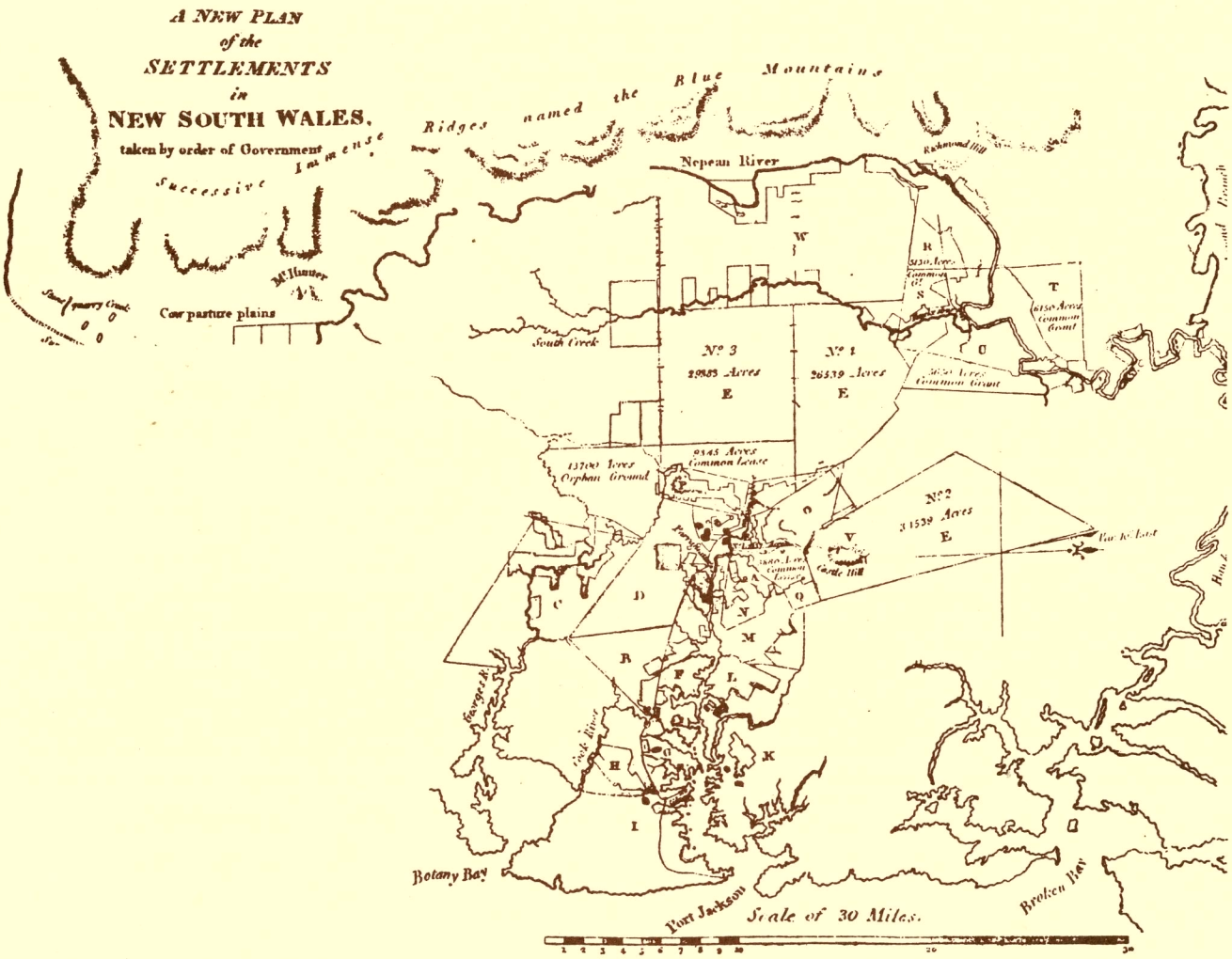
THE USE OF KALMAN FILTERS IN GPS NAVIGATION

BERTRAND MERMINOD



UNISURV S-35, 1989
Reports from
SCHOOL OF SURVEYING

THE UNIVERSITY OF NEW SOUTH WALES KENSINGTON, NSW AUSTRALIA 2033

UNISURV REPORT S - 35, 1989

THE USE OF KALMAN FILTERS IN GPS NAVIGATION

Bertrand Merminod

Received: April, 1989

SCHOOL OF SURVEYING
UNIVERSITY OF NEW SOUTH WALES
P.O. BOX 1,
KENSINGTON, N.S.W. 2033
AUSTRALIA

National Library of Australia

Card No. and ISBN 0 85839 053 0

ABSTRACT

The theory of mathematical filters has mainly been developed by statisticians and electrical engineers. The description of the filter type most commonly used presently for scientific and engineering applications is by KALMAN (1960). It took about ten years for the importance of those principles to be recognised and applied by the surveying community, but at present Kalman filtering is widely used for positioning, especially for hydrographic surveys. Kalman filtering is treated here as an extension of the least squares estimation methods traditionally used by surveyors, but where a kinematic model is added to the measurement model. The kinematic models relevant for navigation using the Global Positioning System (GPS) are classified systematically, and the importance of Kalman filtering in different aspects of GPS data processing is outlined. Refinements such as smoothing and the capability of including constraints on the adjusted parameters are discussed, and new algorithms are proposed for this purpose.

The software package DYNAMO, recently developed at the School of Surveying, University of New South Wales, incorporates Kalman filter and optimal smoother algorithms and is capable of accomodating different choices of kinematic models for relative kinematic positioning using GPS. As well as coordinate and receiver clock parameters for any number of stationary and moving receivers, one bias parameter per satellite can be included in the solution. In such a set-up, the differential corrections in either the measurement or solution domain can be computed and applied, by merely changing the filter "settings". Indeed, these two approaches are particular applications of the more general implementation that allows for simultaneous filtering in both domains, using a single filter. This flexibility sets DYNAMO apart from other differential kinematic programs. Examples highlight the dangers inherent in the filtering process, if the models are improperly defined.

DYNAMO has been tested on a number of GPS datasets. Differential GPS positioning results using pseudo-range and phase-rate data collected in various experiments, and processed using different model options, are discussed. The improvement in the quality obtained by filtering is demonstrated, and the high reliability of filtering and smoothing techniques, as compared with other data processing techniques, is shown.

TABLE OF CONTENTS

ABSTRACT ii
LIST OF ACRONYMS viii
NOTATION ix
ACKNOWLEDGEMENTS x

1. INTRODUCTION 1

1.1. SURVEYORS, SATELLITES AND FILTERS 1

1.2. PREDICTABLE MOTION 3
1.2.1. Static and Kinematic Modes 3
1.2.2. Role of the Kinematic Model 4
1.2.3. Types of Kinematic Models 5

1.3. POSITIONING USING GPS 6
1.3.1. Essential Features of the System 6
1.3.2. Types of GPS Observables 6
1.3.3. Accuracy of GPS Observables 9
1.3.4. Modelling Requirements 11

1.4. DIFFERENTIAL GPS PROCEDURES 14
1.4.1. Data Type 14
1.4.2. Processing Strategy 16
1.4.3. Estimation Algorithm 18
1.4.4. Classes of Differential Kinematic GPS Procedures 19

1.5. FILTERING AND GPS 20
1.5.1. Expected Benefits 20
1.5.2. Possible Applications 21
1.5.3. Approach Selected 25

1.6. SYNOPSIS 28

2. LEAST SQUARES PROCEDURES 29

2.1. LEAST SQUARES IN GEODESY 29

2.2. METHODOLOGY 30
2.2.1. The Two Models 30
2.2.2. The Criterion 31
2.2.3. The Parameters 32

2.2.4. Redundancy 32
2.3. CLASSIFICATION OF LEAST SQUARES PROCEDURES 33
2.3.1. From Static to Kinematic Estimation 33
2.3.2. Why Filtering ? 35
2.4. PRINCIPAL FORMULATIONS OF CLASSICAL LEAST SQUARES PROCEDURES 36
2.4.1. Condition Method 36
2.4.2. Parametric Method 38
2.4.3. Relations Between Both Approaches 39
2.4.4. The Combined Case 40
2.4.5. Decomposition in Least Squares 41
2.4.6. Bayesian Least Squares 43
2.4.7. Sequential Least Squares 45

3. PREDICTION, FILTERING AND SMOOTHING 48

3.1. PRINCIPLES 48
3.2. PREDICTION 51
3.2.1. Initial State Vector 51
3.2.2. Kinematic Model 51
3.2.3. Parametric Expression 52
3.3. FILTERING 53
3.3.1. Measurement Model 53
3.3.2. Basic Least Squares Filter 54
3.3.3. The Kalman Form 56
3.4. FILTER TYPES AND TERMINOLOGY 58
3.4.1. Single Filters 58
3.4.2. Combination of Filters 59
3.5. SMOOTHING 63
3.5.1. Initial State 63
3.5.2. Kinematic Model 64
3.5.3. Least Squares Filter / Smoother: Statement 64
3.5.4. Least Squares Filter / Smoother: Solution 66
3.5.5. Extension to More Epochs, R-T-S Algorithm 68
3.5.6. Modified Bryson-Frazier Algorithm 70
3.5.7. MB-F Rearranged 71

3.5.8. The Adjoint State Vector 72
3.6. NEW SMOOTHING ALGORITHMS 73
3.6.1. Using Filtered Residuals 73
3.6.2. Using Smoothed Residuals 74
3.6.3. Storage Requirements 76

4. NAVIGATION MODELS 78

4.1. RANDOM SEQUENCES AND FUNCTIONS 78
4.1.1. Traditional Assumptions 78
4.1.2. Useful Random Functions 80
4.1.3. Examples of Random Processes 83
4.2. KINEMATIC MODEL 85
4.2.1. Functional and Stochastic Components 85
4.2.2 Classification of Kinematic Models 87
4.2.3 Transition and System Noise Matrices 92
4.2.4 The GPS Receiver Clock 96
4.2.5. Reference Coordinate System 98
4.3. MEASUREMENT MODEL 99
4.3.1. Time and GPS 99
4.3.2. Modelling the Pseudo-range 101
4.3.3. Modelling the Phase-rate 103
4.4. DIFFERENTIAL GPS 105
4.4.1. Principle and Application 105
4.4.2. Extension to Filtering 107
4.5. CONSTRAINTS 109
4.5.1. Principle and Application 109
4.5.2. Distance 110
4.5.3. Azimuth and Heading 112
4.5.4. Vertical Angle 114

5. THE SOFTWARE PACKAGE "DYNAMO" 116

5.1. STRUCTURE AND FEATURES 116
5.1.1. The Software Components 116
5.1.2. User Interface 118
5.1.3. Maintenance 119

5.2. MAIN PROGRAM 119
5.2.1. Characteristics 119
5.2.2. The Filter 121
5.2.3. The Smoother 122
5.2.4. Comparison of Filter and Smoother 123
5.3. AUXILIARY PROGRAMS 125
5.3.1. Prepare the Measurement File 125
5.3.2. Define the Estimation Procedure 126
5.3.3. Control the Printer 127
5.3.4. Select Output 128
5.3.5. Test the Coordinates 128

6. TESTING FILTER PERFORMANCE 129

6.1. THE ROLE OF SIMULATIONS 129
6.2. GPS SIMULATION CAPABILITY 130
6.3. THE SYDNEY-MANLY FERRY 133
6.3.1. Assumptions 133
6.3.2. Results and Comments 135

7. PROCESSING OF REAL DATA 139

7.1. SYDNEY AIRPORT EXPERIMENT 139
7.1.1. Data Collection 139
7.1.2. Data Processing 140
7.1.3. Survey Marks 142
7.1.4. Trajectory 145
7.1.5. Velocity 149
7.2. RIG EXPERIMENT 150
7.2.1. Campaign Design and Data Collection 150
7.2.2. Processing with PoPS 151
7.2.3. Asynchronous Measurements 153
7.2.4. Results Using DYNAMO 155
7.3. OFFSHORE AZIMUTH EXPERIMENT 157
7.3.1. Campaign Design and Data Collection 157
7.3.2. Data Preparation 159
7.3.3. Differential Corrections 162

7.3.4. State Vector and Kinematic Models 166
7.3.5. Position 167
7.3.6. Azimuth 169
7.3.7. Design of a Specific Estimation Procedure 173
7.3.8. Future Investigations 175

8. CONCLUSIONS 177

REFERENCES 180

APPENDIX A: DECOMPOSITION IN LEAST SQUARES 184
A.1. The Combined Case 184
A.2. Adjustment with Constrained Parameters 185

APPENDIX B: BAYESIAN LEAST SQUARES 187

APPENDIX C: LEAST SQUARES FILTER 189
C.1. The Bayes Form 189
C.2. The Kalman Form 190
C.2.1. State Vector Estimate 190
C.2.2. Covariance Matrix of the Filtered Estimate 191
C.2.3. Residuals on the Observations 192
C.2.4. Another Expression for the Gain Matrix 192

APPENDIX D: LEAST SQUARES FILTER / SMOOTHER 193
D.1. State Vector Estimates 193
D.2. Covariance Matrix of the Smoothed Estimate 194

APPENDIX E: SMOOTHING ALGORITHMS 196
E.1. Observation Free Expression for the Smoothed Estimate 196
E.2. From Predicted to Filtered Residuals 197
E.3. From Filtered to Smoothed Residuals 198
E.4. Equivalence of the New Algorithm with R-T-S 199

APPENDIX F: KINEMATIC MODELS 201
F.1. Continuous Form 201
F.2. Transition Matrix 201
F.3. Covariance Matrix of the System Noise 202
F.4. Random Walk as a Special Case 203

LIST OF ACRONYMS

BDOP	Bias Dilution of Precision
DGPS	Differential GPS
ECEF	Earth Centred - Earth Fixed
EDM	Electronic Distance Measurement
FIL	Filtered
F+C	Filtered and Constrained
GDOP	Geometric Dilution of Precision
GPS	Global Positioning System
GPST	GPS Time
INS	Inertial Navigation System
ISS	Inertial Survey System
NAV	Navigation Solution (unfiltered)
PDOP	Position Dilution of Precision
PH	Phase(-rate)
PR	Pseudo-Range
PRN	Pseudo Random Noise
rms	root mean square
rss	root sum square
SLR	Satellite Laser Ranging
SMO	Smoothed
S+C	Smoothed and Constrained
SRB	Satellite Range Bias
WGS	World Geodetic System

NOTATION

As far as possible, the notation is in line with that used in most textbooks on least squares procedures. However, due to the larger number of estimation steps in filter/smoother techniques than in classical least squares, a clear distinction between various quantities, such as the "residuals", is therefore necessary. This is accomplished by a consistent use of centred superscripts, for both parameters and residuals, to denote the stage of the computations.

