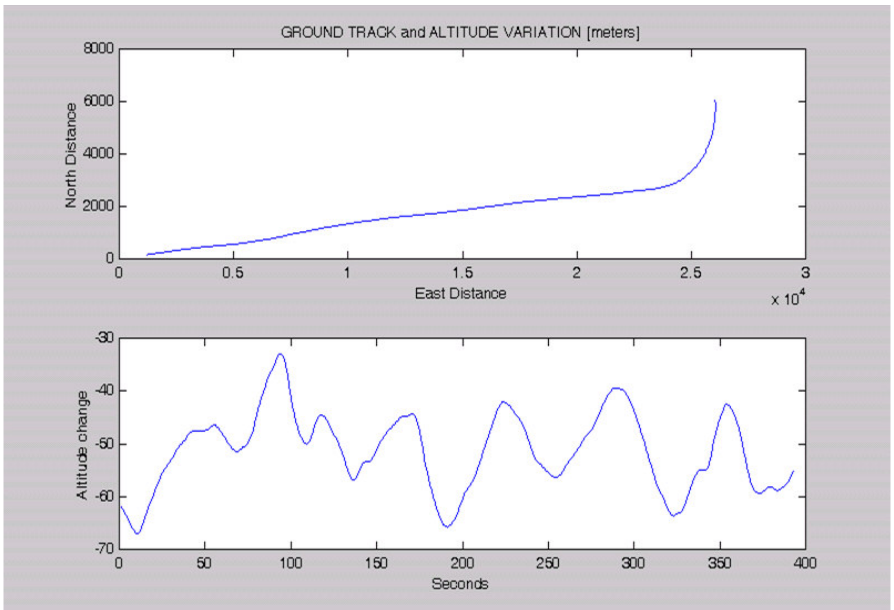

Figure. 8.4 Position history over 7-min flight segment

The comments below precede illustration of velocity and performance histories, in order to enhance their interpretation :

- The bank angle changes from nominal zero to about *13* ^{0} in roughly *20* seconds. During that time the roll rate is little over *0.5* ^{0}/s. The component of lever arm normal to the roll rate vector is about *2* meters long. The product *0.01* rad/s • *2* m, which is then scaled by SV sightline unit vector components, produces maximum products well below the decimeter residual levels obtained throughout this exercise.
- After the roll angle reaches *13* ^{0} the heading rate (therefore the approximate yaw rate at *13* ^{0} roll) is about *1.7* ^{0}/s. The component of lever arm perpendicular to the yaw rate vector is about *4* m. The product (*1.7* / *57.3*) rad/s • *4* m, likewise scaled by SV sightline components, still produces products typically below – but now comparable to – the decimeter residual levels. Since the lever arm is taken into account, that does not explain the performance degradation (higher pseudorange residuals) visible in the turn, especially for the longer (*8*-s) dynamics data block. The next section explains: a quasistatic acceleration model is suitable before or during the turn – but not for transition *into* the turn.

Success of the approach is evident from Figures. 8.6 and 8.7. Phase residuals have a maximum value of essentially *50* cm. The variance of those residuals obviously exceeds $\mathbf{HPH}^{\mathrm{T}}$ {one of two positive constituents); recall (6.45) or the denominator of (2.64) }. If the maximum value can be set to 3σ, then the nominal level of $\mathbf{HPH}^{\mathrm{T}}$ is less than ($50/3$) cm. The magnitude of $\mathbf{H}$ – a difference of two unit vectors – is generally between *1* and *2*. The sequential differences were formed from phase observations one second apart. RMS velocity error is therefore in the decimeter/second range – crude for GNSS but, for constant-acceleration modeling with DC-3 vibration, understandable (and acceptable for many operations).

Aside from a transient during roll change with the *8*-s dynamics data span (discussed immediately below), RMS pseudorange residuals are at or below a meter in Figures 8.6 and 8.7.

With data edit as derived in Chapter 6 activated, no measurements were rejected for the *3*-s data span. With the longer span, however, several of the carrier difference measurements are edited out. Focused subdivision within the flight segment showed that the vast majority of those edits occurred during the bank prior to turning. A transient occurs due to violation of the constant-acceleration model, when held for larger portions of that changing bank period.

For another procedural detail, the results shown were generated without activating rotation of the acceleration vector by (8.1). With this flight path its effect was minor (and in fact, disadvantageous in the case of the *3*-s data span which understandably produced noisy acceleration estimates).

8.1.2.5 Analytical Review And Interpretation

Focus now shifts to contrasting applications involving a standalone receiver without IMU aiding. Issues to be addressed can be clarified by starting the comparison *vs* a much simpler set of dynamic conditions. The limiting case might involve reducing the in-air dynamics just described, in three ways:

- from high-speed motion to stationarity
- from presence of irregular motions containing high-frequency spectral components to quiescence
- from inclusion of acceleration estimates – representing averages – modeled as constant over T to *velocity* estimates held constant over T.

An opportunity is offered in [6] to investigate – apart from distraction by the unknown irregular motions (high-vibration airborne environments) – achievable accuracy in carrier phase sequential difference measurements. They can be quite accurate (*e.g.*, down to the millimeter level in [6] at *50* Hz) with a sophisticated receiver implementation described in [7]. By enabling evaluation of low-order estimation with highly accurate receiver mechanization and the most basic case of dynamics (*i.e.*, no motion at all), that rooftop experiment provides a conceptual point of reference for performance. The remainder of this section will briefly consider proceeding from that point to operation under more general conditions.

The *1*-mm accuracy was applicable to excursions over $\tau = 0.02$ s. Further experimental results in [6] reasonably fit a theoretical phase increment variance formed as a product $\tau \times$ (spectral density) – *i.e.*, proportionality of *RMS* incremental excursion to the *square root of* τ (*e.g.*, adjusted for squaring loss in the *5*-Hz test, excursions over $\tau = 0.2$ s produced about $\sqrt{10}$ times as much error). Thus *velocity* error { incremental excursion error $\div \tau$ *for each individual observation interval* } is *inversely* proportional to $\sqrt{\tau}$ {note the square root of sampling frequency in Eqs. (1,2) of [6]}. Since there are $(T \div \tau)$ of those intervals in T seconds – and *average* error in estimating a constant over (T/τ) independent samples is inversely proportional to $\sqrt{T/\tau}$ – *average velocity error* over T seconds is then inversely proportional to the square root of the *averaging* interval T – for independent observations.

The caveat regarding independent observation errors will be revisited immediately after this summary for dependence of error on an excursion interval τ or an averaging interval T:

- Error in excursion during τ is directly proportional to $\sqrt{\tau}$.
- Error in velocity during τ is inversely proportional to $\sqrt{\tau}$.
- Error in average velocity over T, from T/τ independent samples of duration τ, is inversely proportional to $\sqrt{T}$.

This last item lacks the full theoretical performance benefit offered by the information content in repetitive samples within T. Errors in those samples are not independent; they are negatively correlated. The theoretical advantage gained by accounting for the negative correlations, noted long ago in [8], is addressed at length in Section 5.6 for GNSS applications. In brief that theoretical advantage is significant – and recoverable – only for a large ratio T/τ, and then only if in strict conformance to all model characterizations for periods of duration T. That is not the case for airborne data processed herein (acceleration vectors can reasonably be *averaged* over several seconds but, with the vibrations, they are not at all constant for any significant length of time). For that reason, validated in Section 5.6, negative sequential correlations between measurement errors were not included in results generated for presentation in Section 8.1.2.4.

The basis for the decision just stated is attributable to airborne operation. Although the data generated for that investigation came from a flight test (which naturally places emphasis on airborne environments for the main focus), some further insight can be gained first by considering other applications not subjected to irregular vibrations. Two cases satisfying that condition are

- the stationary experiment of [6] and
- a dynamic propagation model for orbit determination of spacecraft with no high internal dynamics (*e.g.*, no spinning rotors) and with damping of structural oscillations.

Either of these operations could *theoretically* benefit from accounting for the negative sequential correlations.

Pseudorange residuals are consistently well within single-digit values. In fact, all plots after takeoff are bounded by maximum absolute values below 2 m. With no ground station that accuracy is unusual; propagation was exceptionally well-behaved for this test. Again, representative plots (Figures 8.35 - 8.38) for the most dynamic flight segments suffice to validate performance.

In Figures 8.31 - 8.34, the onset of a roll initiates a turn, indicated by change in the drift angle (difference between heading and the horizontal projection of the velocity vector) as expected. As just mentioned that event is accompanied by a speed change (normally an increase unless turning into the wind), to maintain lift thereby holding altitude. Positive pitch angles of up to a few degrees consistently average to not much more than 1^0. Of primary interest in those plots, the leveling adjustments – plotted in mrad – hug the abscissa. Their maximum absolute values rarely exceed 1 mrad, consistent with state-of-the-art (few tenths mrad RMS) tilt performance.

Additional Discussion

Extensive plots of van and flight test results (absent from this limited-size file) are provided in the book . Page 176, preceded by plots of speed, altitude, attitude, and 0.3-mrad RMS leveling corrections, shows the pseudorange residual plots just described. Carrier phase residual scatter plots then follow. The first of those nominal 0.1-hr phase plots shows values almost entirely within ± 3 cm. The other phase plots show somewhat larger values, but still easily support the 1-cm/sec RMS velocity error assertion (statistical analysis is provided, accounting for residuals' inclusion of instantaneous measurement error in addition to effects of imperfect estimation). With data editing based on testing of residuals *vs* integrity criteria, the fraction of observations rejected was spectacularly low – and even those few were generally attributable to known causes (*e.g.*, low elevation angles for rejected pseudoranges, lever-arm adjustments affected by the initially large heading uncertainties during takeoff, etc.). Altitude history during takeoff has the expected appearance, holding steady at local terrain's altitude above sea level until the speed reaches about 70% of its final value – then climbing at near-uniform slope up to a fixed chosen value.

Plots are presented primarily for dynamics rather than position. It proves very little to demonstrate excellent location history from GPS measurements – especially on a scale spanning well over 10 km.

CHAPTER 9

AIDED TRACKING

Determining the path of an object with off-site computations (*i.e.*, computing at a site not onboard that object) can be described as aided tracking if

- measurements from sensors onboard the object or
- cueing observations (data from a "third party" – using sensors not onboard the object and not collocated with the site of the computations)

are used in the path estimation. The first case (labeled as "mutual surveillance" with multiple participants) is clearly cooperative; action occurring on the object helps to track it. The latter case is noncooperative in that sense, but the term "cooperative engagement" applies to action of two or more participants pooling observations of an unknown object. Both mutual surveillance and cooperative engagement are covered here; the latter, when encompassing both functions, is obviously more demanding.

Mutual surveillance is first described here for a scheme wherein each participant uses a Mode S squitter (Section 7.2.1) to transmit and receive raw GNSS observations. Cooperative engagement will then be described with participants

- employing mutual surveillance as just defined with all friendlies, while
- sending raw range and/or line-of-sight (LOS) observations of noncooperative tracked objects from radar and/or optics. In this chapter, "optics" is a general term enveloping infrared, TV sensing equipment, or any electro-optical device.

Exceptions that only partially satisfy these conditions need not be discussed here in any detail, since various modified or extended operations can easily be formed.

To avoid branching into a vast array of related subjects, several limitations in scope are imposed. Stealth and encryption are not addressed. Nor is a step beyond noncooperative – *un*cooperative, involving countermeasures to thwart tracking by emitting chaff, flares, jamming signals, etc. from the object. Communication is treated in an abstract sense only; provisions are taken for granted (within reason; extended squitter messages are used for illustration, and practical restrictions on capacity are given due consideration). As earlier chapters relied on bibliographical references for GNSS, radar or optical sensing and imaging technologies are likewise regarded as covered elsewhere. Aside from citing of a few classic references (*e.g.*, [1] for radar; [2] for image processing), efforts to substantiate applicable characterizations or statements are limited. Exceptions are made, however, for important capabilities that are obscure or subtle; specific references are added in those cases.

Focus in this chapter, as in previous ones, is functional form and achievable performance of estimation in meaningful dynamic scenarios with realistic properties of obtainable measurements. Undue preoccupation with mechanization is avoided. Still it is essential to maintain practicality – sensors used for tracking have intrinsic features and limitations. In some cases this presents no problem – *e.g.*, it is easily accepted that calibration minimizes a host of degradations despite usage of very imperfect hardware. Whether in regard to occasional sets of radome refraction measurements for correcting indicated radar LOS direction, repetitive test signal generation for radar or optical sensor self-calibration in real time, or dozens of other compensation schemes, the perspective adopted here is the same. *Effective* errors in sensor data are amounts remaining post-correction. Without delving into procedures, the errors can be taken into account via RMS values assigned to those uncompensated amounts. Approaches presented in this chapter are independent of implementation, as intended, and performance capability can grow with state-of-the-art.

Other distractions are less easily sidestepped. Radar "blind" spots, multiple sources and types of interference, angle-error-induced cross-range mislocation proportional to distance, plus various subtleties must be taken into account for credible evaluation. The approach taken in this chapter is to furnish a quick resolution of an issue as it arises, while relying on familiar references to provide any needed explanations. Digressions occur as needed but are kept brief.

There are additional operations not fully within either of the two main categories here, but related to one or the other. The tracked vehicle can carry a translator that rebroadcasts its GNSS signals for detection and tracking at another location. Another variation is bistatic or multistatic operation, with transmitters illuminating a tracked object for reception on a displaced platform (Addendum 9.A). Aided orbit determination, ballistic objects, littoral environment, and supporting operations are described in Addendum 9.B, 9.C, 9.D, and 9.E, respectively.

9.1 MUTUAL SURVEILLANCE

Over the past two decades, performance achieved through differential GPS (DGPS) has dramatically improved (and in some cases, revolutionized) many operations. Applications thus far have involved a stationary reference receiver, employed for the purpose of counteracting major error sources. Undeniably the success has been phenomenal. Nevertheless, the logistics involved with a stationary reference receiver can compromise practicality for some applications. That reason was given as an explanation for brevity of Section 5.3.

Certain operations, however, inherently depend on combining data from aboard two or more separate and independent platforms. In those cases, usage of comparative data cannot possibly be avoided (obviously) – but it can be optimized and minimized (apparently that is less obvious). This section addresses differential operation without dependence on a stationary participant.

9.2 COOPERATIVE ENGAGEMENT

Attention now turns to observations of multiple noncooperative objects, shared by participants. The aforementioned flexibility for different message types is useful here, but not to distinguish airborne from ground vehicles as in [3]. Instead, the *48* data bits of an extended squitter message can be allocated to either

- pseudorange differences (*e.g.*, in the *24*-bit pairs) for mutual surveillance , or
- cueing observations for cooperative engagement.

The latter can be LOS (azimuth and elevation) from a radar or optical sensor, or range (from a radar or laser ranging device). Radar doppler could also be included but, in many operations, it is important only for gate placement in preprocessing for target data extraction (spectral separation) – and less important as an observable, because

- radar echoes containing engine modulation can be mistaken for skin return,
- only one component of velocity (along-range) can be sensed, and
- after transients subside, range history alone suffices; doppler is superfluous.

For brevity, restrictions are placed on scenarios to be addressed. Objects to be tracked are either separately observed, or treated collectively as a cluster until they can be resolved. That condition enables correspondence between reports and track files (the *association problem* – with unknown correspondence – is beyond scope). Often this is acceptable for tracking of airborne objects, since

- separate extraction is not difficult at short range, and
- cluster tracking is usually adequate for multiple objects at long range.

Tracking ground objects can be far more challenging. A familiar example calling for sophisticated capability involves a swath of terrain with a few metallic objects plus several rocks (sensed by a radar) and live animals (sensed by an infrared device). Attempts to resolve unknown association between observations and track files could necessitate maintaining an unwieldy set of hypothetical track files – most of them extraneous. To preclude that, reports need to contain *only* responses from objects detected by both sensors. For vehicles moving too slow to be spectrally separated from radar ground return (clutter), that requires fusion of the radar and infrared responses *while still in raster form*. This is theoretically possible for a swath of terrain that is level, or sloped at a known grade; an abbreviated explanation follows.

Responses from terrain to transmissions of a coherent radar are separated into cells with boundaries formed by ground intersections with antenna-centered spheres (equal distance loci) and isodops (equal doppler loci). On a planar terrain swath these form radar image cells bounded by circles and conic sections (mainly hyperbolas). In size, shape, and orientation those cells differ greatly from pixels of optical images (generated from azimuth and elevation angle loci). Image registration achieves correct superposition of responses from different rasters, thereby enabling isolated extraction of doubly detected targets. Development that follows applies only to successfully isolated targets with high output signal-to-noise ratio. Cooperative engagement here provides no means of active assistance (*e.g.*, transponders), but no high-energy impediments to tracking remain after extraction either.

9.2.2.2 Fully Coupled Cartesian States

Simplifications used in Section 9.2.2.1 are quite useful for the case considered therein, which includes the following conditions:

- continuous or near-continuous availability of updating measurements in three axes,
- limited rotation of sensor-to-target sightline during a data window –

otherwise the broader formulation is generally needed. Recall the example in Section 9.2.1 with two aircraft approaching each other at Mach *1* . The familiar *9*-state representation – *without* decoupling on the basis of (9.5) – is fully adequate, for reasons previously explained: simple dynamics in a stable reference frame.

One and the same phenomenon – sensor-to-target sightline rotation – can be

- a detriment to track at very short ranges unless full 3-dimensional coupling is used
- a crucial ***benefit*** in operations with updates missing from one or two axes.

The latter issue is easily explained by considering the sensor's translational excursion occurring between the first and last updates in a data window with passive operation (angles only; no range measurements). From one (simplified) perspective, that change in position effectively provides crossing sightlines emanating from two different locations. From another viewpoint, cross-range information at one instant of time will "morph" into along-range information later. As long as that later time is within the data window, there is some observability along an axis not directly monitored. The *9*-state formulation is quite effective in many operations without full 3-axis sensing.

Immediately an important qualification must be stated. The benefit just identified, maximum for shortest distance, is limited to moderate ranges. At extreme ranges full coupling can actually become a liability for practical reasons. Useful guidelines were provided in [8,9] which contributed the following insights:

- When accuracies are highly mismatched across axes, the effect appears in the error ellipsoid shape. Visualize a "pancake" (good estimation accuracy along range but not cross-range) or a "cigar" (vice-versa).
- Transformation of mismatched covariances must be carefully performed, especially at track initiation. Mechanized algorithms have access to perceived, rather than actual, quantities to be used in formation of sensitivities and transformations.
- With slight discrepancies in perceived directions and grossly mismatched error ellipsoid dimensions at extreme ranges, (2.45) bestows unrealistic decrements in variances along unmonitored axes. Subsequent estimates are adversely affected.
- The decoupling offered in Section 9.2.2.1 can then be more than a convenience; at extreme ranges it becomes a necessity.

A counterpart to passive ranging (angles only) is range-only tracking. In that scheme, 3-dimensional observability requires sightline rotation in both azimuth and elevation. As an intermediate case, the Traffic Alert and Collision Avoidance System (TCAS) includes barometric altitude data transmission, so that sightline rotation in azimuth can provide observability – eventually (*i.e.*, after sufficient tracking duration at sufficiently short range). Both of those caveats are undesirable for collision avoidance; potential conflicts in TCAS today are resolved in the vertical plane.

CHAPTER 10

SUMMARY AND FUTURE PROSPECTS

> *The difficulty lies, not in the new ideas, but in escaping the old ones, which ramify, for those brought up as most of us have been, into every corner of our minds.*
>
> — *John Maynard Keynes*

Material presented herein thus far has offered advances in two main areas, *i.e.*,

- accessible means of implementing GNSS / INS integration, and
- opportunities to capitalize fully on GNSS information for situation awareness

with attention to underlying commonality of navigation and tracking. Much effort is aimed at thorough exploitation of modern capabilities (not only in navigation but in sensing, computing, and estimation as well) to enable a far better cost / benefit tradeoff than commonly realized. Opening discussion here expands on that last remark.

By its intrinsic nature a tradeoff implies inevitable sacrifice of one desirable trait as a price to be paid to gain another. Recent history of computing, however, prompts another viewpoint; lower cost has accompanied vastly improved capability. No attempt is made here to catalog the history of those accomplishments – except to note one feature: *change*. Another example, highly relevant here, involves GPS receiver evolution. With changes in methods and procedures, then, improved performance and economy are clearly not mutually exclusive.

The two main areas of advancement identified to open this chapter are now assessed for realization of twin goals, performance with economy, concurrently. There are low-cost sensors – both inertial (MEMS) and *rf* – with algorithms freely available to process their outputs (*e.g.*, those documented in this book produced the results shown in Chapter 8). Despite that, systems are expensive and slow to gain acceptance. Reasons for those cost and scheduling difficulties are grounded in yesteryear's capability limitations, no longer necessary. The real payoff of advanced techniques can come in the form of urgently needed solutions.

As long as improvement is being sought, it is worthwhile to broaden the goals of navigation (to know "where we are and where we're going") and of tracking (to know "where everything else is and is headed") much further. A list of objectives and means for realizing them – just by making full use of provisions already available, with no need for new inventions – is the next topic.

10.1 INTEGRATED SYSTEM GOALS

Among the more obviously desirable traits of navigation and tracking systems are economy, accuracy, reliability, thoroughness (*i.e.*, providing dynamic states as well as position), and robustness (*i.e.*, ability to perform "no matter what"). A more detailed listing of traits both general and simple (*i.e.*, comprehensive as well as comprehensible) appearing in [1] included the following:

- accuracy commensurate with the best constituent subsystem, continuously rather than just after that subsystem's observations
- generous sharing of data across functions (*e.g.*, usage of the same IMU data for nav plus control for guidance plus sensor stabilization)
- avoidance of duplication to provide size, weight, power, and reliability advantages
- system adaptability to modified parameters and/or conditions
- versatility, enabling flexible design extension to multipurpose applications
- circumvention of delays in development schedules
- maintainability with replacement of moderate size units rather than large assemblies
- testability; comprehensive assessment of equipment performance
- predictability of performance (*e.g.*, via rigorous emulation with insertion of simulated data into bench provisions) before equipment is fully available
- amenability to future growth, accommodating unforeseen added provisions
- extensive backup; graceful degradation taken to the limit (survival of Apollo XIII).

Many current self-contained systems provide the first of these benefits. An integrated system should provide all. Features forfeited in common mechanizations could have been retained by one change – the "raw-data-across-the-board" approach cited near the start of Chapters 5 and 6. A simple example will illustrate: consider the unforeseen addition of a hyperbolic navigation receiver for backup to an existing GNSS/INS design, after all mechanization decisions from Chapter 7 were made and all tests were successfully completed. Insertion of time differences to the nav system interface and limited modification in specific areas of software, including parameter changes plus the logic to control their values, would be straightforward – especially compared to the burden imposed with common interfaces in current use:

- a standard INS interface providing processed outputs, imprecisely timed relative to other equipment and often truncated to *16* bits, with velocity expressed in hybrid units (groundspeed in kts, ascent or descent rate in ft / min), position containing singularities at either pole, and attitude parametrized in a form (angles) also containing singularities and requiring reexpression prior to usage in algorithms
- latitude, longitude, altitude, and velocity vector from the GNSS receiver
- latitude and longitude (with or without horizontal velocity) from the added receiver.

Difference in the task corresponds to contrast in achievable performance, timeliness as well as accuracy. Introduction of custom interfaces has offered no relief, due to unnecessary complexity, added cost, and inflexibility. Utter simplicity afforded by raw data usage (time differences added to the satellite/inertial information defined on the first page of Chapter 5) has all of the advantages – and no disadvantages.

10.2 INTEGRATED SYSTEM USES

Need for the features just listed is evidenced by the variety of tasks that use navigation data in modern applications. Far beyond the need for simple display data that guided early requirements, IMU-derived information has for years supported myriad functions such as

- transfer alignment
- SAR motion compensation
- target designation
- fire control
- flight control
- pointing (*e.g.* laser)
- sensor stabilization
- image stabilization
- fusion of raster images
- antenna servo drive
- tracking (air-to-air, air-to-surface, surface-to-air, surface-to-surface)
- bi-static, multi-static, and passive operations
- surveillance in dense multi-target environments
- cueing, cooperative engagement
- cursor aiding
- GNSS carrier track aiding

Major cost and scheduling burdens involved with these functions arise due to procurement from isolated suppliers committed to using and providing data that can be compromised in multiple ways, *e.g.*,.

- <u>content</u> : In the augmented backup example of Section 10.1, consider a typical interface providing position and velocity pseudomeasurements from both receivers and from the INS. For decades it has been known [2] that "degradations of performance when the correlations in states are ignored is rather dramatic $\cdots$."
- <u>timeliness</u>: Both latency and time tag uncertainty degrade effectiveness of data.
- <u>precision</u>: *16*-bit word length cannot adequately support the functions just listed.
- <u>form</u> : Attitude information is generally needed in the form of direction cosines (computed using trigonometric functions) but supplied as angles (computed from inverses of those same functions)! More latency and imprecision result.

Recovery from nonessential limitations often involves hardware or wiring changes, with obvious impact on economy and delayed deployment. For a future system integrator to perform successfully, a complete interface definition must precede – never follow – design of individual modules or subsystems. That definition, furthermore, must require raw data – with timing, precision, and form needed by the software – availability for central processing. Note that this requirement does *not* preclude federated solutions. A properly configured layout will allow multiple computations, each including or excluding any input combination; every conceivable subset can be obtained for concurrent comparison with fully integrated solutions.

Huge improvements in both cost and performance for navigation are available to tracking as well – simply by adopting raw measurement data as the form for information transfer. This opportunity arises from precisely the same reason for the widely known success of differential operation. The wisdom shown in the 1980s by RTCM SC-104 [3] should be brought to bear on the broad scope of operations.

Even with no augmenting corrections, Section 9.1.1 lists eight major advantages over current practice (paths formed by stitching together coordinates of varying accuracy in different directions with unknown correlations). The obvious impact on velocity – with its direct influence on closest approach time and distance – will command greater attention as closer spacing is needed. Clearly that will be especially important for safety-critical applications, with total annual collision probability dependent on *365* × (*number of aircraft flying daily*).

For thoroughness the abovementioned correlation issue, cited long ago in [2] and confirmed by results reported elsewhere, is briefly revisited here. In principle, correlation data could be sent with coordinates and theoretically taken into account. However, several objections to that scheme arise immediately, *e.g.*,

- Extravagant communication capacity needed to support that step would be impractical, especially since the benefit sought can be obtained without expansion.
- Successive coordinate sets, often derived from different SV's, can impose burdensome tasks of maintaining patterns of correlation sequences.
- Adjustments (starting with timing and corrections but often proceeding further) used to generate outputs modify correlations in ways not always easy to trace. More preprocessing makes the unscrambling task more daunting. †

In the presence of these obstacles, common practice ignores temporal as well as spatial (across axes) correlations between pseudomeasurement errors. Contrast *vs* the proposed approach could hardly be more marked; estimates based on raw uncorrected pseudoranges account for these correlations ignored by conventional practice.

The seemingly unspectacular step of communicating via observables, then, combined with exploitation of modern computing capability, can have a sweeping impact on tracking. Participants broadcast observations instead of their positions. Potential applications noted in [4] included information distribution systems {both commercial – fleet monitoring, <u>V</u>essel <u>T</u>racking <u>S</u>ystems (VTS) – and military}, test range training flights, carrier landings, runway monitoring, collision avoidance, ⋯ . Moreover, as the preceding chapter illustrates, benefits can be extended to radar and optical observations shared by friendly participants monitoring the environment. Adopting the full range of information sharing would further bridge the gap between surveillance and navigation – a historical separation with roots that are institutional, prompted in part by earlier technological limitations, long since overcome.

---

† The above comment does not refer to signal processing performed on the high-volume data ahead of FFT or correlator input; for "raw data" here read "measurements" – again note the first pages of Chapters 5 and 6. To steer between extremes, information should be neither unwieldy nor excessively consolidated.

10.3 CHANGE WITH GROWTH

This book has offered formulations useful for navigation and tracking by the simplest means possible. For all operations detailed herein, linear or linearized estimation is adequate; extension would call for a few iterations at most. No intention is implied to preclude nonlinear estimation from an expanded scope of operations. Complete algorithms for terrain navigation, not addressed herein, would exemplify a need for extension to nonlinear estimation [5]. Rather than an attempt at restriction to linear methods, then, approaches used here indicate usage of sophistication needed – but no more. The door is open for additions when necessary.

A general approach to navigation and tracking must likewise allow room for prospective advances in methodology from other directions. For years the industry has seen efforts – not yet mature but indicative of future expanded usage – toward exploitation of information having less familiar forms. Combined progress toward accumulation of ”many small hints and clues” through means now known by names such as artificial intelligence, fuzzy logic, neural networks, data mining, etc., could eventually be combined, and possibly joined with other approaches (*e.g.*, fractals) to provide enlightenment not yet fully obvious for navigation. Again, what is not now included is not precluded. Just as strapdown IMU update via Kalman estimation – once considered highly advanced – is presented in ”cookbook” form here, the future will likely enable practical implementation of techniques now considered exotic.

No predictions are made here for advances just noted, or additional developments (*e.g.*, Open System Architecture, massively parallel Multiple Instruction Multiple Data “MIMD” processing with dynamic resource allocation, etc.) covered elsewhere. Of greater immediate concern is the need to capitalize on the main benefit offered by differential operation in the first place. The cornerstone of that procedure – usage of separate data from each individual satellite [3] – is readily extendible to relative navigation (Chapter 9) between vehicles moving in two dimensions – ships and driverless cars, or three – airborne applications. For many functions that benefit outweighs further refinement of sensing or processing, which have already improved enormously. Aside from a few exceptions (*e.g.*, imaging – with significant benefits still realizable by better processing of better data while expending less resources), priority can now shift to the one step offering the biggest payoff by far: best selection of data to be routed between subsystems. Evidence of growing (and sometimes blunt) reaction to this awareness can be found in numerous publications, *e.g.*, [6]:

> "⋯ *vendor lock-down, once considered a bitter pill that one had to swallow to deploy a solution rapidly, is no longer deemed acceptable* ⋯ *increased value and liquidity of data and applications that result from the use of standards has become much clearer* ⋯ *proprietary approach simply would not affordably provide interoperability across the broad spectrum* ⋯ "

Again, the realization is widespread: the best sensors imaginable lose effectiveness as their outputs are severely degraded ahead of external processing – and the most ingenious processing methods still need good inputs.