The following **rules** for style apply throughout the text:

General:	lowercase normal:	s	scalar
	lowercase bold:	$\mathbf{s}$	vector
	uppercase bold:	$\mathbf{S}$	matrix

Some important **symbols** used consistently throughout the text are:

l	vector of measurements
$\mathbf{x}$	state vector, that is, vector of parameters
$\mathbf{x}_0$	initial or previous state vector (with epoch subscript)
$\delta\mathbf{x}$	increment of the state vector
$\mathbf{v}$	vector of residuals
$\overset{\circ}{\mathbf{x}}, \overset{\circ}{\mathbf{v}}$	approximate state vector, respective residuals
$\tilde{\mathbf{x}}, \tilde{\mathbf{v}}$	predicted state vector, respective residuals
$\hat{\mathbf{x}}, \hat{\mathbf{v}}$	adjusted (filtered) state vector, respective residuals
$\hat{\hat{\mathbf{x}}}, \hat{\hat{\mathbf{v}}}$	smoothed state vector, respective residuals
$\mathbf{Q}_{ll}$	covariance matrix of the measurements
$\mathbf{P}$	weight matrix of the measurements (subscript l is omitted)
$\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}$	covariance matrix of the predicted state vector
$\mathbf{P}_{\tilde{\mathbf{x}}}$	weight matrix of the predicted state vector
$\mathbf{Q}_{l\hat{\mathbf{x}}}$	covariance between the measurements and the filtered state vector
$\mathbf{A}$	design matrix for parametric adjustment
$\mathbf{B}$	design matrix for condition adjustment
Φ	transition matrix
$\mathbf{w}$	vector of misclosures or system noise (depending on context)
$\mathbf{K}$	Kalman gain matrix

ACKNOWLEDGEMENTS

Firstly, I wish to thank my supervisor, Dr Chris Rizos for his help and guidance. Not only did we have numerous and fruitful discussions, but he also did some of the hard work at the computer, adapting and enhancing some utilities such as for orbit interpolation and plotting. His criticisms of my work were always constructive. In addition to that, I owe him an improvement in my English expression.

It was great to join a strong GPS research group at UNSW. For such a topic, it is best to be where "something is going on", and have access to a seemingly unlimited flow of information. Special thanks go to my fellow student Don Grant for the many discussions which are responsible for my better understanding of the strength of various GPS geometries. The incredible ability of Bernie Hirsch to talk even with the most stubborn computer has helped me on several occasions, and his readiness to make his talent available to others deserves particular mention.

What actually made this research possible was a scholarship from the Swiss Science Foundation. I am most grateful for this opportunity which allowed me to dedicate all my efforts to this project. I hope this thesis justifies the confidence they had in me.

Several practical experiments were conducted during this research, and many organisations provided financial or logistic support. I wish to thank the Royal Australian Survey Corps for their assistance during the "Sydney Airport" kinematic experiment, and Prof. Günter Hein for kindly making available the results of the data processing at the University of the Federal Armed Forces (Germany). Wild Leitz (Australia) Pty. Ltd. and Rod Eckels in particular assisted in the organisation of some experiments. The Commonwealth Scientific and Industrial Research Organisation, Division of Oceanography, provided facilities on its research vessel, the R/V FRANKLIN, for the test of the Constrained Kalman Filter and the University of Tasmania, School of Surveying, helped with the terrestrial measurements. I would also like to thank the American Institute of Navigation, Satellite Division, who sent me an encouraging invitation for their annual meeting in Colorado Springs.

My deepest thanks are for my wife Anne. From the start she was enthusiastic about coming to the other side of the world, and courageously overcame the problem of living with a researcher. So much so that she gave birth to our third son Nicolas during this thesis project. She certainly joins me to thank all the persons -- and they are definitely too numerous to be named -- who helped us have an enjoyable stay "Down Under".

à mes parents ...

1. INTRODUCTION

1.1. SURVEYORS, SATELLITES AND FILTERS

The Global Positioning System (GPS) has emerged as a revolutionary satellite range measurement and positioning technology, and its impact on the surveying profession is already considerable. More importantly, the advent of this new technique reflects the fundamental changes that our profession has undergone in the last few decades, which are twofold:

- The times of Heinrich Wild are over and surveyors no longer design their own instruments.

- As instruments become more complex, their designers strive to make their operation easier, and nowadays collecting accurate measurements does not require skilled operators. Using a GPS receiver in the field allows one the time to ponder this implication!

In short, the survey operation is losing its mystique. The decline in the image of the surveying profession is particularly obvious in Australia. Whereas the early surveyors played a leading role in the exploration and opening up of the country, they now form a tightly structured and rather conservative corporation of professionals. However, the intrusion of other professionals, skilled and unskilled, into the positioning professions is in line with the modernisation of our society, and the situation is no different in the case of many other technical professions. Trying to set back the clock by protecting a specific part of our trade, for example cadastral surveying or geodesy, is a negative attitude. More reasonable is the acceptance of the challenge offered by the new technology: foster our strengths and search for new applications of our measurement and analysis skills.

GPS has the potential to irrevocably change the surveying profession. While it is tempting to restrict the application of GPS to geodetic problems, it is preferable to widen our expertise to areas that could benefit from a strong background in estimation theory. One such application is the positioning of moving objects, also called "kinematic" positioning. In sea navigation for example, the variety of instrumentation available has increased dramatically in recent times. Generally, the integration of these sensors and the interpretation

of the results still relies on sub-optimal procedures, such as the arbitrary selection of results from a preferred or "master" sensor, combined with graphic extra- or interpolation. At the School of Surveying, University of New South Wales (UNSW), the first contribution in this field was by CHISHOLM (1987) with a study of the integration of GPS into hydrographic survey operations. Although the navigator's experience is invaluable, the complexity of modern navigation systems is now such that an optimal combination of data from different instrument sensors requires elaborate computational support. Furthermore, although an electronical engineer is required to design GPS receiver hardware, a surveyor is certainly more familiar with the processing of heterogeneous data than a traditional navigator.

The aim of this thesis is therefore to present the estimation techniques required for kinematic positioning as an extension of the traditional geodetic adjustment methods. The emphasis is on presenting the concepts, an overview and discussion of existing procedures, as well as the development of some new algorithms. Topics from "signal processing" (traditionally associated with electrical engineering) are also included. Of course, mathematical expressions are necessary to illustrate the words, and a special effort has been made to use consistent notation.

Whilst GPS has been the main motivation for this work, because this technology presents definite advantages for kinematic positioning, it is merely the sensor systematically used to illustrate the concepts and apply the estimation techniques presented. Very little background information regarding GPS is presented here, and only characteristics of the system relevant to the design of the estimation algorithms are discussed. These will mainly be problems of time and clocks, rarely encountered in traditional surveying. If the reader is not acquainted with GPS, an excellent background can be found in previous publications of the School of Surveying, UNSW, for example KING et al (1987), ECKELS (1987), CHISHOLM (1987) and GRANT (1988).

The present work is intended for the user of GPS, not the orbit determination specialist. The position of the GPS satellites is taken as error free at all times, and orbital adjustment is not attempted. The motion of interest in this thesis is that of the receiver(s).

1.2. PREDICTABLE MOTION

1.2.1. Static and Kinematic Modes

In the purely random kinematic mode, each determination of the position is independent of the others. Assuming that all other factors influencing the positioning system are constant, the precision of a "fix" is also constant. On the other hand, in the static mode, each new measurement contributes to the determination of the same position parameters. Extending an observation session makes more data available for the estimation of the same number of parameters, hence improving their precision. This behaviour is illustrated in Figure 1.1.

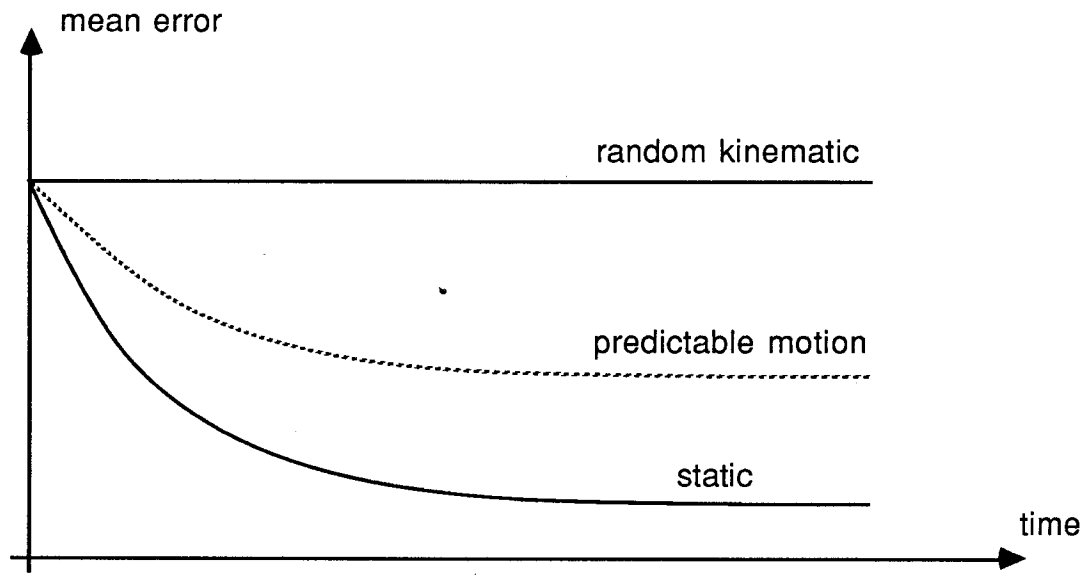


Figure 1.1: Precision of Positioning under Various Conditions

However, the difference in the behaviour of the static and kinematic solutions is not due to the movement itself, but to the *a priori* knowledge of the movement. If we could predict the movement of a vehicle as well as that of a survey mark, the positions at the different measurement epochs could be perfectly related to one another and the precision obtainable would approach that of the static case. This applies, for example, to an object moving along a known trajectory with known velocity. Therefore, it is the **predictability** of the object's motion that permits the link between static mode (surveying) and well-behaved kinematic mode (most navigation) to be established. In addition, this predictability of the motion makes it an ideal candidate for filter estimation

methods. The resultant improvement in the quality of the position determination can be represented by the dashed line in Figure 1.1.

1.2.2. Role of the Kinematic Model

Basically, the kinematic model is the mathematical expression for the predictability of the motion. This predictability is another way of saying that the position parameters (the coordinates) are not entirely random, but have values that are, within certain bounds, related to their values at an earlier epoch. This can be illustrated by considering a vehicle travelling along a straight line. After a couple of positions have been determined, future positions can be predicted by extrapolation. Of course, the uncertainty associated with this prediction grows with time. Whilst a position predicted a few seconds ahead may be more accurate than a determination based solely on new measurements, this will certainly not be the case if positions are predicted over some hours. Therefore, the kinematic model comprises two components:

- the **functional** part -- the prediction of a position based on previous results,

- the **stochastic** part -- the estimation of the precision associated with the predicted position, depending on influences not, or imperfectly, accounted for in the functional part.

Thus, the outcome of the kinematic model takes the form of a direct measurement of some or all of the position parameters, with associated precision estimates. The aim of a filter is to optimally combine these "pseudo-measurements" with new actual measurements, or in other words, combine the position predicted via the kinematic model with a new position determination derived from subsequent measurements.

Often the underlying cause of the system movement is not important, only the resultant trajectory. In these circumstances, it is therefore more appropriate to refer to a **kinematic** process rather than a **dynamic** process. Systems in which parameters change with time will be referred to as kinematic processes or systems, and the relation between parameter estimates at successive epochs will be described by a kinematic model. Unfortunately, the term "dynamic model" is often (mis-)used to describe this relation, even though the

model may not explicitly involve the forces acting on the system and causing the motion.

1.2.3. Types of Kinematic Models

There are two positioning modes involving motion:

(1) positions are required only at selected **discrete locations**,

(2) a continuous **trajectory** is desired. Generally, discrete positions are also computed, but their location on the trajectory is arbitrary, as long as they are distributed in such a manner that reliable inter- or extrapolation is possible.

In the first case, the path between the locations is irrelevant, and no modelling of the trajectory is therefore attempted. For example, the integration of the accelerations measured by (perfect) inertial sensors (INS) along different paths between two stations yields the same difference in coordinates. Indeed, the link between the two locations is established exclusively by the measurements. This positioning mode is gaining in importance with the advent of GPS, and many new terms are used to designate it: intermittent positioning, semi-kinematic positioning, "stop and go positioning", etc. More details will be given in Section 1.5.

Two classes of approaches for modelling a trajectory can be distinguished, whether emphasis is:

(a) on **functional** modelling -- whereby functions of time are associated with the trajectory. Instead of positions, parameters of the functions are estimated from which positions at any time can be derived. Polynomials are commonly used but there are many other possibilities. For example, a sinusoidal function can be defined by estimating its amplitude, period and initial phase.

(b) on **stochastic** modelling -- whereby the uncertainty in extrapolated positions is defined by a suitable "noise" affecting the motion. No predefined analytical form of the trajectory is assumed.

In fact, both approaches involve a functional and a stochastic model. In functional modelling, the stochastic properties of a predicted position are derived from the precision of the estimated parameters (obtained, for example, from previous measurements). For stochastic modelling, a position for which the assumed uncertainty applies must be computed as a function of previous position estimates.

The choice of stochastic modelling as the preferred technique for this research will be justified in Section 2.3.

1.3. POSITIONING USING GPS

1.3.1. Essential Features of the System

GPS range measurements are made using coded time signals transmitted by the satellites. Each GPS satellite transmits a unique signal on two L-band frequencies: L1 at 1575 MHz and L2 at 1227 MHz (equivalent to wavelengths of approximately 19 and 24cm respectively). The satellite signals consist of the L-band carrier waves modulated with a "Standard" or S code (also referred to as Coarse Acquisition or C/A code), a "Precise" or P code and a Navigation Message containing, amongst other things, the coordinates of the satellites as functions of time: the so-called "Broadcast Ephemeris". The S code is intended mainly for civilian use and yields a range measurement precision of about 10m. Once the GPS satellite constellation is fully implemented, the system may be intentionally degraded, according to the **selective availability** policy of the U.S. Department of Defense. However, it is expected that a position precision of 100m will be guaranteed (see, for example, TALBOT, 1988). The P code is intended for military and selected civilian use only, and yields a measurement precision of about 1m.

1.3.2. Types of GPS Observables

Three principal types of observables can be used for GPS positioning:

- The **pseudo-range** -- This measurement is made with the aid of the pseudo-random noise codes modulated on the L-band carriers. The code of most interest to civilian users is the S code (frequency of 1.023 MHz) modulated on the L1 carrier. The code allows the time of

transmission of a signal to be unambiguously determined by the receiver, and the offset with respect to the reception time measured by the receiver gives a measure of the transit time. The range is obtained by scaling the transit time into distance by the speed of light. However, the local receiver time is not perfectly synchronised with the time kept by the satellites (and the transmitted code), and the range derived from the transit time measurement is known as "pseudo-range" because it contains the receiver clock offset from GPS satellite time (a constant for all ranges measured simultaneously to all tracked satellites). The term "code measurement" is also frequently used for the pseudo-range.

- the Doppler count is the difference in the number of cycles of the carrier wave between the received signal and a locally generated signal with the same nominal frequency, as measured over a short time interval (usually less than 1sec). The **phase-rate** is the ratio of the Doppler count and the time interval over which it is accumulated.

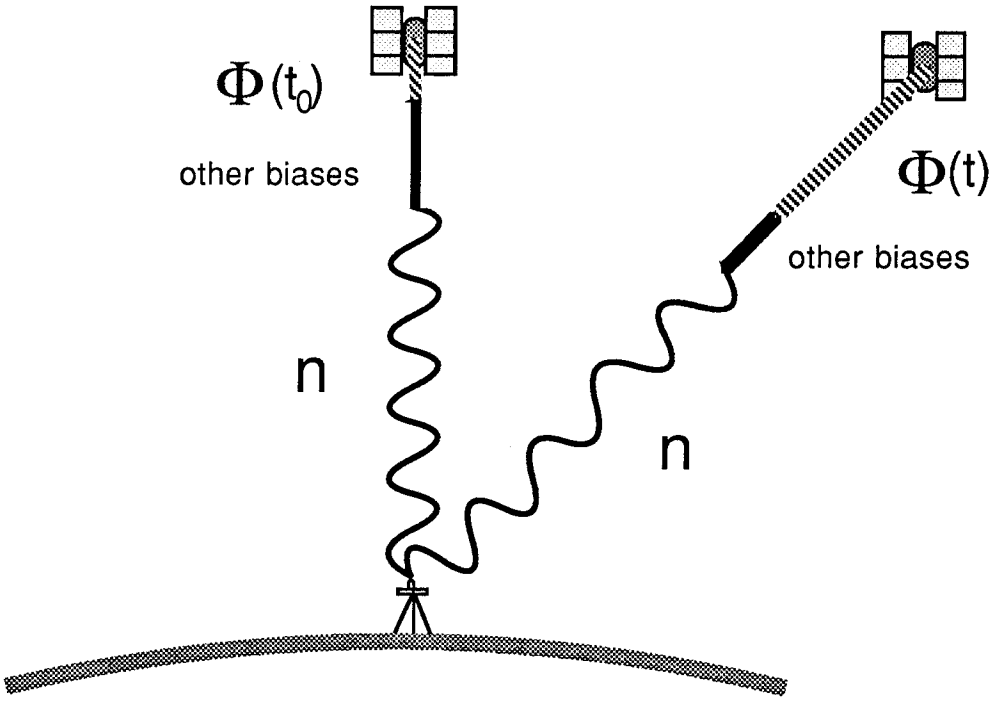


Figure 1.2: Integrated Carrier Phase and Cycle Ambiguity

- the **carrier phase** (or carrier beat phase, or integrated Doppler count) is the difference in phase of the carrier between the received signal and the locally generated signal, measured continuously over a longer period (5min to 3hrs). The full number of cycles is accumulated, but the range is

ambiguous: as the carrier phase measurement is only possible within one wavelength, the initial number of whole wavelengths in the satellite - receiver range, the so-called (integer) cycle ambiguity, is not known. Therefore, the carrier phase gives information about the change in range, and not the range itself. This is illustrated in Figure 1.2, where $\Phi(t)$ is the carrier phase measured at time t, t_0 is the initial epoch and n is the cycle ambiguity term.

A more detailed description of the GPS observables is given in ECKELS (1987) and GRANT (1988). The modelling of GPS observables is affected by a number of biases. The source and the mathematical nature of these biases is briefly discussed here in order to highlight the constraints that they may impact on observation modelling. Three classes of GPS biases can be identified.

1) **Satellite biases**; common to all observables involving that satellite at a particular epoch:

   - satellite clock offset -- caused by instabilities of the onboard oscillator,
   - satellite position error -- due to erroneous orbit parameters in the Broadcast Ephemeris,
   - degradation of the satellite operation -- either accidental or intentional.

2) **Site biases**; common to all observables involving that site at a particular epoch:

   - receiver clock offset -- caused by instabilities of the local receiver oscillator,
   - site position error -- coordinates fixed at erroneous values.

3) **Range biases**; the most difficult to model as they are only present in observables involving a particular satellite - site pair:

   - atmospheric delay -- caused by the ionosphere and the troposphere,
   - multipath -- signal incoming by an indirect path.

These error sources do not affect the different GPS observables in the same manner. For the carrier phase only, two types of range biases must also be considered:

- the cycle ambiguity -- constant over a session,
- cycle slips -- due to occasional losses of lock on the satellite.

A large number of parameters would be required to fully model all these biases, or alternatively, special processing procedures need to be developed to eliminate or minimise some of them (see GRANT et al, 1989).

1.3.3. Accuracy of GPS Observables

Table 1.1 shows the effect of various error sources on pseudo-ranges. Many of these biases are systematic errors, and most of their sources are associated with the satellites. For short GPS receiver separations, some range biases, such as those due to atmospheric delay, equally affect the measurements recorded at two close sites. Therefore, the difference between a pair of site positions is significantly more accurate than a single site position. This is the fundamental principle underlying Differential GPS (DGPS) techniques. With one receiving antenna placed at a known location, it is possible to compute coordinate or range corrections valid for any other site in the vicinity. DGPS is the standard method used to reduce the effect of unmodelled error sources and can be used with any type of measurement. For longer baselines however, refinements in the mathematical modelling of different biases may become essential, as shown in the two right columns in Table 1.1.

The carrier phase is affected by the same sources of systematic errors, though not always in the same manner, and the benefits of DGPS also apply. However, the discussion of the respective characteristics of the pseudo-range and carrier phase observables is more involved.

The **pseudo-range** is strongly affected by multipath, especially in a highly reflective environment. EVANS (1986) reports that errors caused by multipath can increase from about 1m (rms) in a desert-type location to over 4m (rms) for a rooftop or shipboard installation. On the other hand, losses of lock do not reduce the geometric strength of previous or subsequent measurements. The

noise level in the measurement is a limitation on the precision at any time, but an unambiguous range can be derived whenever the coded message is received, without respect to other measurements. This makes the position information gained through pseudo-ranges reliable, unless the satellite constellation geometry is poor or the system is degraded.

Table 1.1: User Equivalent Range Errors (UERE). (from CHEZELLES, 1988)

SEGMENT SOURCE	ERROR SOURCE	ERROR S-CODE (1σ) m	ERROR P-CODE (1σ) m	DIFF. MODE P-CODE	
				NEAR	FAR
SPACE	Clock & Nav Sub-system Stability	3.4	3.4	0	0
	Predictability of Space Vehicle Perturbations	1.0	1.0	0	0
	Other	0.5	0.5	0	0
CONTROL	Ephemeris Prediction Model Implementation	4.2	4.2	0	0
	Other	0.5	0.5	0	0
USER	Ionospheric Delay Compensation	5.0 - 10	2.3	0	2.3
	Tropospheric Delay Compensation	2.0	2.0	0	2.0
	Receiver Noise and Resolution	7.5	1.5	2.1	2.1
	Multipath	1.2	1.2	1.7	1.7
	Other	0.5	0.5	0.5	0.5
1σ UERE		10.8-13.9	6.6	2.7	4.1

The **carrier phase** observation is far less subject to multipath and can be measured much more precisely, to a few mm (rms), but is only useful if accumulated over a certain time interval, so that the cycle ambiguity can be estimated reliably. This problem, together with proposals for the determining of appropriate durations of the tracking sessions, is discussed in MERMINOD (1988) and the efficiency of these proposals is demonstrated in GRANT

(1988). Losses of lock on the satellite do weaken the geometric significance of all the data collected, unless the number of cycles lost can be recovered by other means. However, this is only possible when the observables can be modelled to a precision better than a wavelength of the carrier, 19 and 24cm for the frequencies L1 and L2 respectively.

The problem is therefore: **how to take advantage of the precision of the carrier phase measurement yet minimising the problems of cycle slips and ambiguity?** The benefits of combining pseudo-range and carrier phase measurements have long been recognised, for example by JORGENSEN (1980), and HATCH (1982) proposed an algorithm to use the synergistic properties of both types of observables. Basically, the pseudo-ranges provide a rough, but reliable position, whilst the noise on the pseudo-ranges is smoothed using the more stable carrier phases. The computed positions may not be more accurate than pure pseudo-range solutions, but the relation between positions at adjacent epochs is markedly improved. Over the years, many refinements to this procedure have been proposed, for example KLEUSBERG (1986), ABIDIN (1988), LOMMIS et al (1988). Obviously, incorporating phase and phase-rate data can improve the quality of position and velocity determinations, but there is still a lot of development effort required before such techniques can be used for reliable real-time operations.

1.3.4. Modelling Requirements

Modelling is central to any estimation process, and the choice of a data reduction technique is mainly influenced by the modelling complexity that can be afforded. An increase in the quality of GPS positioning results can be achieved in three principal ways:

1) **Add other types of measurements** -- This could be any range, position, velocity, heading or pitch and roll information, from a variety of sensors such as Inertial Navigation Systems (INS), gyro compass, ship's log, shore-based range or phase measurements, etc. However, as pointed out in Section 1.3.1, GPS provides several types of measurements, and all these types should be used before considering augmenting GPS with data from an other navigation sensor.

2) **Improve the modelling of the measurements** -- The modelling of systematic errors affecting GPS measurements is rather difficult, but DGPS allows for a dramatic simplification, as most biases can be accounted for without requiring any mathematical expression. For example, the difference in simultaneous measurements involving one satellite and two close sites is modelled, rather than each measurement separately. For long baselines however, refinements in the modelling of atmospheric delays and orbits may become necessary.

3) **Improve the modelling of the parameters** -- Some information about the parameters can be obtained from sources other than the measurements. Depending on the nature of this information, it can be expressed as constraints on the parameters (for example when two points maintain the same height at all times), and/or via a kinematic model describing the behaviour of the parameters over time.

Very often several of these options must be resorted to. Some attempts have been made to classify positioning processes according to their accuracy. As the emphasis here is on concepts, a strict classification is not proposed, but two extreme, albeit classical, applications are considered:

(a) Open ocean navigation; for the required accuracy of about 100m, instantaneous single point positions using pseudo-ranges are sufficient, and the carrier phase measurement need not be considered at all.

(b) Geodetic positioning; for the centimetre accuracy desired, the pseudo-ranges cannot make any significant contribution, except to produce initial approximate positions for the phase adjustment. The modelling accuracy required for cycle ambiguity resolution and cycle slip recovery, however, entails two major limitations in the use of the carrier phase :

- The differential positioning mode is necessary for an accurate measurement model,

- The static receiver mode is almost inevitable because it is the only "motion" that can be both easily and accurately modelled (using time-invariant position parameters).

These modelling constraints explain why the carrier phase has only played a central role in GPS surveying, for which purpose special hardware, software and processing strategies have been developed. On the other hand, the operation of GPS as a navigation aid is very straightforward and largely automatic. To summarise, the classical surveying and navigation applications of GPS can be described in terms of 3 attributes, as shown in Figure 1.3. These attributes are not necessarily exclusive. As discussed in Section 1.2.1, there is a continuous graduation between random kinematic and static positioning modes. Furthermore, a single point can be regarded as a degenerate network. Finally, the use of the carrier phase does not preclude that of the pseudo-range.

SURVEYING	Static	Network	Carrier Phase
NAVIGATION	Kinematic	Single Point	Pseudo-range

Figure 1.3: Attributes of Surveying and Navigation Applications of GPS

For all applications that require an accuracy somewhere between (a) and (b), such as for harbour navigation and geophysical exploration, neither of these two well defined methods is suitable and an optimum integration of hardware, special software and appropriate processing strategies is required in order to satisfy the particular accuracy, reliability and operational demands. If we focus our attention on those positioning applications requiring accuracies in the range from the submetre to dekametre level -- whether involving a moving GPS receiver or not -- we may pose the following questions:

- Can we satisfy the accuracy requirements using a single GPS receiver? Or do we require other receivers?
- Can we satisfy the accuracy requirements with the processing of pseudo-range data? Or must we process other data such as carrier phase observations instead (or as well)?
- Do we require the results in real-time? Or can we post-process the data after the data collection is completed?
- What is the benefit of combining data from other sources?