As the industry comes to demand far greater usefulness of data from separate "boxes" supplied by myriad independent sources, timing will increase in importance. A major irony of GPS is its enormous improvement of navigation and synchronization for communication – *but not* synchronization *in* navigation systems ! Ramifications of timing in data from different modules arise constantly in operation. A modern example is GNSS / INS integration; an older example is SAR (with motion compensation governed by excursions while transmitting radar pulses). The subject is discussed further within subsections of APPENDIX III.2.

Tight integration of multiple subsystems driven by separate internal clocks will require closer attention to the timing of signals from each "box." The familiar pre-GPS artificial computational synchronization (with each algorithm reinterpreting time tags of every sample received from every asynchronous module feeding it) imposed a heavy burden in operational performance (*e.g.*, latency) and software complexity. While those procedures might be retained as backup, GNSS timing offers more expeditious methods (Chapter 7).

Implicit in these descriptions is the need for change. That is the main price to be paid for the benefits available. Validation and certification will involve still more change. For example, tests that now focus on solutions from navigation subsystems can be compartmentalized differently – sensor / system / software. Sensor *measurement* time histories can be compared in blind test *vs* sequences chosen from libraries derived, documented, and standardized for any requisite scenario set. Completely independent of all sensor tests, algorithm validation can generate solutions from measurement data streams also corresponding to specified scenarios.

Without the kind of focus and standardization just exemplified, the need to accommodate a broad variety of capabilities can result in performance degraded to a "lowest common denominator" level. Conversion of established procedure is not at all a trivial step, but there is much to be gained. Full usage of methods established over decades, with demonstrated widespread success in myriad applications, can finally curtail constant duplication and reinvention while offering a low-cost way to produce a quantum increase in capability.

The main thrust of this book, with its emphasis on formulations and algorithms, is now complete. It is recognized that such a primary concern may leave some uncertainty (*e.g.*, "How realistic is improvement in both performance and economy when institutional issues and other factors are taken into account ?"). APPENDIX III addresses those items; here a clear payoff description suffices:

- Unprecedented integration within each participant's system. Recall the discussion regarding every possible input combination in Section 10.2 – each subsystem could have prompt access to all information available to or generated by every other subsystem.
- Unprecedented situation awareness . Every operation (command, control, communication, nav, surveillance) involves activity (*e.g.*, sensing) that potentially enhances performance of every other operation; every subsystem belonging to any one participant could maximally support each function for every participant.

References

[1] Saks, S.L., et al., "A modern generic IMU interface standard," IAIN-ION World Congress, 2000.

[2] Wade, M. and Grewal, M.S., "Analysis of a Cascaded INS Calibration Filter," IEEE PLANS, 1988.

[3] Kalafus, R.M., vanDierendonck, A.J., and Pealer, N.A., "Special Committee 104 Recommendations for Differential GPS Service," *ION Journal*, Spring 1986, *pp*. 26-41.

[4] Farrell, J.L., McConkey, E.D. and Stephens, C.G., "Send measurements, not coordinates," *ION Journal*, Fall 1999, *pp*. 203-215.

[5] Neregard, F., et al., "TERNAV, an algorithm for real time terrain navigation and results from flight trials using a laser altimeter," ION GNSS-2006.

[6] Buehler, K., "Interoperability - it's mission critical," *GeoIntelligence*, Mar/Apr 2004, *pp*.36-40.

APPENDICES

APPENDIX I

NOMENCLATURE – MAJOR NOTATION

The purpose of this APPENDIX is to facilitate interpretation of analytical material applicable to multiple areas of the book. This guide will be helpful in developments where results from other sections or other chapters are cited. While there are several such instances, it is also true that many major sections of the book can be followed without this guide. Those include

- chapters having little or no mathematical symbols (1, 8, and 10)
- chapters (such as 7 and 9) wherein most of the notation is specialized – *i.e.*, pertinent to the local section only – and the few exceptions are unmistakably recognizable from discussion in accompanying text.

Likewise in conformance to the latter condition are developments in all Addendum and APPENDIX material. Consequently only the main body (pre-Addendum) from Chapters 2-6 will be included here. Coverage of those areas, which contain most of the cross-referenced analysis, satisfies the purpose just noted.

Brackets have their usual relation to contents within, *e.g.*, $< \bullet >$ is an ensemble mean, $| \bullet |$ is an absolute value for a scalar, magnitude for a vector or, for a matrix, the determinant. Self-evident nomenclature such as any standard function is excluded from this tabulation. Also, some symbols used in only a few places (*e.g.*, E in Section 3.4.3; r in 4.3.2), plus a few items with obvious meanings – the ubiquitous t or τ for time; σ for RMS measurement error; and multi-character variables (*Lat*, *Lon*) – are omitted from the notation table to follow. Letters with multiple meanings are distinguished by typeface in many cases (*e.g.*, $\boldsymbol{q}$ in Chapter 2; $\boldsymbol{q}$ in Chapter 3; $\mathbf{q}$ in Chapter 6), but there are a few exceptions. For example, Section I.1 does not cite the appearance of the following variables in areas where their meanings (differing from tabulations here) are clear from local context:

- R (scalar measurement error variance) in Section 2.3
- q (4-parameter attitude array elements) in Section 3.4.3
- T (direction cosine matrix elements) in Sections 3.3, 3.4.3
- M (dimensionality of data vectors) in Chapter 6.

Separate tabulations for scalars, vectors, and matrices are followed by discussion of subscripts and superscripts. Notes are given at the end of any table needing clarification of unusual notations. For acronyms see the Index.

I.1 SCALARS

Scalars are light-faced and italicized. Most are shown here without subscripts, except where the subscript is an integral part of the designation (as in a_E, R_M, R_P).

Wherever appropriate, a symbol is accompanied by units shown to its right. Although not intended to limit application, those optional units allow equations to be applied with no added conversion factors. Meters (m) and seconds (s) are used in general. For quantities that are dimensionless, including all angles (expressed in radians as defined here), and for quantities that may attach different units to different indices (such as state variables or other array elements), no units are given.

Symbol		Definition	Sections
a_E	m	WGS-84 Earth ellipsoid semimajor axis	3.4
b	m	measurement bias	6.1.1-2, 6.2.3, 6.2.3.1-2, 6.3.4
c	m/s	speed of light	2.8.1-2
e_E		WGS-84 Earth ellipsoid eccentricity	3.4, 5.1.1
f		WGS-84 Earth ellipsoid flattening	3.4
g	m/s²	gravity vector magnitude	2.7, 2.8.1, 3.6, 4.1, 4.3.1-2, 4.5-6, 5.6.1-2
h	m	altitude	3.4-5, 5.1.1
m		measurement number	2.3, 2.6.1-2
n		inertial instrument bias	4.3, 4.3.2, 4.3.4
q		component of **q**	6.1.1
r	m	difference R_P - R_M	3.4, 3.5.3
v	m/s	component of velocity error **v**	4.3, 4.3.1-2
x		scalar state or component of error state **x**	2.4.1, 5.6.1, 6.1.2-3, 6.2.2-3
z		scalar residual or component of **z**	2.3, 2.4.1, 2.5, 2.8.1, 5.2, 6.1.2, 6.2.3, 6.3.2
M		number of measurements in a data block	2.6.1-2, 2.9, 4.3.1-2, 5.6, 5.6.1
M		number of measurements in a data window	2.4.1
P_A, P_{MD}		alarm and missed detection probabilities	6.1.1, 6.2.3, 6.2.3.3, 6.3.4
R	m	distance component	2.8.2, 4.2
R_M, R_P	m	curvature radii along meridian and equator	3.4, 3.5.3, 4.1, 4.3.1-2, 5.1.1, 5.4.2
T_D		data edit threshold	6.1.1, 6.2.3.3, 6.4
T	s	data block duration	2.6.1-2, 2.9, 4.3.1-2, 5.6, 5.6.1
T	s	duration of data window for recursion	2.4.1, 2.9, 4.5, 5.4.2
V	m/s	components of **V** (See Note in Section I.4)	2.1, 2.4, 2.4.1, 2.6.1-2, 3.4, 4.2
W		scalar residual weight or component of **W**	2.4.1, 4.2
W	s⁻¹	Schuler rate, $\sqrt{g/(earth\ radius)}$	4.1, 4.3.1-2
X		scalar state or component of state vector **X**	2.1-4, 2.4.1, 2.6.1-2
Y		scalar observable	2.3, 2.4.1, 2.6.1, 2.8.1-2, 5.2-3, 5.4.1-2
Z	m/s²	scalar acceleration	2.1, 2.6.2

Symbol		Definition	Sections
α		wander angle	3.3, 3.4, 3.4.2, 3.5
α		angle of attack	4.6
ϵ		measurement error	2.4.1, 2.8.1, 5.2-3, 5.6, 6.1.2, 6.3.2
η		process noise spectral density	4.5
κ		cosine of half-angle $\lvert \boldsymbol{\chi} \rvert / 2$	3.4.3
λ	m	transmitter wavelength	2.8.2, 5.6.2
μ		product $\lvert \boldsymbol{\omega}_R \rvert \times$ (computation interval)	3.4.1
χ		component of relative rotation increment	3.4.3
ψ		component of misorientation $\boldsymbol{\psi}$	4.3.1-2
ω_S	s^{-1}	magnitude $\lvert \boldsymbol{\omega}_S \rvert$	3.4.2

I.2 VECTORS

Almost all arrays tabulated here are column vectors, denoted by bold face with applicable dimensions $n \times 1$ where n is either

- 3 (possibly 2 in Sections 2.2, 2.2.1) – in which case the "Index" column below may contain, in parentheses, an indication of the applicable coordinate frame(s) – or
- a value other than 3 – which then shows as the "Index" without parentheses. In that case, N and M are state and data vector dimensionality, respectively.

The only exceptions are row vectors **H**, **h**, **h**. Again most symbols are tabulated here without subscripts (with some exceptions) and units are optional as in Section I.1.

Symbol		Index	Definition	Sections
b	m	M	measurement bias vector	6.1.1, 6.2.3, 6.2.3.1-2
e		N	process noise	2.5, 4.1, 4.3, 4.3.3-4
g	m/s^2	(L)	gravity vector	2.7, 3.4.2, 4.1, 4.6
h, **h**		1×3	scalar measurement sensitivity, partition of **H**	5.2, 5.4.1-2, 6.1.2-3, 6.2.3, 6.2.3.3, 6.3.2
l	m	(A)	IMU-to-GNSS receiver lever arm	5.4.1-2
n		(A)	inertial instrument bias	2.7, 4.1, 4.3, 4.3.3, 5.6.3
q	m/s	(L)	change in IMU velocity	2.2.1
q		4	array of quaternion elements	3.3, 3.4.3, 3.5, 3.5.1
q		M	normalized right-hand column of **Q**	6.1.1, 6.2.3.1, 6.2.3.3, 6.4
r	m	(L)	error in position vector **R**	2.8.1, 4.1, 4.3.1, 4.3.3, 5.1, 5.1.1, 5.2-3, 6.2.1, 6.2.3, 6.3.1
v	m/s	(L)	error in velocity vector **V**	4.1, 4.3, 4.3.1, 4.3.3-4, 5.4.2
x		N	error state vector	2.2, 2.5, 2.6.1, 5.1, 5.1.1, 5.2, 5.4.2, 6.1.1-3, 6.3, 6.3.1-2
z		M	residual vector	2.5, 2.9, 5.1-2, 6.1.1-3, 6.2.3, 6.2.3.2-3, 6.3.2, 6.3.4, 6.4

Symbol		Index	Definition	Sections
A	m/s^2	(A,I,L)	specific force vector	2.7, 3.4.2, 3.5.2, 4.1, 4.3, 4.3.1, 4.3.3-4, 4.5, 5.4.2
E		(A,P)	principal unit vector of nav / vehicle coordinate transformation	3.4.3
F	m/s^2	(A)	specific force vector	4.6
G	m/s^2	(I)	gravitation vector	3.4.2
H		$1\times N$	measurement sensitivity vector	2.5, 2.9, 4.4-5
J		$1\times N$	measurement sensitivity vector to $\mathbf{X}_0$	2.6.1-2
I, **J**, **K**		(A,E,I,L,P)	mutually orthogonal unit vector triad	3.2-4, 3.4.1-3, 4.1, 5.1.1, 5.4.2
L	m	(L)	IMU-to-GNSS receiver lever arm	5.4.1-2
Q	m	(L)	time integral of ***q***	2.2.1
R	m	(L)	position vector	2.2, 2.8.1-2, 3.4.2, 4.1, 4.3.1, 5.1.1, 5.2-3, 5.4.1-2, 5.5.2
R	m	(L)	position relative to moving reference	2.2.1
S	m	(L)	instantaneous satellite position	2.8.1-2, 5.2-3, 5.4.1
U	m/s	(I)	velocity vector w.r.t. geocenter	3.4.2
V	m/s	(L)	velocity vector w.r.t. spinning Earth	2.2, 2.7, 3.4, 3.4.1-2, 3.5, 4.1, 4.6, 5.4.1-2, 5.5, 5.5.2
V	m/s	(L)	velocity relative to moving reference	2.2.1
W		*N*	weight vector for a scalar residual	2.4.1, 2.5, 2.9, 4.4
X		*N*	state vector	2.4.1, 2.5, 6.3.2
Y		*M*	data vector	2.5
Z	m/s^2	(L)	acceleration vector	2.2, 2.2.1
$\Delta\mathbf{V}$	m/s	(L,P)	transformed accelerometer increment vector	4.1-2, 5.4.2
$\boldsymbol{\Delta\theta}$		(A)	vector of small-angle gyro increments	3.4.3
$\boldsymbol{\epsilon}$		*M*	measurement error vector	2.5, 5.1-2, 6.1.1-3, 6.2.3, 6.3.2
$\boldsymbol{\eta}$		(L)	unit vector in direction of $\boldsymbol{\omega}_R$	3.4.1
$\boldsymbol{\chi}$		(A)	$\boldsymbol{\Delta\theta}$ corrected for rotation of nav frame	3.4.3, 3.5, 3.5.1
$\boldsymbol{\psi}$		(L)	vector of small-angle attitude errors	2.7, 3.6.1, 4.1, 4.3, 4.3.1, 4.3.3-4, 4.6, 5.4.2
$\boldsymbol{\omega}$	s^{-1}	Note #1	angular rate vector	2.7, 3.4, 3.4.1-3, 3.5, 4.6, 5.4.2
1		Note #2	unit vector or column of identity matrix	2.2, 2.8.1, 4.3, 5.1-3, 5.4.1-2, 6.1.1-2, 6.2.3, 6.2.3.1-2
0		Note #2	null vector	2.2, 4.1, 4.3, 4.3.4, 5.4.2, 6.1.2-3, 6.2.3, 6.3.2

Notes
1. With coordinate axis subscripts (A, E, L, or P) $\boldsymbol{\omega}$ denotes the absolute (inertial) angular rate of the coordinate frame, with components expressed in that frame. With other subscripts $\boldsymbol{\omega}$ has other meanings –
   - $\boldsymbol{\omega}_S$ represents $\boldsymbol{\omega}_E$ transformed into nav coordinates.
   - $\boldsymbol{\omega}_R$ is the angular rate of the nav frame relative to the spinning earth.
2. Vectors **1** and **0** have dimensions of varying partition sizes, easily determined by context.