Depending on the problem at hand a number of techniques have been developed, generally based on the differential positioning mode.

1.4. DIFFERENTIAL GPS PROCEDURES

The rationale behind the differential or relative GPS positioning techniques was briefly outlined in Section 1.3.3. In the case of differential kinematic GPS the coordinates and velocities of one or more moving receiver are determined relative to a fixed "reference" station (or stations) to a higher degree of accuracy than the absolute position and velocity. There is a wide variety of schemes for implementing differential kinematic GPS, but they can be classified according to a number of attributes:

(1) Data type processed; and if more than one data type is used, the manner of their combination.

(2) Processing strategy adopted; whether the data reduction is executed within the GPS receiver itself (internal processing) or by an external computer (external processing), and whether the processing is distributed or centralised.

(3) Estimation algorithms used; whether the reliance on determining changes in position (and velocity) is placed on the analysis of the measurements alone, or on *a priori* knowledge of trajectory and other parameters, or both as in a Kalman filter algorithm.

Each of these is discussed below in a little more detail.

1.4.1. Data Type

The various types of measurement data that can be used for differential kinematic GPS have already been mentioned. In this section however, the type of measurement data is not only relevant, but also the type of data used to relate the measurements and/or positions obtained at different sites. All GPS receivers measure pseudo-range. It is the ideal observable for determining instantaneous position as 4 pseudo-range observations made simultaneously to 4 satellites permit a navigation solution or "fix" to be obtained: the 3-dimensional coordinates of the receiver and the receiver clock offset from GPS time. There is no redundancy in such a solution, and the GPS receiver will select (automatically or through operator command) the "best" 4 satellites for the strongest solution. If 5 or more satellites are observed simultaneously it is

possible to obtain an optimum solution using a least squares estimation procedure. The major disadvantage of using pseudo-range data is that the measurement precision (and hence the navigation fix) is relatively low for absolute or single point positioning. Two techniques can be distinguished, in which GPS pseudo-range data can be used for differential kinematic positioning:

(1) **Ground station reference schemes**; they are essentially of two varieties:

(1a) **Ground Truth Method**, also referred to as differential correction in the solution domain. A stationary GPS receiver monitors the errors influencing the GPS navigation fix by comparing the stationary receiver position estimate -- the so-called "navigation solution" -- to the known, surveyed coordinates of the point, and informs GPS (mobile) users of the discrepancies. As the navigation solution is sensitive to the satellite constellation used to derive the position, the mobile receiver which receives the correction information to be applied to its own navigation fix should be observing the same satellites as the reference station. When only 4 satellites are visible, only one 4 satellite combination can be selected by users. With 5 satellites visible, 4 combinations are possible. However, with 8 satellites (the maximum expected to be visible) there are 70 combinations of 4 satellites, and 163 combinations of 4 or more satellites ! As each combination corresponds to a set of at least 3 correction factors, a lot of data must be transmitted to ensure that the mobile user can select the appropriate position corrections in any case.

(1b) **Corrected Pseudo-Range Method**, sometimes referred to as differential correction in the observation domain. Using the known (surveyed) position of the reference station, ranges to the satellites can be derived and the offset from the actually observed pseudo-ranges can be computed. If 8 satellites are visible, the reference receiver need only provide corrections to the 8 pseudo-ranges. However, the mobile user's receiver must apply the corrections to the measured pseudo-ranges *before* they are used to obtain a navigation solution.

Both implementations yield virtually identical navigation solutions. These schemes are the most oft used techniques because they are relatively simple to implement, and hence they are the key to real-time differential kinematic applications. The reference station need not even be aware of the other users, as it merely broadcasts correction data. No "raw" measurement data is transmitted either by the reference station or the mobile stations.

Ground station methods rely on a data link between the reference station and all mobile receivers wishing to take advantage of the correction factors derived by the reference station.

(2) **Simultaneous data processing schemes**; these require that the data be transmitted from one or more GPS receiver to a central location (for example the ground reference station), where all the data can be processed together. Such processing can be based on the data differencing techniques used by surveyors to process carrier phase data, or simpler schemes which nevertheless take into account the between-station correlations present in the data from a number of simultaneously operating receivers.

The latter schemes rely on some centralised processing of the observed pseudo-range data. Therefore the "raw" measurements need to be transmitted between receivers, if (near) real-time results are required. For example, simultaneous processing allows several stations to be considered as references and, indeed, ground station reference schemes can be regarded as particular applications of the simultaneous data processing. The main purpose of this thesis is to describe a simultaneous data processing scheme based on a Kalman filter algorithm.

1.4.2. Processing Strategy

There are a number of processing strategies possible, each based on a different combination of internal processing (within the GPS receiver) and external processing, generally using a personal computer (PC). There is an obvious trade-off between a differential kinematic positioning system based exclusively on internal receiver processing (a capability provided by

possible to obtain an optimum solution using a least squares estimation procedure. The major disadvantage of using pseudo-range data is that the measurement precision (and hence the navigation fix) is relatively low for absolute or single point positioning. Two techniques can be distinguished, in which GPS pseudo-range data can be used for differential kinematic positioning:

(1) **Ground station reference schemes**; they are essentially of two varieties:

(1a) **Ground Truth Method**, also referred to as differential correction in the solution domain. A stationary GPS receiver monitors the errors influencing the GPS navigation fix by comparing the stationary receiver position estimate -- the so-called "navigation solution" -- to the known, surveyed coordinates of the point, and informs GPS (mobile) users of the discrepancies. As the navigation solution is sensitive to the satellite constellation used to derive the position, the mobile receiver which receives the correction information to be applied to its own navigation fix should be observing the same satellites as the reference station. When only 4 satellites are visible, only one 4 satellite combination can be selected by users. With 5 satellites visible, 4 combinations are possible. However, with 8 satellites (the maximum expected to be visible) there are 70 combinations of 4 satellites, and 163 combinations of 4 or more satellites ! As each combination corresponds to a set of at least 3 correction factors, a lot of data must be transmitted to ensure that the mobile user can select the appropriate position corrections in any case.

(1b) **Corrected Pseudo-Range Method**, sometimes referred to as differential correction in the observation domain. Using the known (surveyed) position of the reference station, ranges to the satellites can be derived and the offset from the actually observed pseudo-ranges can be computed. If 8 satellites are visible, the reference receiver need only provide corrections to the 8 pseudo-ranges. However, the mobile user's receiver must apply the corrections to the measured pseudo-ranges *before* they are used to obtain a navigation solution.

Both implementations yield virtually identical navigation solutions. These schemes are the most oft used techniques because they are relatively simple to implement, and hence they are the key to real-time differential kinematic applications. The reference station need not even be aware of the other users, as it merely broadcasts correction data. No "raw" measurement data is transmitted either by the reference station or the mobile stations.

Ground station methods rely on a data link between the reference station and all mobile receivers wishing to take advantage of the correction factors derived by the reference station.

(2) **Simultaneous data processing schemes**; these require that the data be transmitted from one or more GPS receiver to a central location (for example the ground reference station), where all the data can be processed together. Such processing can be based on the data differencing techniques used by surveyors to process carrier phase data, or simpler schemes which nevertheless take into account the between-station correlations present in the data from a number of simultaneously operating receivers.

The latter schemes rely on some centralised processing of the observed pseudo-range data. Therefore the "raw" measurements need to be transmitted between receivers, if (near) real-time results are required. For example, simultaneous processing allows several stations to be considered as references and, indeed, ground station reference schemes can be regarded as particular applications of the simultaneous data processing. The main purpose of this thesis is to describe a simultaneous data processing scheme based on a Kalman filter algorithm.

1.4.2. Processing Strategy

There are a number of processing strategies possible, each based on a different combination of internal processing (within the GPS receiver) and external processing, generally using a personal computer (PC). There is an obvious trade-off between a differential kinematic positioning system based exclusively on internal receiver processing (a capability provided by

instrument manufacturers at an additional cost over and above the "basic" GPS unit) and one which uses the GPS receiver(s) as sensors alone, and in which the data is externally processed. The latter can also be relatively expensive as it requires additional hardware and software. In the case of real-time applications, there are further constraints on the processing strategy, and a distributed processing system is preferable to a centralised one. The range of possibilities is best illustrated with some examples:

(a) **Internal / Distributed / Real-time Processing System**: Each GPS receiver would have its own processing software. However, because any external processing is to be avoided there would need to be different software for the reference station, and for the mobile user receivers. This processing strategy is closely associated with the ground station reference schemes. Each of the mobile receivers must be capable of decoding the received correction message, and applying the corrections either *before* the navigation fix is computed (for the observation domain correction scheme) or *after* the navigation fix is computed (for the solution domain correction scheme). Note that the transmitted signals may not be in synchronisation with the computation cycle of the receivers, and a certain amount of "intelligence" has to be built into the differential software.

(b) **External / Distributed / Real-time Processing System**: This would give equivalent performance to the scheme described above except that the internal processing would not need to be of a specialised nature. Off-the-shelf GPS receivers (which would be expected to be cheaper than those specifically designed with a built-in differential positioning capability) would be attached to PC's. The differential corrections in either the measurement or solution domain could be received by a device independent of the GPS receiver and directly connected to the PC. Other configurations are possible, for example, the reference station may have all its functions built into its internal processor, whereas the mobile receivers may be of many varieties each with their own peculiar approach to utilising the transmitted correction information. Hence, the principal advantage of external processing is the flexibility offered to the mobile user.

(c) **External / Centralised / Off-line Processing System**: All schemes that rely on carrier phase data only are of the off-line variety. The data must be transmitted (or transported) to a central facility where specialised software is used to derive the position (and velocity) information well after data collection has been completed. When the capability of transmitting quickly and efficiently the observations is developed then real-time high precision kinematic positioning may become a reality. However it is likely that the "package" (hardware and software) will be expensive, and certainly not as "robust" as the differential positioning techniques based on pseudo-range. Approaches that take advantage of the synergy between pseudo-range and carrier phase data show greater promise (see KLEUSBERG, 1986; SEEBER et al, 1986; ABIDIN, 1988), due to their lower sensitivity to cycle slips. In time such combined pseudo-range and phase data techniques are likely to evolve into internal / distributed / real-time systems.

The processing strategy developed in this thesis can be considered as occupying a niche between the post-processed carrier phase methods and the real-time pseudo-range data methods.

1.4.3. Estimation Algorithm

There is a wide range of algorithms applicable to derive the required position and velocity information from the GPS measurements. Some simply select 4 simultaneously observed pseudo-ranges and solve the 4 equations for the 4 position unknowns (3 coordinates and the receiver clock offset) in a direct analytical manner, then compute the velocity by dividing the change in position by the time elapsed. When more than 4 satellites are observed, it makes no sense to reject some measurements in order to avoid over-determined systems of equations, as these can be elegantly solved using the principle of least squares. When the precision requirements increase, the modelling must also be improved, according to the principles presented in Section 1.3.4. One such improvement is the use of the predictability of the motion, expressed by a kinematic model, thus yielding information on the parameters prior to the adjustment of the measurements. The class of least squares estimation algorithms that make use of *a priori* information are generally referred to as Bayesian Least Squares, and **Kalman filters** when

the processing is sequential (MERMINOD & RIZOS, 1988). Most GPS receivers in fact use a Kalman filter to obtain the standard (single point) navigation solution. Such estimation techniques have a number of advantages:

- data from different sensors can be easily incorporated in a rigorous manner,
- they permit interpolation of position (and velocity) to times other than the discrete observation epochs,
- they can be "tuned" to a particular expected motion,
- they are optimal (in the least squares sense).

As the Kalman filter is a sequential data processor, the results obtained are based only on the previous observations. However, at the completion of data collection, it is possible to go back (if the data has been stored) and process the data sequentially *back in time*, using the principles of **optimal smoothing**. The development of a Kalman filter / smoother algorithm suitable for differential GPS is the basic aim of this thesis project.

1.4.4. Classes of Differential Kinematic GPS Procedures

A classification of the differential techniques can now be attempted:

(1) **Differential Navigation Techniques**: These are basically the pseudo-range techniques, both the solution domain and observation domain correction approaches, but include the new techniques of combined pseudo-range and carrier phase processing (KLEUSBERG, 1986), in which the carrier phase data is often used only to smooth the ranges prior to their use. They all rely on distributed processing, each receiver doing a considerable amount of its own processing. Although the internal processing may be relatively sophisticated (using, for example, a Kalman filter), the differential component is simple, consisting only of the transmission of correction factors. It is therefore comparatively robust, and ideal for real-time applications.

(2) **Carrier Phase Processing Techniques**: These techniques rely on centralised external processing. The data must be transmitted or transported to a central facility, and preprocessed to identify, and if

possible, repair cycle slips. The results are potentially of a very high accuracy, but are only available post-mission (REMONDI, 1985; MADER, 1986). No Kalman filtering is normally carried out, although the cycle slip detection and repair may make use of these techniques, see Section 1.5.

(3) **Differential Kalman Filter Technique**: In this approach there is only one Kalman filter (and smoother) and all the data is processed together. The data may be any of the GPS observables, as well as any other sensor data such as INS, heading, etc. As in the techniques above, the data must be transmitted or transported to a central facility for processing. It is therefore not very suitable for real-time applications. It is however the most "general" of the differential processing techniques in that all other techniques can be considered as being special cases, or simplifications, of the Differential Kalman Filter Technique.

1.5. FILTERING AND GPS

1.5.1. Expected Benefits

Navigation, mainly sea and air, is the traditional field of application of Kalman filtering, and GPS can simply be regarded as a new sensor. The present trend includes a development of techniques for navigation on land. The benefit of incorporating a kinematic model in the estimation procedure to account for the predictability of the motion can be illustrated by comparing static pseudo-range tracking sessions of various durations. The precision of position fixing using GPS, either in the static or kinematic mode, undergoes strong variations with time. This is depicted in Figure 1.4, where the future constellation of 18 GPS satellites is considered over a 12 hour period for Wellington, New Zealand. BDOP3 is an extension of the concept of PDOP (Position Dilution Of Precision) to an observation session in the static case (MERMINOD, 1988). In other words, PDOP is nothing more than an instantaneous BDOP3. The value of these indicators of precision is the ratio between the precision of the position and that of the pseudo-range measurements. In respect of Figure 1.4, the filtered solution may be compared with the results of a short static observation session (say 3 min in length). That is, we may substitute a PDOP precision indicator for the BDOP3 indicator, and the solution is smoothed through geometric trouble spots while information and strength are added throughout.

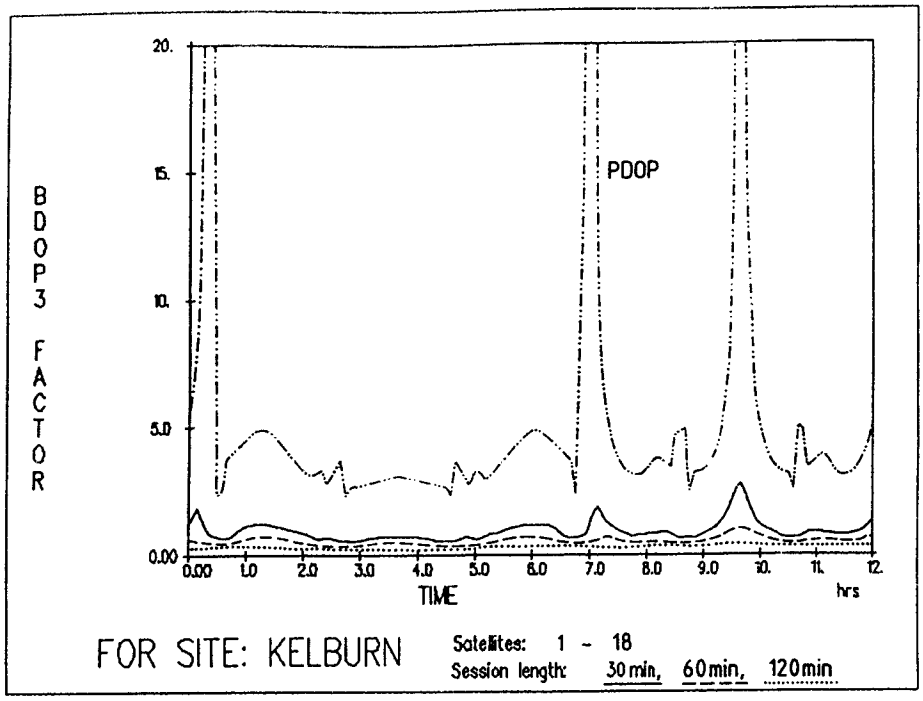


Figure 1.4: Improvement of the Positioning Precision with an Extension of the Tracking Session

1.5.2. Possible Applications

Filtering algorithms have been used long before the emergence of GPS. They have first been applied to the guidance of spacecrafts in the early 60's, see for example BUCY & JOSEPH (1968). The merits of Kalman filtering techniques for marine navigation have also long been recognised (DOVE, 1977). In this type of applications, a kinematic model for the trajectory of the vehicle is incorporated (that is, filtering occurs in the solution domain). GPS data is well suited for this type of processing and, as additional satellites are launched, will tend to become the major input for position determination. Apart from these applications, Kalman filtering techniques have been used for many investigations involving GPS, including:

- Range processing
- Orbit computation
- Ambiguity resolution
- Integrity monitoring
- Cycle slip editing

These applications are discussed briefly below.

Range processing: The range from a GPS satellite can be predicted, as long as the movement of both the satellite and the receiver can be modelled. In this case, it is possible to apply the filter to the residuals of the range measurements. The filtered ranges can then be used to compute the position of the receiver through a resection. This procedure is sometimes referred to as filtering in the observation domain.

Orbit computation: Kalman filtering is a technique used by the GPS Control Segment for the computation of the orbital elements and satellite clock corrections contained in the Navigation Message. The procedure is outlined by RUSSELL & SCHAIBLY (1980). At the end of each day, the tracking data acquired by the GPS Monitor Stations are combined with predicted orbital elements in a Kalman filter, in order to provide current estimates of the satellite positions and velocities. In other words, new tracking data are used to update the orbit estimation. After this update, orbits are extrapolated in the future and the orbital information transmitted by the satellite is based on curve fits to the extrapolated ephemerides. Although single-satellite orbit fits were initially used, SWIFT (1985) describes a multi-satellite filter/smoother implementation.

The use of GPS has also been advocated for tracking satellites other than those of the GPS (but equipped with a GPS receiver), and Kalman filtering has been investigated as an alternative to standard batch data processing procedures for high precision satellite orbit determination. In this context, YUNCK et al (1987) describe a " reduced dynamic strategy" and claim that sub-decimeter accuracy can be achieved even for low orbiting satellites such as TOPEX. AXELRAD & PARKINSON (1988) propose a similar Kalman filter approach for the support of satellite manoeuvres. Note that the motion of a satellite in its orbit is very well behaved, and can thus be accurately predicted.

Ambiguity resolution: This application of Kalman filtering is specific to static positioning using GPS carrier phase. The commercial post-processing software NOVAS (WANLESS & LACHAPELLE, 1988) includes a multi-hypothesis testing scheme based on Kalman filtering algorithms for the resolution of the cycle ambiguities, following a proposition by HWANG & BROWN (1985). This task may soon be carried out in real-time, and data collection could then stop as soon as a reliable set of integer biases is

determined. This would result in time savings and provide a useful field assessment of the quality of the data.

Integrity monitoring: The reliability of the GPS system is a cause for concern for all users, and integrity monitoring is becoming a major area of research. The problem is to test the reliability of the signals transmitted by the satellites. In case of a system degradation, the erroneous satellite(s) must be identified and health information provided, so that appropriate action can be taken. If 3 coordinates and the receiver clock must be estimated, there is some redundancy in the adjustment when 5 satellites are available. In addition, for a gross error caused by any one satellite to be detected, each subset of 4 satellites must provide a strong geometry. Even when the full GPS satellite constellation is implemented, this requirement will often not be fulfilled. This is why filtering can contribute to this task. Some implementations also include a multi-hypothesis testing scheme (McBURNEY & BROWN, 1988).

Cycle slip editing: Using GPS carrier phase data alone, MADER (1986) reports an airborne experiment in which agreement with a laser altimeter at the decimetre level was obtained. Also using the GPS carrier phase, but in a differential intermittent mode (see Section 1.2.3), REMONDI (1985) reports survey tests where centimetre accuracy was achieved within seconds. Once the cycle ambiguities have been determined on a pair of static sites, and as long as lock is kept on all satellites, it is sufficient to set the roving antenna for a few seconds on each survey mark. Even the initial determination of the cycle ambiguities can be accelerated by "swapping" the antennae between a site pair, yielding the double of the vector between the two sites.

As a result of improvements in the reliability of GPS receivers, and under favourable observing conditions, the occurrence of losses of lock can be minimised. However, it is worth pointing out that these techniques place a **complete reliance on the carrier phase measurements**. In effect, no assumptions are made as to the nature of the motion, rather it is assumed to be random. Cycle slips result in a systematic positioning error that can only be discovered by resetting the roving antenna back at a known location. The effect of a cycle slip on the estimated mean error and on the true error in position is illustrated in Figure 1.5. The relative position of both antennae is known at epochs e_A and e_Z.

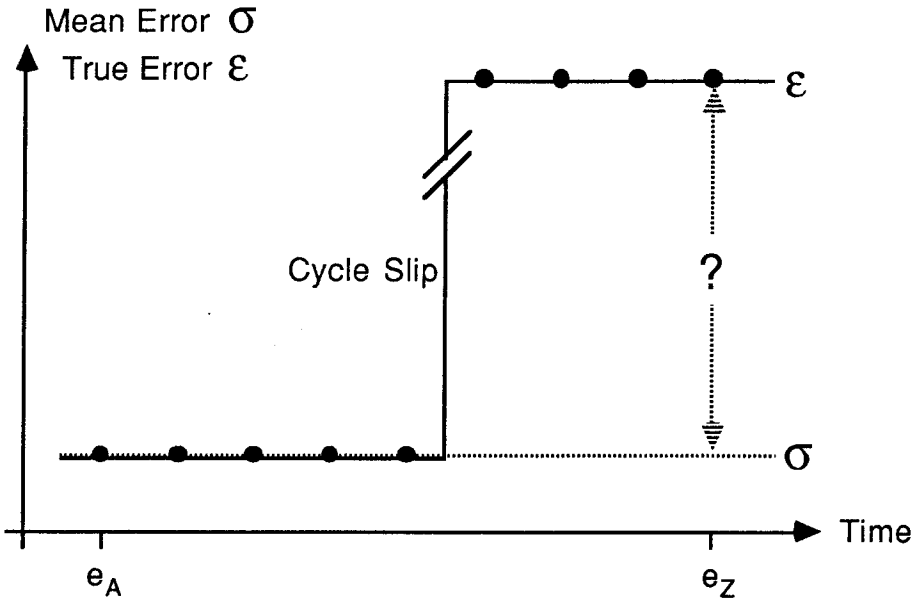


Figure 1.5: Cycle Slip in Intermittent Mode

The problem of cycle slip repair is therefore a critical one for such GPS kinematic positioning strategies. Indeed, unlike the situation of GPS integrity monitoring, the degradation has its origin in the user segment, but the problem is very similar, albeit at a higher precision level. Again, an abundance of available satellites provides error detection capability, but cycle slip recovery techniques can be developed that do not rely on GPS measurements alone:

- Incorporating an inertial sensor into the system -- the very predictable short term behaviour of data from such sensors provides an ideal counterpart to the GPS carrier phase observations (WONG et al, 1988). However, the measurement model for a strapdown inertial sensor is rather complicated, and the integration of the different types of data in a Kalman filter is not a simple task.

- Filtering the phase residuals -- to be efficient, the unmodelled phase residuals must be well behaved, so that a kinematic model can be used. This method is only applicable with dual frequency receivers, because the variations in the ionospheric delay must be accounted for. At this level of precision in the modelling, the range residuals are often called ionospheric residuals, as the ionosphere contributes the major component. Additional difficulties arise, however, when cycle slips occur on both frequencies at the same time (STOLZ & HEIN, 1988).

Both methods resort to filtering techniques. They have been tested in a research environment, but it may take some time before they are truly operational.

1.5.3. Approach Selected

The solution chosen in the present work consists of considering the phase-rate observable instead of the carrier phase. Therefore, the cycle ambiguity need not be resolved and short losses of lock on a satellite do not affect the processing of previous and subsequent valid measurements. It is clear that this technique cannot deliver the ultimate accuracy that GPS may provide, but this is a concession that must be made for a technique that is more robust and relatively simple.

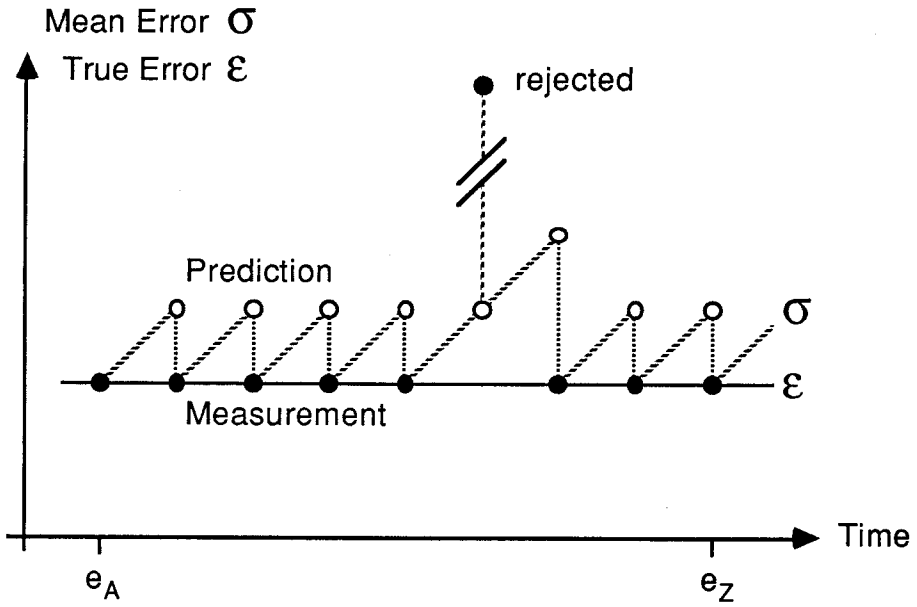


Figure 1.6: Phase-rate and Prediction

The phase-rate observation is almost independent of the position of the receiving antenna, but is closely related to its velocity. By estimating the velocity of the antenna, together with its position, and relating them through a kinematic model, the redundancy in the determination of the position is increased. In addition, the position at a subsequent epoch can then be more accurately predicted. This predicted position is used to model the incoming observations. If an actual observation differs greatly from its predicted value, it may then be rejected. Thus, the kinematic model provides a plausibility test for the measurements. Figure 1.6 shows the evolution of the mean and true

errors in position during a step-by-step estimation procedure that includes a kinematic model. The increase in position uncertainty associated with its prediction is balanced by a decrease due to new measurements. The capability of rejecting erroneous measurements ensures consistent results.

Figures 1.5 and 1.6 represent different approaches to ensuring the continuity of the positioning process: the integration of the GPS carrier phase or the use of the phase-rate, together with a kinematic model for the position parameters. The comparison is further illustrated in Figure 1.7, where shaded rectangles represent the epochs of carrier phase measurement.