I.3 MATRICES

A bold Arial character denotes a matrix. Except for 3×3 arrays, dimensions in most cases conform to size of either the state N or a data vector (normally denoted M but modified within the context of Section 5.6). A few, such as **I** or **O**, are dependent upon partition sizes.

Sym	Size	Definition	Sections
$\mathbf{A}$	$N \times N$	dynamics matrix	2.2, 2.5
$\mathbf{C}$	$M \times M$	inverse of $\mathbf{L}$	5.2
$\boldsymbol{D}$	$N \times N$	Bierman's diagonal factor of $\mathbf{P}$	5.5.1, 6.3.2
$\mathbf{E}$	$N \times N$	process noise	2.5, 2.9, 4.5
$\mathbf{H}$	$M \times N$	measurement sensitivity to current state	2.5, 2.9, 4.4, 5.1, 5.1.1, 5.2, 5.4.1-2, 6.1.1-3, 6.2.2, 6.2.3.1, 6.3.2
$\boldsymbol{H}$	$M \times N$	modified measurement sensitivity	5.2, 6.1.3, 6.2.3, 6.2.3.3
$\mathbf{I}$		identity	2.2, 2.5, 4.1, 4.3.3, 4.5-6, 5.4.2, 6.1.1, 6.2.1, 6.3.2
$\mathbf{J}$	$M \times N$	measurement sensitivity to initial state	2.6.1-2, 5.6
$\mathbf{L}$	$M \times M$	normalized square root of measurement error covariance matrix	5.2
$\mathbf{O}$		null matrix	2.2, 2.5, 4.1, 4.3, 4.3.3-4, 4.5, 5.4.2, 6.1.1, 6.3.2
$\mathbf{P}$	$N \times N$	state uncertainty covariance matrix	2.5, 2.9, 4.4, 4.5.1, 5.5.1, 6.2.3.3, 6.3, 6.3.1-4
$\mathbf{Q}$	$M \times M$	orthogonal factor of normalized measurement sensitivity $\mathbf{UH}$	6.1.1, 6.1.3, 6.2.2, 6.2.3.1, 6.2.3.3, 6.3.2
$\mathbf{Q_P}$		$M - N$ columns on the right of $\mathbf{Q}$	6.1.1, 6.2.3.1, 6.2.3.3, 6.3.2
$\mathbf{Q_x}$	$M \times N$	N left columns of $\mathbf{Q}$	6.1.1, 6.2.3.1, 6.2.3.3, 6.3.2
$\mathbf{R}$	$N \times N$	upper triangular factor of normalized measurement sensitivity $\mathbf{UH}$	6.1.1, 6.1.3, 6.2.2, 6.2.3.1, 6.2.3.3, 6.3.2
$\mathbf{R_x}$	$N \times N$	top N rows of $\mathbf{R}$	6.1.1, 6.2.3.1, 6.2.3.3, 6.3.2
$\boldsymbol{R}$	$M \times M$	measurement error covariance matrix	2.5, 2.9
$\mathbf{S}$	$M \times N$	measurement sensitivity $\mathbf{J}$ normalized	5.6, 6.3.2
$\mathbf{T}$	3×3	direction cosine matrix	2.7, 3.3, 3.4.1-3, 3.5, 3.5.1-3, 4.1, 4.3, 4.3.3-4, 4.5, 4.5.1, 4.6, 5.4.1-2
$\mathbf{U}$	$M \times M$	weighting matrix for measurement block	5.6, 6.1.1, 6.2.1, 6.2.3, 6.2.3.1, 6.2.3.3, 6.3.2
$\boldsymbol{U}$	$N \times N$	Bierman's upper triangular factor of $\mathbf{P}$	5.5.1, 6.3.2
$\mathbf{W}$	$N \times M$	measurement residual weighting matrix	2.5, 2.9
$\mathbf{X}$	3×3	vector product operator $(- \boldsymbol{\eta} \times \;)$	3.4.1
$\boldsymbol{\Gamma}$	$M \times M$	normalized inverse square root of $\mathbf{U}$	6.1.1-2, 6.2.1, 6.2.3.3, 6.3.2
$\boldsymbol{\Phi}$	$N \times N$	state transition matrix	2.2, 2.2.1, 2.4.1, 2.5, 2.6.1, 2.9, 4.1, 4.4, 4.5, 5.4.2, 5.5.1

I.4 SUBSCRIPTS AND SUPERSCRIPTS

Rather than conformance to a rigid pattern, an opportunistic approach is taken for usage of subscripts. Following are the main categories:

- nav reference coordinates (A,E,I.L,P) for vectors such as **R**, **V**, etc.
- double-coordinate subscripts separated by a slash for a direction cosine matrix, *e.g.*, $\mathbf{T}_{A/L}$ is the transformation from nav reference to vehicle coordinates; $\mathbf{T}_{L/E}$ is the transformation from Earth to nav reference coordinates.
- $1,2,3$ or x, y, z for componens of a vector.
- a, ω – accelerometer or gyro error, respectively, as applied to **n** or **e**.
- m – sometimes combined with other subscripts – denoting occurrence at time t_m.

Other subscripts are also used in combinations. Components of velocity V, for example, sometimes appear with both a coordinate axis designation (P) and an integer (1, 2, or 3) for direction – or the direction can be explicit (North, East). Alternatively, 0 is used denote initial or reference value. In any case, subscripts of V can be easily interpreted from context.

For superscripts the familiar notations apply –

- a dot above for time derivative,
- a circumflex (^) for apparent (observed or estimated) value,
- a tilde (~) for error in apparent value,
- -1 for inverse and # for generalized inverse of a matrix
- T for transpose of a vector or matrix
- -T for transposed inverse of a matrix
- $(-)$ to denote *a priori* (pre-observation) value
- $(+)$ to denote *a posteriori* (post-observation; post-update) value

APPENDIX II

FREE-INERTIAL COAST

For several years the navigation community has been (1) accustomed to achieving high accuracy by using satellite signals and (2) interested in backup methods to be used when satellite signals are unavailable. A free-inertial operation, beginning with the last GNSS update received, has been advanced as a possible means of backup. For that reason, after focusing this entire book on short-term operation supported by availability of repetitive updating, the opposite subject – navigation over a long term without available updates – is finally treated here.

The need to assess coast capabilities and limitations has prompted various investigations, with varying conditions affecting different details in methodology. While those efforts show a welcome awareness, an ever-widening array of methods and programs could produce a wealth of results derived for many different configurations, conditions, characterizations, etc.; decisions would not readily be made on the basis of one-to-one comparison. A program written by this author [1] is offered as a draft that could eventually enable agreement on a set of IMU error propagation characteristics applied to standard test scenarios, universally available for evaluating performance of any candidate IMU in practical cases of interest to all. Certain phrases in that stated goal carry implications regarding means of implementation; these include

- "universally available" – This would preclude blind attempts to compare unequal sets of results based on unverifiable statements accompanied by proprietary claims.
- "standard test scenarios" – Candidate IMUs must be compared on an equal footing.
- "practical cases of interest to all" – Each test scenario should be chosen to convey information that will clarify either the error propagation phenomenon itself or its effects in flight segments where coast performance determines success or failure.
- "enable agreement" – Means of evaluation can undergo a sequence of draft versions, each intended to improve and/or broaden the scope of its predecessor.
- "set of IMU error propagation characteristics" – A future draft should eventually become a final evaluation tool, with all dominant degradation sources present.

It goes without saying that realism is of paramount importance here. The final form of what eventually evolves is intended to place bounds on performance that will actually result in flight. Claims and specifications should be subject to real-world verification.

A modular MATLAB®† program furnished here facilitates future program extensions. Sample results are included with supporting covariance analysis for a subset of conditions addressed. Particular attention is drawn to perishability of pre-coast "calibration" (an overstatement of capability., since errors maintained at null by pre-coast update have higher order derivatives "learned" from conditions that change after corrections stop), especially with limited time constants of low-cost instruments.

The program "coast.m" uses scripts and functions including "init.m" (initialization), "scenar.m" (scenario), "xrand.m" (random error generation), "eprop.m" (error propagation), "bkpg_.m" (various bookkeeping modules), and "grafx_.m" (plotting modules). With this structure, a scenario can be added in the future with little or no change to the error propagation module – or vice-versa. Three main scenarios with perfect position and velocity at coast initiation are included here:

- Dispatch: aircraft stationary on ground, correctly showing and maintaining zero velocity. Without GNSS, takeoff begins at zero time, followed by ascent to altitude where limited (*i.e.*, a planned subset of existing) VOR updates can be obtained.
- Straight-and-level flight, with GNSS fully operational, initially achieves zero velocity error. All GNSS updating cuts off at $t = 0$ when a single 180° turn begins.
- Same as just described, except that the turn is the first quadrant of a repetitive sequence of holding pattern cycles, each cycle consisting of four *1*-min flight legs wherein semicircle quadrants alternate with straight segments.

Another simple scenario ("#*0*") was added for validation, beginning with the same conditions as the second case but without the turn.

Each run begins with $t = 0$ at the instant of GNSS data cutoff. Position and velocity errors here will rise from zero at that time, in accordance with propagation characteristics resulting from a total effective drift rate vector $\mathbf{N}_{\omega}$ and a total acceleration error vector $\mathbf{N}_A$ containing accelerometer offsets combined with vertical deflection effects. Although GNSS updating will maintain position and velocity errors at negligible levels prior to $t = 0$, there will be nonzero initial conditions applicable to higher-order states due to Kalman filter operation with imperfect modeling. Specifically the estimation process – implemented with prompt responsiveness to counteract high drift rates of low-cost gyros – attributes acceleration offsets to tilt variations (an attempt to include accelerometer states, not distinguishable within the short response time, offers no solution for that initialization error).

Earlier work [2] produced a non-modular listing, providing approximate effects of gyro cross-axis errors only. This first draft of a modular program enables broader investigation that can grow to cover all major sources. It begins by adding gyro scale factor plus band-limited vertical deflections as well as accelerometer and gyro biases. Plots of sample results, with typical time constants adopted from [3] and [4], are presented with interpretations and suggested future enhancements that other industry participants are invited to add. Complete MATLAB listings then follow.

† MATLAB is a registered tredemark of The MathWorks, Inc.

APPENDIX III

IMPLEMENTATION ISSUES

A brief discussion is provided here to consider
- some constraints imposed by institutional considerations
- further interface details
- software ramifications

including factors affecting overall coordination. In these discussions brevity is essential. There is no distraction from mathematical developments, exhaustive bibliographies, devices unrelated to navigation and tracking, nor unusual case histories. Only cursory attention is given to examples from specific designs

No attempt was made to be thorough in presenting the material that follows. Full coverage would have consumed disproportionate length, distracting from the objective underlying the entire approach built throughout this book – integration carried to the limit of achievable potential. Ramifications of that objective extend in myriad directions, many of which are peripheral here. Still, their discussion in these supplements helps to establish a principle of crucial importance, *i.e.*, integration at an unprecedented level is reachable. That is the main message carried by this APPENDIX: while prescribing significant departures from common practice in several areas, practical issues have not by any means been overlooked.

III.1 INSTITUTIONAL CONSTRAINTS

While formulations and algorithms presented hare have general (air / land / sea) adaptability, institutionalization for land vehicles is embryonic by comparison to the others. Also, for ships even a basic advance – from paper charts to Electronic Chart Display and Information Systems (ECDIS) – has encountered appreciable delay of official approval. Reasons for that, though of course valid, involve pertinent international administrative and legal issues which cannot be pursued here. This immediate section concentrates most heavily on airborne operations (where safety issues exert an obvious dominant influence) within the National Airspace System (NAS) – with further emphasis on commercial aviation (where procedures are most heavily restricted).

III.1.1 Air Traffic Control (ATC) Applications

It is universally acknowledged that today's ATC methodology, bearing little resemblance to any design that would be devised today, evolved by gradual accumulation over many decades. Throughout almost all of that growth period, techniques described in the earlier chapters were beyond reach. Other approaches, limited by earlier technological capabilities, naturally influenced the procedures that were adopted and institutionalized. Understandably a rapid pace of change in those procedures is inhibited by risk. At the same time, undue delay in adopting newer methods forfeits the twin benefits (high performance concurrent with economy) that could otherwise become available. Following is a simple step applicable to aviation in achieving those goals while still observing institutional constraints.

Some complications that arose from enhancement of ATC capabilities, solved long ago, can serve as a model for smooth transitions in the future. An immediate example involves the pilot's display, which has for many years accommodated upgrades by adoption of well-known features including

- replacement of multiple gauges by multiple screen selection options
- access of each individual reading (*e.g.*, pressure, precipitation, microburst, wind shear, ⋯) to specific information relevant to it
- presentation of information in a form not requiring pilot retraining.

The latter item offers an opportunity to accommodate customary guidance methods while exploiting advanced means of generating the requisite commands. The technique, already decades old, can be described as in [1]; some main points are paraphrased below:

A differential GPS aircraft landing system uses airborne and ground GPS receivers, plus all provisions needed for data communication and computation. Positions of two points (one airborne and the other established by a ground station) define a prescribed glide path. The horizontal projection of that path is compared *vs* a runway centerline, from which follows a lateral deviation. The path slope is compared to a nominal glide slope, producing a vertical deviation. Historically those deviations were sensed by comparing intensities of different modulating tones in signals received from localizer and glide path transmitters in the vicinity of a landing strip;

- a difference in depth of modulation (DDM) measured by a localizer receiver is essentially proportional to the lateral deviation.
- A DDM measured by a glide path receiver is essentially proportional to the vertical deviation.

Simply by generating DDM signals corresponding to computed lateral and vertical deviations, the familiar steering signals are derived. Observing all institutional guidelines, that “transparent display” approach and additional analogous procedures can accommodate further modernization, in the insertion of any nav and tracking information from earlier chapters.

III.2 DATA FLOW

To many in the electronics field the word *interface* invokes thoughts of input/output (I/O) conventions (1553, RS-232, etc.) concerning issues such as signal formats, voltage and impedance levels, modulation methods, or bus topology. Here we presuppose successful choices made in all those important decisions, concentrating instead on the information carried from every source to each destination. A further subdivision follows, distinguishing flow within a participant's system (internal I/O) from data transmission and reception among multiple participants (external I/O). The role of timing is revisited, with added considerations arising from improved tagging.