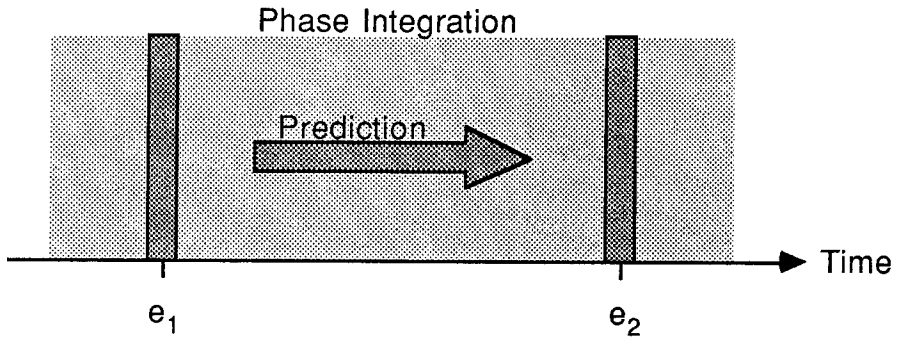


Figure 1.7: Two Ways of Connecting Epochs

The method used in this work can be described as the filtering approach to computing phase smoothed pseudo-ranges, and has already been implemented by SCHWARZ et al (1987). The software package DYNAMO for differential kinematic GPS positioning developed during the present thesis project incorporates this capability. A justification for this choice is also provided by ABIDIN (1988): the use of the phase-rate data, together with a prediction model, has proven efficient during periods of cycle slips. Figure 1.8 presents the situation of DYNAMO with respect to the 3 attributes generally associated with either surveying or navigation techniques, and introduced in Section 1.3.4.

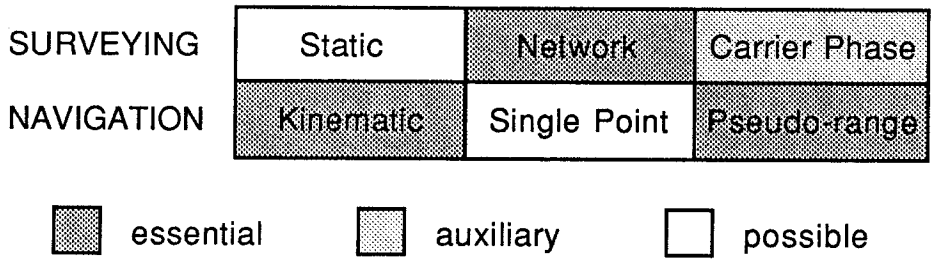


Figure 1.8: Present Implementation of DYNAMO

Future developments will permit the incorporation of the carrier phase as an observable. This was borne in mind throughout the software development, so that the structure of the package will not require any substantial change when this extension is undertaken. The carrier phase data will then assume the major role and the pseudo-range will only assist in the determination of the cycle ambiguities and the detection of cycle slips. The consequence is illustrated in Figure 1.9. Clearly, improvements in the modelling bring the approach closer to the surveying methodology, especially regarding the field and data processing procedures. Indeed, the only strictly navigation attribute that remains is the fact that the object is moving !

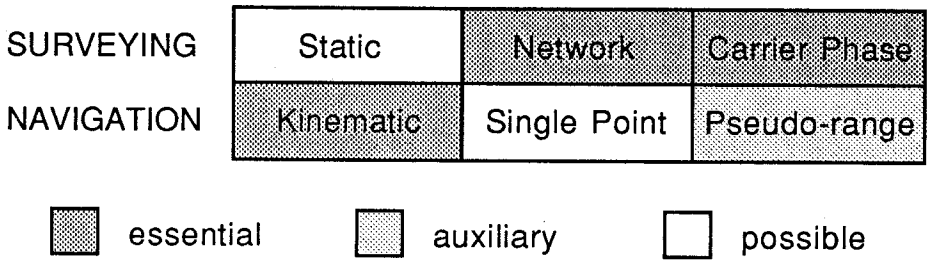


Figure 1.9: Future Implementation of DYNAMO

The evolution from Figure 1.8 to Figure 1.9 is not surprising: in the coming years, the greatest impact of carrier phase processing is likely to remain in surveying applications, even in the kinematic mode. Examples are the fast capture of survey data on land (in intermittent mode), precise positioning of platforms at sea, and aircraft positioning in support of aerial photography.

1.6. SYNOPSIS

Chapter 2 presents a short revision of the least squares methodology and a classification of the different estimation procedures. Some useful derivations are in Appendices A and B.

Sequential estimation methods are developed in Chapter 3. Most derivations have been placed in Appendices (C through E). In the main text, emphasis is therefore on principles rather than mathematical details.

In Chapter 4, the scope of navigation filters is introduced and the different functional and stochastic models -- as implemented in the software package DYNAMO -- are described in detail. These include the modelling of the motion of the receiver, the GPS observations and some constraints applicable for position and velocity parameters.

The software package DYNAMO is described in Chapter 5. In particular, the data-flow through the different components is shown.

Procedures for monitoring and testing filter performances are discussed in Chapter 6, where results using simulated data are also presented.

Chapter 7 presents results obtained with DYNAMO, using real data collected during experiments on land and at sea.

Chapter 8 draws the conclusions and indicates areas where further research is needed.

2. LEAST SQUARES PROCEDURES

2.1. LEAST SQUARES IN GEODESY

The basic concepts of modelling a physical system, of making measurements in such a system, and of carrying out an adjustment in order to obtain best estimates for the model variables of the physical system are central to all experimental sciences. In the case of surveying, almost all operations are essentially concerned with estimating such properties of a system as position, orientation, size and shape to a high precision using direct and indirect measurements. In addition, surveyors are generally known for dedicating their attention to remarkably static objects, such as hilltop trig stations and property boundaries. Traditionally, the analysis of movements is limited to those slow enough to be neglected during the actual measurement process but which may be evident only over a long period of time. A typical example is the monitoring of the deformation of a dam wall. This approach certainly restricts the field of applications of a surveyor's measurement and analysis expertise to those that are essentially **static** in nature. On the other hand, this has resulted in surveyors having a special committment to ensure that the results obtained are of high precision and good quality. Some characteristics of survey operations therefore are:

- **Measurement accuracy** -- these usually have an accuracy (relative to the system dimensions) greater than in most other sciences. Geodesy is therefore, in many respects, a "measurement science".

- An insistence on obtaining **redundant observations** -- encapsuled in such dictums as "one measurement is no measurement".

- The importance placed on **error analysis** -- a committment to precision and reliability requires that the quality of the results be assured. Results without an accompanying "quality assessment" are held in suspicion.

These factors have tended to stimulate the development and refinement of a number of procedures, based on the principle of least squares, that are designed to give the most precise and reliable results possible from a set of observations.

Although much of the pioneering work on least squares methods was done by applied mathematicians and astronomers, the role of the geodesist has nevertheless been, and continues to be, important. Geodetic studies were amongst the first practical applications of the method of least squares. Within the context of "classical" geodetic adjustment, generations of surveyors have been trained in the correct application of techniques such as adjustment by the parametric or condition method, with a minimum attention paid to the underlying statistical concepts. However the scope of least squares is much wider and the selection of the quadratic form to minimise is open to some choice, even in a limited static process such as a geodetic network adjustment. This chapter reviews the traditional parametric and condition methods of adjustment, as well as introducing some generalisations of the least squares principles.

Classical least squares formulations have been found wanting as the bases for data processing procedures for new technologies such as the Inertial Surveying System (ISS) and the Global Positioning System (GPS). The underlying mathematical models of both of these positioning technologies contain variables that change with time, and hence rely on new (for geodesists and surveyors) estimation techniques such as least squares filtering and smoothing, which constitute the topic of Chapter 3.

2.2. METHODOLOGY

2.2.1. The Two Models

The least squares approach is based on two models being assigned to the measurements:

- the **functional model**, relating the measurements and the parameters, for example, range = f(coordinates of 2 points, characteristics of the medium, etc.).

- the **stochastic model**, describing the statistics of the measurements, for example, all measurements are independent and have the same precision.

Neither the measurements, nor the functional model are perfect, and redundant measurements are not usually compatible with the chosen functional model. After the adjustment is completed, there remains a fraction of the measurements that cannot be accounted for by the estimated parameters: the residuals. Any **sufficient** subset of functionally independent observations can be used to assess the functional model (for example, derive the parameters of interest in the model), but each minimum subset would yield a different result. In order to obtain a unique result from redundant observations, it was necessary to introduce an optimality criterium that the model variables had to satisfy. "Least Squares" is a group of well defined methods used to obtain a unique and "optimal" solution from redundant measurements containing observational errors.

The principal advantage of Least Squares is the capability of checking the validity of the models: assumed measurement errors are propagated into the parameters and residuals via the functional model, yielding their expected stochastic properties. Discrepancies between the expected and actual statistics of the residuals reveal shortcomings in either the functional or the stochastic model. This provides a check of both models, and the notion of feedback is therefore inherent to the least squares estimation method.

2.2.2. The Criterion

It is worth emphasising that least squares is inherently **deterministic** in nature. It is a rigorous method by which a unique result to an overdetermined problem can always be obtained, no matter what the error characteristics of the observations are. Least squares adjustment does not require the observations to have errors that are normally distributed, or to have any other type of distribution for that matter. The least squares estimates of the model variables are defined as those which minimise a specific quadratic form of the residuals:

$$\mathbf{v}^T \mathbf{P} \mathbf{v} \rightarrow \text{minimum}$$

where $\mathbf{P}$ is the matrix that selectively weights the observations according to their respective quality, and $\mathbf{v}$ are the "corrections" to the observations necessary for the functional model to be satisfied. The least squares estimate is known as the Best Linear Unbiased Estimate (BLUE). Although it is fashionable these days to develop the theory of least squares by first **defining**

the statistical properties required of the estimates and then to **derive** the least squares process (including the minimum quadratic form), the end result is identical.

2.2.3. The Parameters

Parameters in the models can, in principle, be divided into a number of classes:

- **unadjusted parameters** -- they are considered fixed and remain unaffected by the estimation procedure (for example, the speed of light),

- **nuisance parameters** -- they are considered as variables in the functional model and estimated, but their estimated value is irrelevant (for example, GPS clock errors),

- **parameters of interest** -- the desired output of the estimation procedure (for example, the coordinates of the GPS antenna).

Of course, any particular parameter can be associated with any of the above classes. For example, GPS satellite orbital parameters are considered unadjusted parameters for the navigator, but nuisance parameters for the geodesist and parameters of interest only for the orbit determination specialists. This does not alter the basic structure of the least squares approach, only the models. For navigation applications, typical parameters of interest are latitude, longitude, heading and speed. Although heading and speed can be estimated directly, it is often more convenient to derive them from the estimated rate of change in latitude and longitude. Other parameters such as height and slope are only estimated in land surveying operations.

2.2.4. Redundancy

Surveyors who regularly use the GPS positioning technology invariably process the carrier phase measurement, because of its high resolution. This observable contains one unknown constant -- the cycle ambiguity (Figure 1.2) -- which can only be estimated reliably after all the sources of variations in the ranges to the satellites have been accounted for with a precision better than the wavelength of the carrier signal. This can only be achieved in the static

differential positioning mode during which data is collected over a finite time period. Indeed, it is because of the high redundancy available in the adjustment of measurements acquired in the static mode that the rather complicated modelling of the carrier phase observable can be attempted.

On the other hand, not all surveys need to deliver the highest accuracy. In some cases (for example, prospected points in a geophysical exploration survey), use of the pseudo-range observable is adequate. The reduced number of unknowns (no cycle ambiguity parameters in the solution) is an advantage, particularly when the redundancy in the measurements is low. This is also the case for most kinematic applications. The gain in redundancy realised by using an observable that is easier to model (less nuisance parameters) is virtually lost when kinematic applications are considered (as there are in effect more position parameters, for example, one set for each epoch). That is why the predictability of the motion is so precious. By adding positioning information -- the term "pseudo-observations" is appropriate -- redundancy can be improved and the least squares techniques can be applied.

In summary, the principle of least squares yields efficient estimation procedures as long as there is sufficient scope for varying the complexity of the models and enough adjustment redundancy can be assured.

2.3. CLASSIFICATION OF LEAST SQUARES PROCEDURES

2.3.1. From Static to Kinematic Estimation

In order to relate the various least squares procedures used by geodesists, an attempt is made to classify them with respect to three criteria:

- the **available data** -- either measurements only, or, in addition, available *a priori* estimates of the parameters,

- the **processing mode** -- either batch (that is, all of the data at once) or step-by-step (sequential),

- the **system state behaviour** -- either static or time-varying (kinematic).

The most common example of a time-varying system that is likely to be encountered by surveyors is one in which the instantaneous position of a moving platform is to be estimated using GPS or ISS. For such kinematic applications, step-by-step procedures are the key to the delivery of results in real-time. However, the two concepts should not be confused. The choice of the processing mode is not imposed by the system behaviour, but rather it is often a matter of data organisation. Step-by-step estimation procedures can be very useful even in the static case. They have proven to be very efficient for upgrading geodetic networks as new survey data becomes available, as for example the incorporation of EDM measurements into a previously triangulated network. Estimation procedures appropriate for static case and kinematic systems are presented in Figure 2.1.

STATIC		**Data**	
		measurements	measurements + parameters
Mode	batch	Classical	Bayesian
	step-by-step	Sequential	Bayesian Sequential

KINEMATIC		**Data**	
		measurements	measurements + parameters
Mode	batch	Classical	Bayesian
	step-by-step	Step-Classical	Filter

Figure 2.1: Least Squares Estimation Procedures for Static and Kinematic Systems

Considering a set of n observations and u parameters at each of e epochs, the change from a static system process to a kinematic system process can be summarised for the different cases:

- **Classical** and **Bayesian**: The number of parameters increases. However, a relation between the parameters at different epochs can be modelled. Typically, truncated time polynomials are used for such a purpose. The total number of parameters therefore is between u (static)

and $u \cdot e$ (independent states at all epochs). Thus, even if n is smaller than u, it is possible to model the system with enough epochs and a limited complexity of the kinematic parameter model (for example, by a moderate order of polynomial).

- **Step-Classical**: By definition, only measurement data are used. The parameters of the measurement model are not related between steps, and no kinematic model can be considered. Thus, n must be larger than u at all epochs. This procedure can be considered as a succession of classical adjustments. However the stringent requirements on the number of measurements makes this method of limited use for many applications other than real-time navigation.

- **Filter**: The state vector at one epoch can be related to the previous ones. Thus, n may be smaller than u. A filter can be defined as Bayesian sequential estimation in a kinematic environment. Filters are therefore considered here as the most general process. A filter with weights attributed only to measurements reduces to sequential least squares and one step of the filter is equivalent to Bayesian least squares.

2.3.2. Why Filtering ?

From the brief discussions in Sections 1.2.3 and 2.3.1, it is clear that there are different ways of modelling time-varying parameters. The rationale behind the functional approach to kinematic modelling is to replace the time-varying parameters of interest (for example the coordinates), by auxiliary constant parameters (for example the coefficients of a polynomial). Therefore, with regard to the estimation procedure, **the kinematic system becomes a static one !** The auxiliary parameters replace the parameters of interest in the functional model of the measurements. Thus there is no distinct kinematic model, but it is included in the measurement model. However, it must be noted that step-by-step processing of the measurements is still possible, as for any other static problem. All information from previous measurements is contained in the estimated auxiliary parameters and their covariance matrix, and can be input as *a priori* parameter data in the next processing step. This combination of Mode and Data results in a Bayesian Sequential procedure, see Figure 2.1. A number of drawbacks result from such a functional modelling of kinematic systems:

- The parameters of interest must be substituted for auxiliary parameters in the functional model of all (types of) measurements.

- The parameters of interest are not output directly by the least squares algorithm, but must be computed for each selected epoch, from the estimated auxiliary parameters. The same applies for their precision.

- If the signature of substituted parameters does not match the expectation (for example, the sign of the trend changes more often than expected), the entire processing must be repeated with other auxiliary parameters (for example, a higher order of polynomial).

In summary, this approach is only appropriate for very well-behaved motions.

By comparison, stochastic modelling allows for a selective weighting of the information. The noise added in each extrapolation progressively reduces the weight attributed to previous position determinations. This "fading memory", together with the absence of predefined signature for the trajectory, increases the ability to adapt to a new system behaviour. This characteristic is highly desirable for navigation. The stochastic approach to kinematic modelling is the basis of filtering techniques. Measurement and kinematic models are independent and the drawbacks of the functional approach listed above disappear. The main difficulty in filtering is the choice of the appropriate system noise model, and this problem will be addressed in Chapter 4.

2.4. PRINCIPAL FORMULATIONS OF CLASSICAL LEAST SQUARES PROCEDURES

2.4.1. Condition Method

Condition equations express properties that the observations should satisfy. The general form of a condition equation is $f(\mathbf{l}) = 0$, where $\mathbf{l}$ is the vector of observables (ideal observations). Actual observations l are generally biased by a number of errors and therefore do not satisfy this condition. A vector of misclosures can be computed as $f(l) = \mathbf{w}$. The adjustment aims at computing corrections $-\mathbf{v}$ to the observations such that the corrected observations satisfy both the relation $f(l - \mathbf{v}) = 0$ and the least squares

condition $\mathbf{v}^T\mathbf{P}\mathbf{v} \rightarrow$ minimum. The convention used in this thesis is that a correction must be added to the observation, and that the correction has the same magnitude as the residual but is of opposite sign.

The linearisation of the condition equation is based on a Taylor's series expansion to the first order:

$$f(l - \mathbf{v}) = f(l) - \mathbf{B}\mathbf{v}$$

where $\mathbf{B}$ is the design matrix, containing partial derivatives of $f(\mathbf{l})$ with respect to $\mathbf{l}$ about the actual observations l. The covariance matrix $\mathbf{Q}_{ll}$ of the observations is assumed known. The computational procedure can be summarised in a few steps:

Linearised form: $\mathbf{B}\mathbf{v} = \mathbf{w}$ with weight matrix: $\mathbf{P} = \mathbf{Q}_{ll}^{-1}$ (2-1)

Solution for the residuals: $\hat{\mathbf{v}} = \mathbf{Q}_{ll}\mathbf{B}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}\mathbf{w}$ (2-2)

Covariance matrix of the residuals: $\mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll}\mathbf{B}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{Q}_{ll}$ (2-3)

Covariance of the adjusted obs.: $\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q}_{ll} - \mathbf{Q}_{\hat{v}\hat{v}}$ (2-4)

The advantage of the condition method is that the unknown terms are simply corrections to the observations. The size of the matrix $\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T$ to invert is $r \times r$, where r is the number of conditions, which is in fact equal to the number of redundant observations. However, there are a number of drawbacks. The derivation of adjusted values for functions of the observables (for example, the coordinates) is tedious, as is the derivation of their respective precisions and correlations. Furthermore the construction of (2-1) requires a sound geometrical understanding of the situation, as only independent conditions must be used. Consequently, the setting-up of the equations is not easily automated for a computer. However this method was especially popular for geodetic network adjustments when computers were not available.

2.4.2. Parametric Method

This method of adjustment makes use of observation equations, where observables are expressed as a function of some or all of the parameters, in the general form: $\mathbf{l} = f(\mathbf{x})$. To satisfy this relation, actual observations need to be corrected or "adjusted". The linearisation of the relation is performed about an approximate set of parameters $\overset{\circ}{\mathbf{x}}$:

$$\begin{aligned} l - \mathbf{v} &= f(\mathbf{x}) \\ l - \mathbf{v} &= f(\overset{\circ}{\mathbf{x}} + \delta\mathbf{x}) \\ l - \mathbf{v} &= f(\overset{\circ}{\mathbf{x}}) + \mathbf{A}\delta\mathbf{x} \\ (l - f(\overset{\circ}{\mathbf{x}})) - \mathbf{v} &= \mathbf{A}\delta\mathbf{x} \end{aligned}$$

The expression in brackets on the left hand side of the equation is the "observed minus computed term", or approximate residual, and is denoted by $\overset{\circ}{\mathbf{v}}$. The covariance matrix $\mathbf{Q}_{ll}$ of the observations is assumed known. As $\overset{\circ}{\mathbf{v}}$ differs from l only by a constant, it has the same stochastic behaviour. The computational procedure therefore is:

Linearised form: $\overset{\circ}{\mathbf{v}} - \mathbf{v} = \mathbf{A}\delta\mathbf{x}$ with weight matrix: $\mathbf{P} = \mathbf{Q}_{ll}^{-1}$ (2-5)

Solution for the parameters: $\delta\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$ (2-6)

with covariance matrix: $\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}$ (2-7)

The adjusted observation residuals can be computed in two different ways:

- The direct approach, by including the adjusted state vector in the functional model of the observations:

  $\hat{\mathbf{v}} = l - f(\hat{\mathbf{x}})$, where $\hat{\mathbf{x}} = \overset{\circ}{\mathbf{x}} + \delta\hat{\mathbf{x}}$

- The indirect approach, by including the adjusted state vector increment in the linearised functional model of the observations (2-5):

  $\hat{\mathbf{v}} = \overset{\circ}{\mathbf{v}} - \mathbf{A}\delta\hat{\mathbf{x}}$

The second method clearly illustrates the relation between the approximate and adjusted states. Indeed, this is the main justification for the choice of the unusual symbol $\mathring{\mathbf{v}}$ to denote the vector of approximate residuals, hence ensuring complete consistency between quantities related either to $\mathbf{x}$ or $\mathbf{v}$. The covariance matrix of the residuals is derived from the indirect computation of the residuals using the "law of propagation of variances", and assuming the stochastic independence of the measurements and the approximate state vector:

$$\mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{A}\,\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{A}^T \qquad (2\text{-}8)$$

If the desired results (for example, the coordinates) are selected as the parameters, the solution of the system leads directly to the answer. There is exactly one equation per observation, and its form can easily be defined according to the type of observation. The size of the matrix to invert is $u \times u$, where u is the number of (unknown) parameters, and the setting-up of the equations can easily be automated in a computer program. The linearisation of the problem requires some *a priori* approximate knowledge of the parameters, which is usually available. For geodetic purposes, it is generally no problem to obtain a converged solution of a non-linear problem through an iterative process.

2.4.3. Relations Between Both Approaches

If any problem can reduce to either the condition or parametric case, there is actually a choice between the two methods of setting-up and solving the least squares problem. A demonstration of the passage from the parametric to condition equation method is given in, for example, BJERHAMMAR (1973). Given n linear observation equations with u (unknown) parameters, the elimination of all the parameters leads to a system of $r = n - u$ condition equations. The reverse of this process is much harder because there is an almost infinite number of possible parametrisations. A formal proof is of no practical importance: if a parametrisation has to be chosen at this stage, it could also be defined at the time of constructing the equations, thus leading to the parametric case. This indicates that the condition equation approach in fact represents a more fundamental concept, because it is based solely on intrinsic geometric properties of the observables. In some cases, for example if there are fewer redundant observations than unknown parameters, the

condition method offers computational advantages. However, the dramatic improvement in the power of computers has made these advantages largely irrelevant and the adjustment by parameters is now the standard solution approach to almost all overdetermined systems. Unfortunately the practical importance of the condition equation approach has been reduced to such an extent that not only its use, but also its fundamental features, are now often overlooked.

2.4.4. The Combined Case

It is also possible to formulate relations involving both observables and parameters. Functions of the observables are related to functions of the parameters, through the general form: $f(\mathbf{l}, \mathbf{x}) = 0$. For example, this method is useful when solving for transformation parameters (HARVEY, 1985). A linear relation is obtained following the usual procedure:

$$f(l, \overset{\circ}{\mathbf{x}}) = \mathbf{w}$$

$$f(l - \mathbf{v}, \overset{\circ}{\mathbf{x}} + \delta\mathbf{x}) = f(l, \overset{\circ}{\mathbf{x}}) + \mathbf{A}\delta\mathbf{x} - \mathbf{B}\mathbf{v} = 0$$

Linearised form: $-\mathbf{A}\delta\mathbf{x} + \mathbf{B}\mathbf{v} = \mathbf{w}$ with weight matrix: $\mathbf{P} = \mathbf{Q}_{ll}^{-1}$ (2-9)

Adjusted parameters: $\delta\hat{\mathbf{x}} = -(\mathbf{A}^T(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\mathbf{A})^{-1}\,\mathbf{A}^T(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\mathbf{w}$ (2-10)

Covariance matrix: $\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = (\mathbf{A}^T(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\mathbf{A})^{-1}$ (2-11)

Formulae for the residuals are given, for example, in CROSS (1983). This case is also sometimes referred to as "condition equations with parameters" (PACHELSKI, 1980), which has the merit of indicating that such a hybrid form is basically a condition equation. In fact, the condition and parametric methods can be thought of as two special cases. To demonstrate the equivalence between these two approaches, it suffices to rearrange (2-9) and consider particular design matrices:

Condition: $\mathbf{B}\mathbf{v} = \mathbf{w} + \mathbf{A}\delta\mathbf{x}$ with $\mathbf{A} = \mathbf{0}$
Parametric: $-\mathbf{w} + \mathbf{B}\mathbf{v} = \mathbf{A}\delta\mathbf{x}$ with $\mathbf{B} = -\mathbf{I}$

In the parametric case, the misclose vector $\mathbf{w}$ has the same magnitude as $\mathring{\mathbf{v}}$, the observation residual computed using the approximate parameters $\mathring{\mathbf{x}}$, but is of opposite sign.

2.4.5. Decomposition in Least Squares

It may be impossible to express all relations according to one of the models discussed above. In many cases however, it is possible to decompose a complicated processing step into elementary ones. This concept is best illustrated by examples. For the two cases presented here, the full derivation is given in Appendix A.

a) The Combined Case

The linearised form (2-9) can be written as a condition equation:

$$\mathbf{Bv} = \mathbf{w} + \mathbf{A}\delta\mathbf{x} \qquad \text{with weight matrix: } \mathbf{P} = \mathbf{Q}_{ll}^{-1} \qquad (2\text{-}12)$$

In a first step, no attempt is made to estimate $\delta\mathbf{x}$. Instead, a standard condition adjustment $\mathbf{Bv} = \mathbf{w}$ is performed. The residuals of the adjusted measurements are obtained using (2-2). In a second step, the relation between the residuals and the parameters is expressed as a system of observation equations. The parametric method is applied, and the adjusted increments $\delta\hat{\mathbf{x}}$ are obtained using (2-6). The final result is equivalent to (2-10), as shown in Appendix A. Thus the combined case can been solved by applying the condition and the parametric equation methods in succession.

b) Adjustment with Constrained Parameters

Let $\mathbf{x}$ be a vector of parameters, to be determined from observations $\boldsymbol{l}$, with covariance matrix $\mathbf{Q}_{ll}$. In addition, the parameters are subject to a constraint. The situation can be described by the following system of equations:

$$\mathring{\mathbf{v}} - \mathbf{v} = \mathbf{A}\delta\mathbf{x} \qquad \text{with weight matrix: } \mathbf{P} = \mathbf{Q}_{ll}^{-1} \qquad (2\text{-}5)$$

$$\mathbf{U}\delta\mathbf{x} = \mathbf{t} \qquad (2\text{-}13)$$

This parametric system of equations can be solved by minimisation of the quadratic form $\mathbf{v}^T\mathbf{Pv}$, subject to the constraints defined by the second

equation. Appendix A illustrates how the problem can also be solved by applying the parametric and the condition equation methods successively. This approach is easier to implement than a first look may suggest. The different steps are:

1) From the unconstrained parametric adjustment: $\mathring{\mathbf{v}} - \mathbf{v} = \mathbf{A}\delta\mathbf{y}$ compute the unconstrained solution to yield: $\delta\hat{\mathbf{y}}$, $\mathbf{Q}_{\hat{y}\hat{y}}$

2) Consider the parameters adjusted without constraints as observations of the constrained parameters, that is:

   $\delta\hat{\mathbf{y}} - \mathbf{z} = \delta\mathbf{x}$ where $\mathbf{z}$ is the vector of residuals.