III.2.1 Internal I/O

The data bus long ago replaced an antiquated "one-wire-one-signal" heavy cabling structure but, as implemented thus far, this advance has inflicted a hidden penalty on some operations. Properly shielded wiring conveyed signals as promptly as finite bandwidths would allow; aside from minimal residual crosstalk and minimal limitations intrinsic to existing passbands (transport lags, distortion due to nonlinear phase shift), no flaws were inserted by the signal path. Data bus arrangements introduced opportunities to deposit information temporarily at intermediate points along *indirect* paths to ultimate destinations. With data dispatch and receipt controlled by asynchronous computers, latency is not governed by finite data rates. Often the performance degradation is compounded by *multiple* asynchronous computers in the path. The problem is important enough to warrant further scrutiny given below to data handling.

Implications of asynchronism can be described in many ways. For an introductory oversimplified "tip of the iceberg" example consider angles summed for a climbing aircraft at zero roll, carrying a radar with its beam instantaneously elevated while at airframe-reference azimuth. From subsystems with uncorrected time stagger, indicated airframe pitch plus radar elevation could produce a spatial inclination that was never reached. Compensations often introduced to alleviate the degradation are conceptually simple but, from the standpoint of performance and procedures, crude. A more involved operational example will now be discussed, for which the undesirable impact of that crudeness is made clear.

Descriptive material immediately following Figure 9.3 is now revisited. Direction to the target's anticipated location at the time of the next angle measurement is to be compared *vs* vehicle direction *at that same future time* (that latter caveat was taken for granted in Chapter 9, with further attention shunned to avoid distraction in explaining Figure 9.3). Success of the action is clearly affected by the ability to time-align signals *and to extrapolate*. Imagine stabilization of a sensor sightline during a snap roll; the data informing a controller's action will obviously degrade when signals arrive late. The ramifications in this case can include temporary absence of a tracked target from sensor field-of-view:

- When the radar-carrying aircraft is a fighter, the amount of rotation between successive observations can exceed the allowable beam stabilization error (*e.g.*, a half-beamwidth). This highlights a *disadvantage* of recent technology advances: Older systems had gyro-stabilized platforms and gyro-stabilized servo-driven antennas – both highly responsive. With a strapdown IMU and an array antenna (conformal or not), the beam being stabilized is an intangible line in space. Mixing "yesterday's" aircraft attitude with "today's" sensor pointing angles can be avoided, but only by staying current while minimizing timing offsets between them.
- Performance with pre-GPS artificial computational synchronization (Section 10.3) is vulnerable. Modest data rates produce latency while sample timing, neither closely controlled nor consistent, produces uncertainty in time stamps.
- Extrapolation based on sampled values, inherently a procedure with performance limitations and pitfalls, is further degraded by the item just described. A similar statement applies to interpolation, with reduced uncertainty but increased latency.
- It often occurs that a computational synchronization must be constructed within constraints of an interface configuration long since frozen. Data to be brought together can then traverse multiple independent processors before meeting – producing *cumulative* lags and uncertainties incurred along the data flow path.
- Interrupts and protocols can differ among mechanizations of various "vintage" – and seemingly minor variations can introduce unexpectedly greater complications.
- Epochal events (*e.g.*, clock walk-through, recycling), if undetected, can produce catastrophic failure.

Discontinuity of sensor reports due to nonessential data lag exhibits a flaw in system design methodology. There is a correction strategy: ***Before specification of any data bus*** provisions, a transfer matrix (*e.g.*, Appendix III of [2]) can define the source and destination for every signal: *who* needs *what*, *when*, how *often*, and from *where*. The computer-controlled data transmission system is ***then*** devised to answer the question *how*. The question of *what* is needed includes not only information content but also its form and precision. With signal paths dictated by need, layout imposes conditions; timing requirements will sometimes ***veto*** convenience of assembling a path from multiple bus branches.

When data transfer requirements are given full preeminence to lead (not follow) data bus definition, state of the art in digital data transmission – far advanced from an implementation standpoint – will provide a matching advancement in system performance. A well-conceived interface layout will be the backbone of a successful avionics / vetronics / shipboard electronics design. Controls and displays can then conform to their roles as final destinations – not defining specifications – for information generated. The former (controls) can involve conversion (*e.g.*, to analog form and amplification) as necessary. Only after all system requirements have been satisfied, preparation of information for control and display signals can then be generated by modular functions, governed by programmable algorithms whose

- inputs are derived from the perceived state of the system and
- outputs can influence *only* the disposition (*not* that perception) of the state.

III.2.2 External I/O

As opposed to internal transmission (*e.g.*, along a data bus), this section addresses communication of information externally (*e.g.*, to a wingman, a fleet, or a command center). Developments over recent decades have enabled profound improvements in this field also. Digitization of course allowed expression of time periods as precise multiples of diminutive intervals, producing opportunities that include

- freedom to choose those multiples to be integrally related – so that carrier frequencies can be exact multiples of frequencies that govern coded messages, thus constraining code epochs to start at zero carrier phase.
- freedom to choose the diminutive intervals small enough to prevent significant propagation of time quantization effects – even down to atomic clock periods.
- ability to make hardware component substitutions – such as replacement of a voltage-controlled crystal oscillator ("VCXO") with a numerically controlled oscillator ("NCO") – thus stabilizing frequencies as never before.

These developments, not surprisingly, benefit every conceivable facet of communications. Modulation in many instances became simple switching of amplitude, frequency, or phase. Demodulation was reduced to simple detection of those keying events. Multiplexing and demultiplexing (time, frequency, and code) improved for similar reasons. Synchronization at all levels (page, frame, word, character, bit) remains limited by the quality of incoming data – but, with optimized algorithms, to a much reduced extent – and no longer captive to basic hardware limitations. The same applies to tracking of carrier frequency and phase, and to extraction of weak signals (whether narrowband or spread spectrum) in the presence of noise and intermodulation or other distortion. Algorithmic approaches likewise systematized the encryption/decryption process, as well as error detection/correction and lossy or lossless bandwidth compression.

If all of the above can be called "old news" then what remains for a system equipped with all bus types plus hardware and software provisions to exploit advances just noted? From an integration standpoint, the answer lies in rules governing choice and transmission of data to be communicated. One enhancement could be acceptance of a unified time base (to exploit precise timing capability for packet switching or for squitters emitting interrogations at integral seconds of GPS time). For data selection everything previously said in connection with nav and surveillance applies here directly. That is hardly surprising, despite association of nav data with onboard equipment and surveillance data with external observations (with traditional separation of those functions in civil aviation). Both concern instantaneous state of objects in a scenario. In marked contrast to communication that actually is largely counterproductive (Section 7.2.1 described one case [3]; others have been reported), methods of Chapter 9 efficiently enable accurate determination of those states.

For allowable data volume, finite bandwidth or power must accommodate overhead (preambles, error detection / correction, etc.) in addition to messages. Information capacity is thus somewhat reduced, but not drastically so.

III.2.3 Timing: Further Details

The self-evident requirement – data subject to rapid variation must have accurate time tagging to be fully effective – has already been discussed. Likewise already covered are the expeditious methods, mentioned in Chapter 7 and cited near the end of Section 10.3, for achieving desired outcomes. A few timing issues, not yet fully discussed, can have ramifications that either need to be more thoroughly examined or are not necessarily limited to internal or external I/O (potentially affecting either or both). These include the following:

- Delay in generating feedback signals for a control loop can compromise performance and even threaten stability. Not all decision loops are purely internal.
- Communications often waive punctuality requirements while still disallowing arbitrary timing (so that, for example, different packets belonging to the same message can dependably be stitched together – though not always with promptness and not necessarily without quiet periods between message constituents).
- Again irrespective of punctuality, data *refresh* rates for surveillance must support correct association of sensor responses with corresponding track files. This does not imply fixed refresh rates – aperiodicity is generally acceptable – but extended quiet periods between consecutive unidentified reports (*e.g.*, optical updates) would produce observations with inconclusive file association.
- Lags, arbitrary timing, and infrequent data are all acceptable for navaids updating a dead reckoning device, as long as *uncertainty* in an observation's timing is less than (RMS measurement error) / (rate of change in quantity being measured).
- When a time base governing all equipment on a single participant's platform is unified, a quantum improvement in that individual's outputs can result. A unified time base over *all* participants in a scenario can produce a quantum improvement evident across an entire fleet.

Nuances of this type can exert major influence on system performance. The subject of control, in particular, projects data timing issues in a unique light. Modernization introduces changes that affect implementation, but not basic requirements. Section III.2.1 discussed the replacement of mechanical beam steering by computation of phase shift commands for an array antenna. Timing demands for a beam being driven

- to produce a desired scan pattern, or
- in response to changing position of an object being tracked, or
- for stabilization (*i.e.*, changes in direction relative to the airframe, to counteract an unwanted spatial redirection caused by airframe rotation)

are independent of whether some control functions are transformed into processing chores (*e.g.*, digitization of some – formerly mechanical – actuator driving tasks). Similar considerations apply for antenna null steering and, with some focal plane arrays, for mechanical stabilization of an electro-optical sensor sightline. Irrespective of mechanization details, timing requirements are imposed on the generation of an input to a digital-to-analog converter (DAC) or an actuator.

It is worth adding an observation that, where physical motion is to be controlled, conventional servos are often adequate (with no pressing need for optimal controllers). Reasons include the following:

- Many of the controllers needed do not involve the complications and challenges (*e.g.*, highly nonlinear characteristics, several inputs and outputs) accommodated by the more powerful modern methods. Moreover, some involving nonlinearities with multiple outputs and inputs (such as flight control via aileron/elevator/rudder deflection) have well-established linearized models applicable through overlapping speed/altitude flight regimes.
- Modern control theory's ability to maximize or minimize virtually any parameter within reason (speed of response; control energy consumed; excursions experienced during a specified interval; smoothness of transition to a desired state; some weighted combination of these) present designers with choices – often conflicting – for optimality criteria. Resulting candidate configurations can have dramatically different degrees of stability and robustness in the presence of variations in system parameters or disturbance characteristics.

These factors are sometimes offered as explanations why modern control has often been a specialization, as opposed to modern estimation (with minimum variance as a far more widely applicable optimization criterion).

Digitization has introduced another (not universally recognized) opportunity for control. Many conventional analog controllers exhibited a loop-within-a-loop feature, exemplified by a stabilization loop for a radar antenna within a servo-driven tracker. For stability of that approach the inner (stabilization) loop had to be far more responsive than the outer (track) loop. A digital implementation, however, knows both the direction of the actual sensor sightline and the estimated direction of the tracked object. Deviation between those directions in Figure 9.3 is used to decouple the two loops, enhancing accuracy (tracker error does not contain stabilization offset) as well as stability.

Even with all data evaluated at the same instants of time, information is sometimes needed at intermediate moments. On occasion that requirement has led to overdesign with unnecessarily high data refresh rates, easily avoidable by using extrapolation or interpolation in computing control inputs from dynamic estimator output histories. That procedure is sometimes permissible even with little allowable delay (such as the aforementioned stabilization example), and usually permissible when promptness is less critical. Application to flight control will depend on the specific driving requirement, which will vary with different purposes –

- a lead pursuit or lead collision course at long range
- high dynamics at short range
- maintaining a nominal separation from aircraft in the same general area
- tight formation
- exotic flight modes (such as aircraft yaw without rolling)
- active autopilot control of otherwise aerodynamically unstable airframes
- terrain follow at low altitude with a sensor-augmented terrain map data base.

Examples just cited include operations both purely internal (a connotation often implied by the phrase *control system)* and less localized *command and control* – carrying a different connotation. The latter would apply to a control room containing provisions for communication, processing, and display, for purposes of conveying all available scenario assessment information to a command generator. Not every object in a scenario is successfully identified, and not every successfully identified object will accept guidance from the command center. To the best of the command generator's ability, though, decisions will be made – telling which participant to do what or go where, when. Timing is related to command effectiveness as limited by

- coordination / interoperability and
- nature of the data informing the guidance decisions.

The first item involves political, institutional, or military considerations which, aside from the highly emphasized interface compatibility issue, received only very limited discussion in Section III.1. For the latter item, with commands generated from instantaneous estimates for each object's state (position and derivatives), minimization of latencies and uncertainties is clearly desirable. For surveillance near a large busy airport or for any other rapidly changing dynamic scenario involving many objects concentrated in a limited area, with a deficiency of valid observational data, prompt and accurate expression of all information is the goal. Not all operations are equally demanding (*i.e.*, some require split-second timing) but all are supported by approaches used throughout this book.

III.3 SOFTWARE CONSIDERATIONS

A high-priority purpose of this section is to address the phenomenon labeled a "software crisis" – actually a misnomer – and briefly discuss some steps that could help alleviate it. As distinct from another realm (commercial software, with its own very different set of problems), software of interest here of course involves development for application to navigation and tracking.

Among the main steps – quite obvious and yet so often unobserved – would be clarification for the roles of functional design and software. As one who has

- devised both formulation and code, with results as documented in Chapter 8, and
- seen many projects governed by completely different methods and procedures,

this author can identify several points that could steer future developments toward success. Several issues, potentially affecting the logical organization of tasks into blocks of code, are now listed and followed by further discussion:

- magnitude of each task
- repeated appearance of a task in multiple places; amenability to standardization
- specific traits of the language to be used operationally
- availability of similar algorithms from past developments or public domain
- implications of a block's inputs and outputs on data flow and testability.

One of the earliest steps – compartmentalization into software "blocks" – involves some traits that are obvious and some subtleties. Among the obvious is a need to steer between extremes in length; blocks of code needed for functions of intermediate complexity provide a good compromise. Examples might include covariance matrix decrementing, or a set of postprocessing as shown in Section 9.3. Counterexamples could include operations needing only one or two lines of code (not needing separate function status) or, at the other extreme, a long and unwieldy collection of unrelated operations in unnecessary proximity.

A more subtle packaging issue is its potential dependence on the language. That can be explained by citing experience in conversion (FORTRAN to C and MATLAB) for the program producing the results in Chapter 8. It was found that exploiting traits (*e.g.*, syntax) peculiar to each individual language produced the clearest and most efficient code. Some functions used in one language underwent a thorough overhaul for conversion; a few small "worker" functions were even deleted or replaced by "one-liners" – no longer requiring separate designation – and that experience was not limited to MATLAB conversion.

Many projects are preceded by a history of developments that already produced an infrastructure of locally available functions – extensive in scope, ranging from the mundane (*e.g.*, trigonometric) to the sophisticated (*e.g.*, image processing). In that case, before any new design starts, those that are applicable can be identified and thoroughly utilized. In addition there are opportunities to take advantage of public domain software – often unfamiliar to coders inventing data manipulations, unaware of pitfalls adeptly avoided in algorithms perfected over years of refinement.

Some public domain software is available in source code (*e.g.*, LINPACK [4], EISPACK [5], and various other algorithms) or in the form of tried-and-true object libraries for a vast array of numerical and graphical manipulation. New code is not necessary for operations covered thereby. When similar-but-not-identical source code is available, a modified version of an existing function is an attractive alternative to new code. This is clearly a step in the direction of standardization, to subdue the Tower-of-Babel syndrome from multiple devices, signal characteristics, interfaces, formats, protocols, languages, operating systems, support packages, algorithm libraries, etc.

Helpful though standardization can be, imposition at the right level is crucial. For primitive elements, one set suffices for usage everywhere (self-evident as that is, there have been operational systems with multiple independent trigonometric libraries). As the level of complexity grows, a need for some flexibility emerges. At a high level, configuration control and design freezes tend to impose uniformity over time within subsystems. If taken early in an "accelerated development" (wherein stages that would normally be sequential commence in parallel) – especially when an old configuration is declared a baseline for a new system – fundamental design steps become "revisions." The accelerated schedule creates a tendency for unchecked accumulation of changes, retaining adjustments originally intended to be temporary. Revisions can eventually outweigh the original configuration in size and complexity.

The situation just depicted produces "soft"ware which, due to absence of clear understanding, becomes less flexible than hardware, clearly failing to provide the "quick start" desired. For a complex system, an older design often falls short as a "baseline" amenable to quick programming revisions by a coder. Furthermore that "baseline" may have come from a still older system (thus influenced by limitations of still earlier technology) and even more remote from present needs. The methodology is vulnerable to continued propagation of increasingly dated (thus increasingly obscure and less relevant) design criteria. A rebuild of the entire program becomes the only way to recover the software's once-vaunted flexibility.

Between the case just described and the opposite extreme (multiple sets of primitive computations) there is, fortunately, a middleground permitting extensive software reuse. With commonality of "blocks" used in several operations (*e.g.*, different nav and track modes having the same transition matrix), the opportunity is not at all surprising. As one illustration, this author has used modern estimation for purposes as diverse as integrated nav, bit synchronization, tracking of multiple maneuverable fighter aircraft plus supersonic projectiles, determination of orbits, spacecraft attitude, and flexural deformations. No single standard Kalman filter would have fit any of these purposes, but an appropriate set of modules could have supported all of them. Again the rule-of-thumb:

- the more elementary the task, the greater the potential benefit of unification (and the greater the loss without it);
- the more far-reaching the task, the more an attempted unification can become a strait-jacket.

Modularity, then, has crucial importance. Properly done it can capitalize on previous related experience – and *not* try to capitalize on previous *un*related experience.

Armed with a software repertoire gained from past activities, development based on interfaces conforming to clear specification of requirements can begin with assembly of modules and all support activity needed. Functions not covered by existing code are then supplied (possibly with predesign via pseudocode). An opportunity for parallel efforts exists during lead time for system design, unencumbered by demands to provide immediate coding instructions (and premature documentation). In one branch a properly conducted simulation can verify the concept and also generate input/output test data files. At the same time, processing resources to be used for implementing the design operationally can be undergoing selection. System designers (preferably with ability to read and debug code) can work closely with software engineers (preferably with ability to understand the functional design), to support generation of the operational program. Thorough understanding can then produce a configuration tailored to

- specific functional operations intended
- required performance in that application
- available technology appropriate for the pertinent time period.

Adequate test provisions then enable a vital final software development milestone.

III.3.1 Validation (General)

Although software development is often discussed as an issue separate from design and test, close interaction is essential. Without a blending of cultures – to the extent that designers, coders, and testers can comprehend each other's documents – the best procedures in all areas can be in vain. The *good* news about simulation and test is that a flawed algorithm cannot succeed. Algorithms that fail in simulation will likewise fail a properly conducted bench test – and therefore in the field. That information can prevent waste from premature operational effort.

At least initially, simulations and tests with limited scope can enable corrective measures by providing greater insight than field test. That will occur if test provisions include "hooks" enabling access to data of interest and a capacity for controlled execution of programs with time-freeze, plus capability to resume without changing the relative timing between all signals. That is no small order – particularly when there can be

- reluctance to provide access to intermediate data not otherwise appearing on I/O specifications,
- presence of high dynamic conditions (forces, angular rates) in scenarios of interest,
- feedback, in that some of the instantaneous input data can depend on previous outputs (example: radar signal inputs depend on antenna pointing accuracy – which depends on an algorithm's perception of a tracked object's location).

Often the full set of needs can be met only by a combination of simulations and tests complementing one another. Hardware and software modules can first be tested separately. Simulation often validates concept, but not software or hardware. *Emulation* on a host computer can validate concept plus software and / or processor firmware but not hardware. Success in the field validates all but, when problems arise (especially with immature instrumentation), it is difficult to debug conclusively. Thus another intermediate level – bench test [§] – is needed. Unprepared field operation can otherwise wreck a project's budget.

In a typical development, corrections accumulate via measures that are temporarily expedient for narrow purposes but lacking system-wide compatibility. With passage of time and changing conditions their meaning can become obscure. Loss of continuity is prevented by repetitive incremental recompiling and relinking for the changed portion – and occasionally, even a full reconfiguration. The latter step, although often resisted, can produce streamlining with flexibility for further adaptation.

[§] One example directly from this author's experience involved generation of processor input histories in accordance with a known I&Q sequence for a SAR input. In an early test the processor's failure to produce a known corresponding output was traced to the necessity for a "no-op" (literally, a no-operation) between two specific processor firmware instructions in succession. If not caught at the bench, that minuscule detail would have delayed flight success for an unknown duration.

An ever-present need for adaptability arises from changes in requirements, input characteristics, technology, equipment, and resource availability. Even for a system flawlessly providing every desirable performance feature, changes could be imposed by subsequent redefinition or extension of requirements. Strategies just noted enable coping with accumulated revisions; modular software in "packages" of moderate size (not fragmented nor unwieldy) accounting for design functionality can minimize reinvention while maintaining flexibility, control, and ease of debug.

To speed processing, programs are sometimes unfolded into lengthy and largely repetitive sections, thus avoiding overhead. Features just described then became a casualty of expediency. With sufficiently widespread observance of standards, modern capabilities fortunately can render that step unnecessary. There can even be an overabundance of computing resources to accommodate future expansion of processing load. In that one area it is becoming progressively less expensive to expect the unexpected.

III.3.2 Validation (A Specific Example)

In addressing one particular validating operation – integrity test with high confidence – several issues raised previously here and in Chapter 6 come together. With repeated test trials for random events (in this case, integrity alarms or missed detections), high confidence can come only from a significant number of occurrences – and the event statistics differ substantially from what intuition would suggest [6]. Since both alarms and missed detections must be low-probability events, a significant number of them translates into a very high number of test trials. Those tests could be performed, especially with today's processing capabilities – but to this day no requirement exists for such a high number of end-to-end "*rf* in, event-count out" test trials.

Reason for concern about the absence just identified is offered in [7], previously cited with discussion at the opening of Chapter 6. In complete opposition to high confidence of integrity tests verifying proper performance, [7] inescapably shows confidence to be highly inadequate. Discussion below aims toward solution.

With subsystems compartmentalized as described near the end of Section 10.3, end-to-end test can be subdivided into "*rf* in, measurement errors out" and "measurement error in, event-count out" stages. The first, more sophisticated operationally, is less demanding in terms of repetition. The second, requiring many run trials to provide confidence, needs a less demanding test setup – easily compatible with probability scaling, mentioned in Section 6,4 and briefly summarized here.

A measurement error ϵ at $4.9\,\sigma$ would produce 10^{-6} alarm probability. With a 10% increase in σ, ϵ – now normalized at $4.9/1.1$ – produces almost nine times as many alarms. This sensitivity, quite inconvenient operationally, offers a benefit (lower run trial count) for test. This and many additional validation issues are described in [8] which advocated rigorously demanding blind test of actual code (never pseudocode, when high confidence is required), with input and output sequences in full correspondence to operational conditions.

III.4 COORDINATION

The rapid pace of scientific breakthroughs greatly complicates the effort to stay current while planning, designing, building, and maintaining systems – that is universally understood. Nevertheless, much difficulty encountered in meeting the integration challenge can be overcome by exploiting untapped flexibility offered with digitization. For the industry to capitalize on its latent promise, system design plans must reflect expectations of recurring, sweeping changes in capability. Adaptability to those changes can then be designed into system concepts and architecture. Methodology used throughout these chapters is entirely consistent with that purpose, with design, coding, and test fully coordinated.

With the need for integration allowed to transcend other agendas, the name *system integrator* will acquire a literal meaning, with clear connotation of a much deeper technical involvement.

References

[1] Class, K.A., Hartman, R.G., and Helgeson, B.J., "Differential GPS landing assistance system," U.S. Patent Appl. No. 07/970,320, Nov, 1992.
[2] Farrell, J.L., *Integrated Aircraft Navigation*, Academic Press, 1976. (Now in paperback)
[3] van Sickle, G.A., "GPS for Military Surveillance," *GPS World*, Nov. 1996.
[4] http://www.netlib.org/linpack
[5] http://www.netlib.org/eispack
[6] Farrell, J.L., and vanGraas, F., "Statistical validaton for GPS integrity test," *ION Journal*, Summer 1992, *pp*. 203-215; also *ION Redbook of Selected Papers* vV 1998, *pp* 89-100.
[7] Nisner, P.D. and Johannessen, R., "Ten million data points from TSO-approved GPS receivers: results of analysis and applications to design and use in aviation," *ION Journal*, Spring 2000, *pp*. 43-50.
[8] Farrell, J.L., and vanGraas, F., "Integrity testing for GNSS Sole Means," ION Nat'l Tech Meeting, 1994.

Subject Index

For assembly of index topics it is common practice to omit entries that appear most frequently in the text. In many respects that practice is followed here, especially for words that could appear on practically any page (*e.g.*, GPS, GNSS, pseudorange, position, velocity, acceleration, navigation, initiation / initialization), plus other common terms such as tracking, aiding, strapdown, short-term, validation, test, integration – and familiar terms connected with

- motion (dynamics, translation, angular rate, rotation, roll, pitch, attitude, altitude, latitude, longitude, gravity, cardinal directions)
- a "low-cost" modifier for terms such as IMU, INS, and the like; the preponderance of inertial material involves high-drift instruments.
- estimation, *e.g.*, linear(iz•••), state, model, dimension, sensitivity, performance, accuracy, error, average, variance, covariance, RMS, random, bias, measurement, observation, observable, optimal, residual, processing, weighting, recursion, update, integration, *a priori*, *a posteriori*, etc.

The Table of Contents is more likely to provide useful direction for the ubiquitous terms, which could otherwise distract from the main purpose of the index (*i.e.*, to locate discussion involving items that are less prominent). The index does, however, include some expressions for these terms appearing as part of a phrase.

Certain entries (*e.g.*, "transition matrix") can occur in verbal form or as a symbol – or in both forms on the same page – or in expanded form within an equation with or without its symbol notation or its verbal expression; Section numbers given in APPENDIX I can then help to locate additional occurrences. Partly for that reason, pages in APPENDIX I are not included in the index. Likewise omitted from index generation are terms used in bibliographical titles at the end of each chapter.

Another circumstance applies here: Items appearing in APPENDIX III only – and nowhere else – are also omitted from the index, for a different reason. Those topics are peripheral; of interest to only a subset of readers. Many could be distracted by addition of so many side issues. Those wishing to pursue APPENDIX III will be able to follow the side discussions without help from table lookup.

When a term appears literally in text – but with a meaning other than that implied (*e.g.*, "settling" on *p.* 24; "heading" on *p.* 105, and several others) – the page number is intentionally omitted. The converse is also true; an effort was made to consider context without requiring all entries to have exact terminology.

Some topics have subheadings and a few have another lower-level subheading. Page numbers for subheadings are repeated with headings in some situations (*e.g.*, when both the specific and the more general context appear), but this is not always the case.

1pps (1-pulse-per-second), 135-138, 140, 141

A

accelerometer (See inertial, measurement, accelerometer)
acquisition (See also pull-in), 23, 24, 86, 135, 148, 190, 199, 212
A/D (analog-to-digital) conversion, 136
adaptability (/ flexibility), 1, 4, 28, 71, 97, 100, 114, 117, 119, 188, 190-192, 206, 226, 259, 266-271
air data, 55
  airspeed, 1, 54, 246
  angle of attack, 1, 54, 183, 246
alarm probability, 111, 126, 270
alert limit, 108, 127, 202
alignment (mechanical), 54, 145, 220-221
  instrument mounting (See misalignment)
along-range, 17, 18, 190, 191, 195, 202, 212, 213, 222
alpha-beta tracker, 197, 199, 212
alpha-beta-gamma tracker, 197, 199, 212
altitude channel, 7, 19, 42, 51, 53, 57, 249
antenna, 22, 69, 75, 102, 104, 155, 162, 166, 170, 191, 201, 215, 222, 227, 262, 264, 265, 269
  electronically steered array, 192, 199, 221, 223, 262, 264
    conformal, 220, 221, 262
ascent rate/descent rate, 7, 226
  of satellite elevation, 88, 89
aspect angle, 205
association (file), 191, 218, 221, 223, 264
asynchronous observations, 69, 189
attitude error (See misorientation)
attitude pseudomeasurements, 54, 87, 96, 178, 182
autocorrelation, 26, 58, 59, 242
autonomous, 54, 55, 109, 151, 152, 160, 207, 212, 246
avionics, 137, 262

B

backup, 121, 136, 226, 227, 230, 239
ballistic coefficient, 212-214
ballistic object tracking, 186, 209-214
band limiting, 58, 66, 149, 192, 205, 240-242, 244, 251, 257
bandwidth, 62, 136, 146, 149, 192, 199, 261, 263
baseline, 23, 72, 75, 119
  configuration, 267, 268
beamforming, 220, 221
beamwidth, 193, 215, 262
bearing, 201, 205
  azimuth bearing, 215-217
  relative bearing, 205
  true bearing, 205
Beidou 1, 21, 67
bench test, 226, 246, 269
bistatic/multistatic tracking, 26, 186, 206, 227
bleed-off (of "calibration" data), 245, 246, 257
block estimation, 1, 10, 15, 16, 80, 81, 84, 87, 91-97, 152, 154, 196
bookkeeping, 87, 199, 204, 240, 247, 251

C

calibration, 38, 39, 65, 85, 99, 150, 186, 204, 240, 245, 246
IMU, 40, 56-66, 135-142
in-flight "calibration", 36, 38, 39, 58, 59, 76, 77, 82, 83, 85, 99, 101, 240, 241, 245, 246, 257
laboratory pre-calibration, 36, 38, 39, 61, 66, 99
thermal, 65
radar or optical sensors, 186, 204
carrier landing, 228
carrier phase, 1-5, 21, 40, 67, 69, 82-84, 86, 88, 96-98, 126, 128, 263
discontinuity, 51, 78, 85, 86, 96
sequential change, 2, 5, 13, 22- 24, 29, 69, 72-96, 101-106, 122, 125, 144-146, 151, 154, 155, 158, 160, 164, 167, 170, 171, 177-180, 182
SV motion within subtraction interval, 87, 101-105, 144
centripetal acceleration, 192, 213
certification, 230
cg (center of gravity), 144
chi-squared, 109, 128
clock, 88, 136-138, 230, 262, 263
user, 5, 21, 69, 70, 78, 87, 112, 113, 115, 117, 119, 128, 145, 155, 189
satellite, 21, 68, 72, 111-114, 116, 120, 189
closest approach, 2, 192, 218, 219, 228
clustering, 203
of position solutions, 112, 116
of tracked targets, 91, 203, 215
clutter (See ground return, radar)
coast (free-inertial), 28, 29, 40, 42, 57, 60, 83, 92, 98, 99, 150, 239-258
collision avoidance, 2, 187, 202, 203, 218, 219, 228
compensation, 23, 28, 40, 42, 88, 127-150, 187, 245
GNSS errors (See also clock, satellite or propagation, ionosphere and propagation, troposphere)
IMU (See calibration, IMU, laboratory pre-calibration)
motion compensation (See radar, SAR, motion compensation)
radar or optical sensor error (See calibration, radar or optical sensors)
time lag, 136, 261
computational synchronization (See synchronization, computational)
confidence coefficient, 127
conflict resolution, 203
coning, 5, 28, 29, 35, 36, 40-42, 65, 66, 135, 147-150
conservatism (See also overbounding), 5, 24, 26, 52, 53, 56, 66, 76, 83, 88, 89, 108, 118, 121, 125, 126, 165, 180, 181, 198, 218
constellations, 4
mixed constellations, 4, 142, 144, 145
containment, 60, 244, 246
CONUS (Continental U.S.), 4
cooperative engagement, 185, 191, 203-205, 209, 220, 221, 227
coordinate frame, 196, 202
earth, 30, 31, 33, 39, 69, 204, 206
ECEF (earth-centered,earth-fixed), 22, 69, 103-105, 145, 156, 209, 210, 216, 217
geographic, 42, 55, 156, 200, 205, 210, 216, 241
inertial (nonrotating), 19, 20, 33, 34, 36, 38, 209
navigation (wander azimuth), 4, 5, 20, 30-36, 38, 39, 42, 46, 54, 55, 58, 64, 68, 69, 73, 75, 77, 87, 99-104, 192, 193, 196, 204, 212, 220, 241, 242
platform (apparent nav), 30, 39, 46, 75
rotation of, 3, 20, 26, 30-33, 35, 38, 39, 54, 58, 75, 77, 87, 99, 100, 155, 192, 212, 242
vehicle or IMU reference, 20, 30, 34, 35, 38, 47, 54, 55, 65, 73, 98, 99, 148, 193, 201, 220

correlation, 15, 17, 26, 62, 63, 65-67, 70, 71, 79, 80, 83, 84, 97, 129, 132, 222, 228, 242
coefficient, 70, 80
cross-axis (geometric), 41, 42, 187, 196, 198, 222, 227, 228
measurement, 26, 87, 90, 91, 106, 118, 155, 161
due to across-SV differencing, 5, 67, 79, 71, 78, 79, 84, 87, 113, 115-118, 155
sequential, 42, 71, 73, 79-85, 87, 90-97, 100, 129, 155, 161, 228
correlator, 228
cost / benefit tradeoff, 225-230
coupling, 62, 65, 101, 124, 202
deep integration, 146, 151, 182
kinematical (See kinematics, kinematical coupling)
loose, 1, 146
tight, 3, 77, 86, 96, 98, 120, 146, 151, 165, 230
ultratight, 96, 146, 151, 165
covariance decrement, 15, 16, 91, 202, 212, 267
covariance propagation (See propagation, covariance propagation)
cross-axis
IMU (See misalignment, inertial instrument or nonorthogonality)
tracker states (coupling or decoupling), 195, 196, 202-204, 214
crosshairs, 193, 194, 201, 203, 204
cross-range, 17, 186, 195, 202, 203, 212, 213
CSC (carrier smoothed code), 77, 86
cue, 153, 154, 185, 191, 216, 227
cursor, 218, 227
curvature radius, 31, 32, 39, 156, 215, 216
cycle slip / loss of cycle count, 3, 78

D

data acquisition, 135, 136, 148
data collection, 104, 137, 154, 162
data editing, 95, 96, 107-109, 123-125, 127, 160, 171, 173, 178, 180
data fitting, 24, 76, 83, 97, 100, 183, 207, 221
data rate, 1, 28, 29, 41, 42, 48, 58, 84, 136, 137, 140-143, 147, 170, 181, 190, 199, 205, 223, 261-265
data span (data averaging duration), 2, 10, 16, 25, 29, 45, 50, 52, 56, 59, 76-79, 82-85, 96, 97, 99-101, 135, 154, 180, 190, 208, 214, 216, 219
data block, 16, 80, 81, 84, 87, 91-97, 156-158
data window, 2, 25, 74, 96, 97, 100, 126, 154, 156, 158, 159, 172, 178, 183, 202
datum reference, 19, 187, 189
deceleration, 18, 212
deep integration (See coupling, deep integration)
deformation (structural), 220, 221, 268
delay (system time lag; See also propagation), 21, 36, 77, 86, 125, 137, 139, 141, 142, 148, 199, 205, 206, 227, 230, 261, 262, 264-266
delayed-state, 51, 101
depression angle, 205, 217
detectable bias, 111, 114, 118, 120, 124, 126
detection statistic, 110, 111, 117, 124, 133
detection threshold, 109, 111, 119, 124
differencing, 95, 111, 115-120, 123, 131, 132, 144-146, 180
across receivers, 23, 72, 75, 77, 86, 145, 155, 191, 203, 260
across satellites, 5, 67, 69, 70, 72, 73, 77-79, 84, 86-88, 93, 96, 102, 105, 115, 119, 125, 128, 129, 145, 155, 165, 171, 172, 178, 181, 189
correlations from (See correlation, measurement, due to across-SV differencing and correlation, sequential)

double, 71-73, 77, 115, 187, 189, 205
in time (See carrier phase, sequential change)
triple, 73, 75, 77, 92, 170
ATD (adjusted triple difference), 92, 93
differential GNSS / differential operation, 72, 75, 77, 101, 109, 152, 170, 186-190, 228, 229, 260
dispersion, 112, 116
of position solutions, 109, 112, 116
resolution of tracked targets, 191, 203
disturbance, 96, 265
gust, 154
perturbation (See orbit determination, perturbation)
vibration, 42, 43, 56, 58, 59, 63, 64, 66, 149, 154, 160, 171, 178, 222, 246
divergence, 125
DOP (dilution of precision), 68, 71, 112-114, 119, 189, 190
double differencing (See differencing, double)
drag, 183, 211-214
drift angle, 151, 162, 163, 166, 168, 171, 172, 175
drift (rotational), 20, 26, 29, 38, 40, 42, 47, 49-53, 56, 58, 59, 64, 74, 77, 78, 82-84, 92, 93, 97-101, 147, 149, 170, 178, 240-242, 245, 246, 253, 256-258
DSP (digital signal processor), 136, 148

E

eight-state filter (position, velocity, user clock), 69
EKF (extended Kalman filter), 14
electro-optical (See optics)
elevation angle, 88-90, 191, 209, 215
ellipsoid
earth ellipsoid, 31, 39, 145, 156, 195, 202, 215-217
evolute, 215, 216
error ellipsoid, 193, 195, 202, 212
error suppression / cancellation (See also compensation), 23, 28, 40, 42, 58, 88, 118, 125, 126, 136, 145, 147-150, 186, 218, 221, 222, 227, 230, 261
exclusion (See integrity, identification/exclusion)
external trigger (See sampling rate, external trigger control)
extrapolation, 76, 78, 129, 156, 209, 211-213, 218, 223, 261, 262, 265

F

false observability (See observability, false observability)
FDE (fault detection/exclusion) (See integrity, identification/exclusion)
FDI (fault detection/identification) (See integrity, identification/exclusion)
federated system, 227
FFT, 146, 192, 228
field of regard, 203, 211
field of view, 203, 261
final value, 104, 205
flight test, 2, 20, 50, 60, 151, 152, 154-161, 171-181, 246
fusion (registration of raster images), 191, 204, 227
future capability, 4, 6, 21, 29, 79, 83, 84, 87, 92, 95, 97, 145, 182, 187, 225-227, 229, 239, 240, 246, 260, 266, 270

G

GALILEO, 1, 4, 21, 67, 142, 144, 145
garble, 143
GDOP (geometric dilution of precision) (See DOP)

geocenter/geocentric, 34, 69, 75, 105, 207, 209, 210, 215, 216
geoid, 215, 217
geometry, 5, 65, 68, 86, 112, 117, 119, 120, 125, 126, 152, 187-190, 209, 246
GIS (geographic information system), 218
GLONASS, 1, 21, 67, 144
GMT (Greenwich Mean Time) (See time bases, GMT)
graceful degradation (See backup)
gravitation, 19, 22, 34, 57, 145
gravity anomalies, 46
ground return (sensor response), 191, 192
  radar (clutter), 191
ground station / ground receiver (See receiver, ground station / ground receiver)
GTRF (Galileo Terrestrial Reference Frame), 145
gyro (See inertial, measurement, gyro)

H

HDOP (horizontal dilution of precision) (See DOP)
heading
  off true North, 4, 30, 54, 98, 108, 121, 151, 157, 158, 162, 163, 171, 175, 178, 181-183, 189, 201, 205
  magnetic heading, 1, 178, 246
holding pattern, 60, 240, 244, 245, 252
hyperbolic nav (See also LORAN), 226

I

image registration, 191
IMU (See inertial, measuring unit)
increment, 48, 55, 78, 87, 121, 125, 204, 207
  incrementing task list, 5, 36, 38, 39, 136
  rotational, 1, 5, 28, 30, 34-36, 38, 40, 41, 43, 45, 55, 67, 147-150, 204, 220
  time, 1, 12, 24, 32, 33, 35, 36, 38, 39, 45, 85, 137-142, 155, 156, 263
  translational, 1, 3, 5, 12, 23, 28, 32, 33, 36, 38-41, 45, 48, 51, 62, 67, 73, 101, 137, 150, 161
inertial, 1-6, 19, 26, 35-45, 66, 87, 96-98, 108, 136, 147, 148, 182, 206-209, 214, 220, 225
  acceleration (absolute), 19
  angular rate (absolute), 19, 20, 33, 35, 45, 56, 66, 148, 149, 242, 256
  instrument error, 2, 5, 20, 29, 36, 40-42, 45-47, 49, 51, 56-58, 76-78, 82, 96-101, 190, 222, 257
    flicker noise, quantization, random walk, 41, 58, 147, 149, 246
    motion-sensitive, 29, 40, 56-66, 78, 85, 86, 96, 99, 148, 149, 222, 244, 246
  measurement, 4, 19, 142
    accelerometer, 1, 19, 27-29, 34, 36, 38-40, 42, 45-47, 49, 51-54, 56-59, 61-63, 65, 67, 76-78,82-86, 93, 97, 98, 100, 101, 135-137, 150, 151, 240-242, 245, 249, 253-258
    gyro, 1, 19, 27-29, 35, 36, 38, 40, 41, 45, 50, 52-65, 67, 76, 77, 83-87, 97-99, 135, 147-151, 170, 178, 182, 204, 240-242, 245, 249, 254-258, 262
  measuring unit (IMU), 1, 2, 20, 28, 29, 38, 40-42, 45, 49, 54-57, 60-62, 73, 75-77, 82, 83, 85, 98, 99, 101, 135-138, 140, 142, 144, 146, 151, 152, 154, 155, 160, 170, 171, 181, 214, 220-222, 226, 227, 229, 239, 262
    specifications, 53, 56, 60, 61, 65, 76, 83, 85, 99, 170, 182, 239, 244
  navigation system (INS), 20, 21, 25, 28, 36, 46, 54, 60, 77, 78, 81-86, 92, 95, 96, 98, 107, 120-122, 124, 146, 152, 162, 171, 182, 190, 193, 205, 207, 220, 221, 225-227, 230, 258
INS (See inertial, navigation system)
institutional considerations, 108, 127, 228, 230, 259, 260, 266
integer ambiguity, 3, 23, 24, 67, 77, 78, 86, 100, 144, 145
integrity (See also RAIM), 3, 5, 129, 131, 133, 134, 187, 246, 270
  detection, 6, 108-111, 113, 115, 116, 118-120, 125-127, 132

identification/exclusion, 6, 109, 113, 116, 118-120, 125, 127, 131, 132
multiple flawed signals, 5, 6, 109, 113, 114, 119, 186, 206
parity, 5, 107-127, 132, 133, 187
snapshot, 5, 120, 121, 125
interface, 6, 28, 42, 107, 151, 192, 220, 226, 227, 259, 261, 262, 266-268
intermeasurement interval, 26, 75, 76, 87, 89-94, 98, 101, 126, 198, 199, 211
interoperability / compatibility, 4, 107, 144, 145, 229, 266
interpolation, 6, 138-143, 189, 262, 265
data samples, 6
interrogation, 143, 263
ionospheric propagation (See propagation, ionosphere)
isodop, 191, 222
ITRF (International Terrestrial Reference Frame), 145

K

Keplerian, 207-212
kinematics (of classical motion analysis; not implying carrier phase), 39, 99, 100
kinematical coupling, 43, 45, 62, 65, 187, 196, 202, 204

L

Lambert's theorem, 209, 210
latency, 136, 137, 139, 142, 205, 227, 230, 261, 262, 266
lever arm, 6, 43, 62, 67, 73, 74, 78, 87, 92, 104, 105, 142, 144, 155, 157, 162, 171, 178, 181, 182, 220
littoral environment tracking, 186, 214-217
LORAN, 40, 119
LOS (radar or optical sensor line of sight), 185, 186, 191-194, 201, 203, 213, 214, 221
LOS rate, 192, 205, 213, 218, 219
LSB (Least Significant Bit), 136, 143, 203

M

maneuver, 4, 25, 151, 169, 181, 190, 196, 199, 211, 216, 219, 245, 246, 268
speed change, 29, 85, 105, 151, 152, 158, 166, 172, 174, 175, 183
turn, 29, 60, 85, 96, 98, 99, 105, 151-154, 157, 158, 160, 162, 163, 166, 169, 172-175, 178, 183, 203, 240, 243-245, 252
mask, 78, 88, 126
matrix factorization, 95, 125, 129
QR, 109-111, 114, 116, 118, 120, 122, 128-130, 187
UD, 78, 87, 90, 91, 122
MEMS (microelectromechanical sensor), 28, 42, 56, 65, 225
misalignment (cross-axis mounting offset), 58-61, 85, 87, 96, 98, 99, 245, 252
inertial instrument, 38, 40, 41, 53, 58-61, 63-65, 85, 87, 96, 98, 99, 147, 170, 216, 240, 242, 245, 249, 252
misorientation, 4, 20, 26, 30, 36, 38, 42, 43, 45-47, 49-53, 56, 74, 77, 87, 98, 99, 101, 178, 182, 221, 241, 242, 256
[co]variance adjustment, 53, 87
leveling error (tilt), 5, 26, 36, 43, 45, 49-51, 53, 57, 58-60, 77, 78, 82, 83, 92, 93, 96-98, 100, 151, 162, 163, 166, 169-172, 175, 179-181, 220, 240, 241, 244, 245, 249, 252, 253, 256
pseudomeasurements (See attitude pseudomeasurements)
missed detection
in integrity testing, 111, 126, 127, 270
radar, 223
mixed-constellation observations, (See constellation, mixed constellation)
Mode S, 143, 185, 187
modularity, 240, 262, 268, 270

motion-sensitive error (See inertial, instrument error, motion-sensitive)
multi-constellation (See constellation, mixed constellation)
multistatic tracking (See bistatic / multistatic tracking)
mutual surveillance (See surveillance, mutual surveillance)

N

narrowband, 43, 61, 62, 64, 66, 204, 212, 263
nonlinearity, 13-15, 22, 26, 65, 69, 146, 182, 192, 193, 207, 211, 229, 261, 265
nonorthogonality, 39, 65, 242, 244, 245

O

observability, 54, 78, 90, 96, 99, 100, 114, 151, 152, 171, 178, 202, 203, 212, 214
  false observability, 202, 212
observer, 197-199, 204
optics (electro-optical or infrared sensing), 185, 186, 188, 190, 191, 193, 204, 214, 220, 221, 223, 228, 264
orbit determination, 6, 24, 161, 162, 182, 183, 186, 207, 208, 211, 268
  perturbation, 207, 211
overbounding, 24, 26, 89, 108, 111, 126
overdetermined, 110, 111, 114, 117, 120, 125, 132, 187
ownship, 25, 144, 188-190, 193, 196, 204-206, 216

P

parabolic blending, 137-140
parity (See integrity)
passband, 148
passive ranging, 201, 202, 218, 227
PDOP (position dilution of precision) (See DOP)
pixel, 191
position estimator (3-state, uncoupled), 46, 52, 67, 77, 79, 86, 101, 122, 125, 145, 153, 154, 178, 190
position pseudomeasurements, 146, 227, 228
preprocessing, 1, 29, 40-42, 135, 149, 191, 192, 228
PRF (See radar, PRF)
probability scaling, 127, 270
process noise, 15, 25, 26, 53, 56, 58, 74, 76, 78, 79, 83, 84, 90, 91, 96, 97, 101, 121, 126, 128, 129, 247, 253, 254
projectile tracking, 214, 268
propagation, 73, 77, 83, 87, 92, 105, 126, 144, 176, 189, 197, 207, 263
  covariance propagation, 15, 16, 52, 56, 76, 78, 196, 211, 212, 240, 253
  error propagation / eror growth, 2, 4, 5, 20, 28, 29, 40, 41, 45-47, 49-51, 57-60, 69, 78, 84, 93, 96-98, 147-149, 161, 182, 205, 211, 223, 239-258
  ionosphere, 21, 70, 72, 88, 89, 155
  troposphere, 21, 70, 72, 88-90, 155
protection level, 125, 127
pseudomeasurements, 96, 146
  attitude (See attitude pseudomeasurements)
  position (See position pseudomeasurements)
  velocity (See velocity pseudomeasurements)
pull-in, 24, 25, 86

Q

quadrature, 62-65, 147, 222
quality indicator, 187

quasistatic, 25, 51, 57, 58, 66, 74, 97, 154, 155, 157
quaternion (4-parameter attitude representation), 30, 34-36, 38, 42, 203
QZSS, 1, 21, 144

R

radar (See also optics), 17, 18, 24, 185, 186, 188, 190-194, 198, 204, 206, 209, 211, 212, 214-217, 222, 223, 228, 230, 261, 262, 265, 269
clutter (See ground return)
PRF (pulse repetition frequency), 192
SAR (synthetic aperture radar), 85, 181, 218, 220, 222, 230, 269
ISLR (integrated side lobe ratio), 222
motion compensation, 218, 222, 227, 230
PSLR (peak side lobe ratio), 222
RAIM (Receiver Autonomous Integrity Monitoring), 109-120, 125, 131
single-observation, 5, 109, 121-126, 113, 133
range (distance), 17, 18, 21, 88, 185, 190-196, 201-204, 212-219, 221-223, 265
range-only, 201, 202
range rate, 17, 24, 192, 205, 218, 219
training flight (See training flight)
raw data, 1, 4, 5, 27, 28, 36-40, 42, 67, 77, 108, 143, 145, 151, 185, 187, 189, 206, 226-230
reacquisition (See reinitiation)
receiver, 21, 22, 69, 75, 101, 102-104, 107-109, 119, 146, 160, 171, 192, 206, 214, 218, 225-227, 260
ground station/receiver, 23, 71-73, 75, 77, 86, 87, 91, 101, 119, 164, 166, 167, 170, 186-188, 260
for orbit tracking, 207, 209
track loop, 85, 96, 146, 151, 165, 182, 265
rectification, 40-43, 58-60, 63, 64, 149
recursion, recursive estimation, 10, 11, 13, 16, 72, 75, 76, 84, 92, 96, 97, 152
reentry, 209-213
refraction, 186, 209, 214-216
reinitiation, 23, 45, 69, 83, 86, 107, 121, 122, 125, 137, 141, 211
relative angular rate, 31-33, 38, 39, 75
relative motion (translational), 9, 13, 20, 26, 104, 187, 190, 192, 196, 203-205, 212, 213, 218, 219, 229
reprocessing, 97, 125, 135, 141, 142
residual monitoring, 126
response, 27, 41, 60, 63, 84, 85, 97, 143, 146-149, 182, 191, 192, 196, 198, 215, 218, 222, 240, 257, 262, 264, 265
risk analysis, 127
robustness, 3-5, 22-24, 32, 42, 51, 67, 69, 84, 226, 265
runway, 179, 190, 243, 260
runway monitoring, 228
RVCC (Range-Velocity Cartesian Coordinates), 212

S

sampling rate, 1, 6, 28, 29, 41, 42, 45, 83, 136, 137, 147, 161, 199
sample timing, sampling interval, 75, 76, 138, 140, 141, 262
external trigger control, 135-137
scale factor (inertial instrument), 36, 38, 39, 53, 58-63, 65, 85, 87, 96, 98, 99, 136, 240, 242, 244, 245
scenario, 20, 56, 59, 143, 147, 186, 188, 189, 191, 192, 196, 214, 219, 223, 230, 239, 240, 242-245, 247, 250-252, 256, 257, 263, 264, 266, 269
Schuler, 5, 29, 46, 49, 50, 59, 60, 82, 84, 97-99, 248, 251, 253, 256
sculling, 5, 28, 29, 36, 38, 40-43, 148, 150
segmented estimation, 2, 3, 4, 9, 45, 48, 50-52, 56, 67, 76-79, 81, 84, 86, 87, 91, 92, 96, 97, 100, 101, 122, 144, 145, 154
settling (See also steady state), 96, 196, 198, 199, 212

shipboard electronics, 137, 262
sidereal, 33, 39, 54, 58, 102, 104, 145, 156, 209, 242, 248
sightline, 264, 265
of tracked object, 9, 189, 193-196, 200, 202, 203, 212-214, 217, 261, 264, 265, 269
between SV and user's receiver, 119, 144, 157, 188, 189
radar or optical sensor-to-target (See LOS (radar or optical sensor line of sight))
simulation, 41, 50, 64, 148, 149, 151-154, 216, 217, 226, 245, 258, 268, 269
situation awareness, 143, 189, 192, 205, 225, 230
sparse data, 151, 152, 154, 187
specific force (nongravitational acceleration), 19, 20, 34, 43, 45, 46, 49, 54, 56, 60, 62-64, 66, 75, 78, 87, 214, 241, 250
spectral density, 16, 52, 56, 58, 76, 84, 121, 161, 196, 254
squitter (extended), 143, 144, 185, 187, 189-192, 203, 206, 263
stability, 24, 30, 57, 86, 121, 146, 165, 264, 265
stabilization, 204, 226, 227, 261-265
of azimuth, (See wander angle / wander azimuth)
of computation, 39
of estimator, 15, 26
state selection, 49, 51, 56, 87, 92
steady state, 5, 171, 172, 178, 198, 222
transient preceding, 3, 25, 54, 56, 59, 78, 96, 105, 154, 160, 163-165, 170-172, 178-180, 191, 195, 196, 198, 199, 211, 212
sub-wavelength, 3, 23, 77, 83, 84, 97, 146, 218, 222
surveillance, 199, 203, 205, 206, 214, 227, 228, 230, 263, 264, 266
mutual surveillance, 185-187, 189, 191, 203, 205, 206
swath, 191, 222
synchronization, 6, 135-137, 142, 189, 220, 230, 263, 268
computational, 72, 135, 137-142,, 220, 230, 262
system integration, 83, 98, 108, 226, 227, 271

T

TCAS (Traffic Alert and Collision Avoidance System), 202, 203
TDOP (time dilution of precision) (See DOP)
TEC (total electron count), 88, 172
temperature, thermal properties, 56, 61, 65, 136
testability, 226, 266
three-state position estimator (See position estimator (3-state, uncoupled))
tilt (See misorientation, leveling error)
time bases, 4, 21, 135-137, 145, 263, 264
GMT (Greenwich Mean Time), 4
UTC (Universal Time Coordinated), 143, 189
time constant, 101, 149, 198, 199, 205, 240-242, 245, 246, 248, 251, 253, 255-258
time shift, 143
time slot, 143, 223
time tag / time stamp, 93, 107, 121, 136, 143, 205, 227, 230, 261, 262, 264
time-to-go, 218, 219
track file retention / maintenance, 49, 91, 112, 190, 222, 223
training flight, 228
transfer alignment, 204, 220-222, 227
transit time, 21-23, 192, 223
transition matrix, 3, 8, 9, 11, 17, 18, 25, 47, 49, 50, 52, 75, 76, 78, 87, 92, 99, 100, 128, 196, 200, 205, 208, 209, 211, 214, 247, 253, 268
cascaded, 45, 47, 75, 76, 87, 92, 93, 209
translator, 186, 206, 214

transmitting station, 1, 67
transponder, 143, 191
triangular matrix, 67, 71, 72, 78, 87, 95, 100, 110, 122, 123, 130
tropospheric propagation (See propagation, troposphere)
tuning (Kalman filter), 25, 29, 96, 180
TWS (Track-While-Scan), 216

U

ultratight coupling (See coupling, ultratight)
unobservable (See observability)
UTC (See time bases, UTC)

V

van test, 73, 84, 142, 151, 162-171, 181
variance multiplier, 25, 165
VDOP (vertical dilution of precision) (See DOP)
velocity pseudomeasurements, 146, 227
vertical deflections, 46, 59, 85, 240, 242, 245, 246, 256-258
verticality, 4, 49, 50, 54, 55, 60, 61, 84, 100, 170, 216, 256
vetronics, 137, 262
vibration (See disturbance, vibration)
vibration-sensitive error (See inertial, instrument error, motion-sensitive)
VOR (VHF Omni[directional] Range), 240
VTS (Vessel Tracking Systems), 228

W

wander angle / wander azimuth, 30-33, 36, 39, 42
WGS (World Geodetic Survey), 4, 31, 145, 210
wind, 54, 55, 175, 183, 213, 246, 260
WLS (weighted least squares), 92-96, 110, 124
word length, 147, 226, 227, 262
world view, 4, 145, 187, 189

Author Index

Adams, J.A., 150
Anoll, R.K., 44, 66, 258
Ashley, H., 224
Axelrad, P., 6, 106, 134
Babu, K., 134
Barbour, N., 44, 66
Bass, C.A., 106
Bayliss, E., 150, 224
Bechhoefer, E.B., 224
Bierman, G.J., 6, 78, 87, 90, 91, 94, 106, 122, 134, 152, 184
Bisplinghoff, R., 224
Blackman, S.B., 224
Blakelock, J.H., 184
Blanchard, R.C., 224
Bortz, J.E., 27, 44, 148, 150
Brenner, M., 134
Brooks, W.R., 134
Brown, R.G., 6, 26, 66, 106, 134
Bryan, R.S., 184
Bryson, A.E., 6, 106
Buehler, K., 231
Chin, G.Y., 134
Class, K.A., 40, 64, 204, 271
Conley, R., 134
Daum, F., 224
Enge, P., 6, 106
Erickson, G.W., 66
Etkin, B., 184
Farrell, J.L., 6, 26, 44, 66, 106, 134, 150, 184, 224, 231, 258, 271
Feng, Y., 184, 224
Fitzgerald, R.J., 224
Fried, W.R., 258
Gonzales, R.C., 224
Goodman, L.E., 150
Grewal, M.S., 150, 231
Gunawardena, S., 184
Gupta, S.K., 106
Hartman, R.G., 271
Hedman, E.L., 106
Hegarty, C., 6, 106
Hein, G., 150
Helgeson, B.J, 271.
Ho, Y., 6, 106
Hoech, R.W., 150
Hwang, P., 6, 26, 66, 106
Ignagni, M.B., 148, 150
Jazwinski, A., 6, 106
Johannessen, R., 134, 271
Jordan, J.W., 150
Kalafus, R.M., 134, 224, 231
Kalman, R.E., 1, 6, 14, 26, 32, 36, 38, 39, 41, 42, 53, 57, 77, 78, 90-92, 96-100, 122, 124, 126, 152, 154, 155, 180, 187, 198, 199, 207, 219, 224, 229, 240, 256-258, 268
Kaplan, E., 6, 106
Karmokolias, C., 106
Kayton, M., 258
Khatri, A., 106
Kiran, S., 150
Kline, P., 134
Kovach, K., 134
Kovach, J.P., 134
Lancaster, E.R., 224
Laning, J.H., 27, 44, 148, 150
Lee, S., 106, 150, 184
Lee, Y., 134
Lemay, J.L., 66
Leondes, C.T., 224
Liang, A.C., 66
Mansfield, R.L., 224
Mark, J.G., 27, 44, 148, 150
Marti, D., 134
Martin, E.H., 106
Mathews, B.D., 224
McConkey, E.D., 44, 66, 150, 224, 231, 258
McGraw, G.A., 134
Miller, R.B., 27, 44, 148, 150
Mims, J.H., 224
Misra, P., 6, 106
Morrison, N., 6, 19, 26, 81, 82, 84, 106
Neregard, F., 231
Newland, D.E., 150, 224
Nisner, P.D., 134, 271
Ober, P.B., 134
Parkinson, B.W, 6, 106, 134
Pealer, N.A., 224, 231
Quesinberry, E.C., 224
Reasenberg, R., 106, 184
Reid, L., 184
Robbins, G.M., 134
Robinson, A.R., 150
Rogers, D.F., 150

Saks, S.L.., 44, 231
Savage, P., 27, 41-44, 148, 150
Schifferns, E., 258
Schlegel, L.B., 134
Schmidt, G., 44, 66
Shin, D., 106
Skolnik, M.I., 224
Soloviev, A., 184, 258
Son, S., 106
Spilker, J.J., 6, 106
Spitzer, C.R., 184
Stephens, C.G., 150, 224, 231
Stoddart, L., 134
Sturza, M.A., 134
Sullivan, J., 150
Sung, T., 106
Tazartes, D.A, 44, 148, 150.
Teunissen, P.J., 134
Trommer, G.F., 106, 184
Unger, F., 66
van Dierendonck, A.J., 224, 231
van Graas, F., 44, 66, 106, 134, 150, 184, 258, 271
van Dyke, K., 134
van Sickle, G.A., 150, 271
Wade, M., 150, 231
Wendel, J., 106, 184
Wiles, G., 224
Wintz, P., 224
Young, R.S., 134
Zarchan, P., 224
Zhou, N., 184, 224