   The constraint condition on the parameters becomes: $\mathbf{U}\delta\hat{\mathbf{y}} - \mathbf{U}\mathbf{z} = \mathbf{t}$

3) Form the misclosure vector: $\mathbf{w} = (\mathbf{U}\delta\hat{\mathbf{y}} - \mathbf{t})$, compute the condition adjustment: $\mathbf{U}\mathbf{z} = \mathbf{w}$. The solution yields $\hat{\mathbf{z}}$ and $\mathbf{Q}_{\hat{z}\hat{z}}$

4) The constrained solution is thus obtained as:
   $\hat{\mathbf{x}} = \hat{\mathbf{y}} - \hat{\mathbf{z}}$ and $\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q}_{\hat{y}\hat{y}} - \mathbf{Q}_{\hat{z}\hat{z}}$ (2-14)

If the condition equations are linearised around $\hat{\mathbf{y}}$, the expression for the misclosure vector simplifies to $\mathbf{w} = -\mathbf{t}$. An example of a constraint that may be applied to the parameters is that a distance between two points on a ship is to remain unchanged, whatever the overall movement of the ship may be. Note that in such a case the condition could also be replaced by a fictitious observation of the distance, equal to the specified value and with an infinite weight. This method is not rigorously applicable as weights must be finite for computations, though it is possible to consider quite large weights without causing numerical problems.

This method of decomposition in least squares therefore yields an elegant and rigorous solution. The computation burden can be kept within limits as the dimension of the matrix to invert is equal to the number of conditions on the parameters. For example, if one distance is to be kept fixed the central operation degenerates to a scalar inversion. Furthermore, **any** condition on

the parameters can be enforced, such as four points on a circle or fixed area of a polygon.

What makes this method particularly attractive is the possibility to apply a constraint on the parameters **after** the adjustment has been completed, without having to perform the entire adjustment again. In fact, constraints may be applied according to the principle of least squares to any set of parameters $\hat{\mathbf{y}}$, as long as the full covariance matrix $\mathbf{Q}_{\hat{y}\hat{y}}$ is available. This information usually results from a previous adjustment, but this need not be the case.

2.4.6. Bayesian Least Squares

In many cases, a fairly good *a priori* knowledge of the parameters is available. Thus, it is reasonable to require that the adjusted value of a parameter should not be too different from its *a priori* value. This condition can be imposed in a number of ways:

- The *a priori* values of the parameters are considered as observations. Suitable weights, relative to that of the measurements, are required. Consequently, the size of the system of parametric equations is augmented, but not that of the normal equations, as the number of parameters remains unchanged.

- The increments of the *a priori* values of parameters are included in the quadratic form to minimise, with appropriate weights.

The formal proof of the equivalence of both approaches for the parametric case is given in Appendix B. The resultant normal equations are:

$$(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{\mathring{x}})\,\delta\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}} \qquad (2\text{-}15)$$

The extension of the quadratic form to $\mathbf{v}^T\mathbf{P}\mathbf{v} + \delta\mathbf{x}^T\mathbf{P}_{\mathring{x}}\delta\mathbf{x}$ represents a generalisation of the classical least squares method. With $\mathbf{P}_{\mathring{x}} = \mathbf{0}$, that is, no *a priori* information on the parameters is available, Bayesian least squares reduces to the classical definition. The minimisation of the extended quadratic form can also be applied to the combined case. However, Bayesian least squares is not defined in the pure condition case, due to the absence of

parameters. Two examples will illustrate the effect of the generalisation of the quadratic form.

a) The Combined Case

The complete derivation for the combined case is given in, for example, KRAKIWSKY (1975). Compared with the solution for the classical method (2-11), the covariance matrix of the parameters is slightly modified:

$$\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{A}^T(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\mathbf{A} + \mathbf{P}_{\mathring{x}})^{-1} \qquad (2\text{-}16)$$

The solution for the parameters is modified accordingly and becomes, compare with (2-10):

$$\delta\hat{\mathbf{x}} = -(\mathbf{A}^T(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\mathbf{A} + \mathbf{P}_{\mathring{x}})^{-1}\,\mathbf{A}^T(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\mathbf{w} \qquad (2\text{-}17)$$

These expressions are correct as long as the *a priori* estimates of the parameters are also used for the linearisation of the equations, otherwise correction terms must be considered. The appropriate formulae are found in BOSSLER (1972).

b) Adjustment with Constrained Parameters

If $\delta\mathbf{x}^T\mathbf{P}_{\mathring{x}}\delta\mathbf{x}$ is added to the quadratic form in (A-3), the second diagonal term of the 4×4 matrix in (A-4) is replaced by $\mathbf{P}_{\mathring{x}}$. Following the same procedure from (A-4) to (A-8), it appears that the following substitution should be made:

$$\delta\hat{\mathbf{y}} = (\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{\mathring{x}})^{-1}\,\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}} \quad \text{and} \quad \mathbf{Q}_{\hat{y}\hat{y}} = (\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{\mathring{x}})^{-1} \qquad (2\text{-}18)$$

This is exactly the form of the solution of the Bayesian parametric case. It suffices to compare with (2-16) and (2-17) replacing $\mathbf{B} = -\mathbf{I}$ and $\mathbf{w} = -\mathring{\mathbf{v}}$. Thus the procedure developed in Appendix A.2 is still applicable.

In both examples, it was shown that adding the *a priori* weight matrix of the parameters $\mathbf{P}_{\mathring{x}}$ to the normal matrix of the observations leads directly to the normal matrix of the Bayesian adjustment. In (2-18), it is assumed that the *a priori* estimates of the parameters are also used for the linearisation of the equations, that is, $\mathring{\mathbf{x}}$ is used to compute $\mathring{\mathbf{v}}$.

2.4.7. Sequential Least Squares

We can differentiate between batch and sequential (or step-by-step) processing modes. The batch processing mode is the one most commonly encountered in classical geodesy. The least squares adjustment is carried out once all the data has been acquired. However, a sequential treatment of least squares problems may be preferable for several reasons:

- it divides a large computing burden into smaller parts, and reduces the requirements on both processing capability and storage.

- it is the key to real-time applications.

It is therefore tempting to treat successive batches of measurements sequentially, though it should be kept in mind that the results of a sequential least squares adjustment will be identical to that of a batch solution once all the data has been processed. However, this approach to the processing of data may influence other options in the formulation of the least squares problem, such as the choice between the condition and the parametric method. The formulae used in sequential least squares estimation are given in, for example, CROSS (1983). The primary considerations that should be borne in mind when using sequential least squares vis a vis other estimation procedures are briefly discussed below.

Storage: The use of a sequential adjustment process makes sense if it is possible to condense the information gained from some observations and make it available at the next processing stage. A very handy condensed form is a set of parameters and their associated covariance matrix. Apart from testing the consistency of the observations, this is the main reason why surveyors compute networks: coordinates and their covariances include the contributions of all observations, but usually are much more convenient to handle than a list of measurements. Nevertheless, archiving the original observations allows for a later modification of their respective weights, which would not be possible if only coordinate information were available.

Partitioned Parameter Set: Sequential processing is particularly useful if, in a set of observations, some parameters are not related to subgroups of observations. An example are the so-called local or epoch parameters which

describe the behaviour of a GPS satellite clock. Parameters of this kind can be eliminated from the normal matrix once the group of observations have been processed using, for example, Helmert-blocking techniques (CROSS, 1983). This reduces the amount of data and parameters needed for subsequent computations.

Heterogeneous Data: Measurements of different types, such as distances, directions and azimuths, can be modelled as functions of the same set of parameters. Furthermore, when it comes to kinematic applications, it is necessary to relate the observations and the movement. This is best achieved by expressing both the measurement and kinematic models with common parameters. Thus a suitable parametrisation allows for the processing of heterogeneous data in a step-by-step basis. Positions can be determined first using one type of measurement data, and then another, and so on.

Choice of the Functional Model: The choice of a parametrisation is more critical for sequential adjustments as the parameters are the link between the different groups of observations. In this respect, the evolution of adjusted observations, as obtained from condition adjustments, has limited value. Its explicit use is restricted to some particular applications like monitoring the deformation of a triangle by checking the variation of its angles. Writing a condition equation implies the simultaneous availability of several measurements. If one is missing, others may become useless, or other conditions need to be formulated. The parametric method, on the other hand, allows for more flexibility, as any measurement can be incorporated into an observation equation.

Correlation Between Steps: In some cases, measurements at different epochs are correlated. Whilst this simply results in a non-diagonal weight matrix in the batch processing mode, the effect is more difficult to deal with in a step-by-step adjustment. Two possible remedies are:

- decorrelate the measurements, that is, compute a set of linear combinations of the measurements so that its weight matrix is diagonal. Such a procedure, known as Gram-Schmidt orthogonalisation, is discussed in, for example, KING et al (1987).

- consider the measurement noise, which causes the correlation, as a "disturbance" of the state parameters. A suitable relation between the disturbance at different epochs is then input as a "pseudo-observation" in the functional model (see, for example, GELB, 1974).

The second method is preferable because it requires an increase in the size of the state vector rather than a new algorithm. Furthermore, being part of the estimation process, the correlation can be more easily adapted to changing circumstances, for example by modifying the weight of the pseudo-observation. However, this method is relevant only in the context of Bayesian estimation (Section 2.4.6).

3. PREDICTION, FILTERING AND SMOOTHING

For many applications of new surveying technology such as ISS and GPS the parameters of interest (usually position and velocity), or the dominant system errors, or both, are time-varying. Furthermore, the time variation is more or less predictable. For such applications, the data processing techniques that are the most efficient and optimal, and therefore the most appropriate, are those based on the principles of least squares prediction, filtering and smoothing.

Least squares filtering has its origins in electrical signal processing, and consequently the literature mainly reflects a preoccupation with signal processing and communication engineering applications. The Kalman filter is perhaps the best known of the techniques that have gained wide acceptance across a wide spectrum of physical and engineering sciences. The synergy between classical least squares procedures used in geodesy (and briefly discussed in Chapter 2), on the one hand, and the new filtering techniques on the other hand, has been recognised for over a decade (KRAKIWSKY, 1975). However, it was only with the advent of ISS and GPS that considerable interest has been aroused by geodesists in least squares filtering as a data processing tool.

3.1. PRINCIPLES

The three concepts of prediction, filtering and smoothing are closely related and are best illustrated through an example. Let us assume a moving vehicle, and that the parameters of interest are its instantaneous position at some time t. The process of computing the vehicle's position in real-time (that is, observations are taken at time t_k , position required at t_k) will be referred to as **filtering**. The computation of the expected position of the vehicle at some subsequent time t_k , based on the last measurements at t_{k-1} is properly termed **prediction**, while the estimation of where the vehicle was (say at time t_k), once all the measurements are post-processed to time t_{k+1} , is referred to as **smoothing**. The three steps of the estimation procedure are illustrated in Figure 3.1.

- the **prediction step**: based on past positioning information together with a kinematic model, the expected position and its precision at the next epoch of measurement is computed. The kinematic model is therefore composed (as is the measurement model) of functional and stochastic components.

- the **adjustment** or **filtering step**: this is a classical least squares adjustment, except that a fairly good *a priori* estimate of the parameters is already provided from the prediction step. Basically, the resulting parameter estimates are weighted combinations of predicted quantities and measurement data. As both the kinematic and the measurement model are composed of functional and stochastic components, four models must be considered. Given a particular application and a certain data type, the filter design process is therefore one of selecting the appropriate models. The Kalman filter is a particular form of this general least squares filter.

- The **smoothing step**: by which all the measurements are reprocessed after the last measurement has been made and the filtering step has been completed.

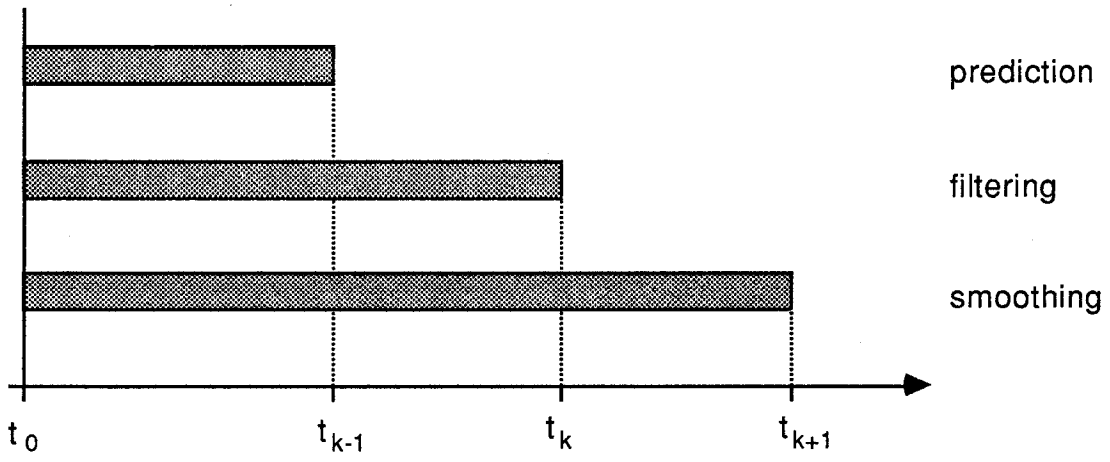


Figure 3.1: The Concepts of Prediction, Filtering and Smoothing

The estimation procedure is described in this chapter, and each new concept is introduced with as few definitions as possible in order to highlight the actual preconditions necessary for that concept. Assuming an initial state vector and a kinematic model, the **prediction** of parameters is first presented, independently of any other consideration. This prediction is then merged with

measurements and the derivation of an **optimal filter** based on least squares principles is given. The term optimal is precisely defined in this context: the filtered parameters must be unbiased and the trace of their covariance matrix must be a minimum. This algorithm is called the "Basic Least Squares Filter" and its most important features can be demonstrated using only an estimate of the state vector at one epoch, a kinematic model and a set of observations at a later epoch. The equivalence of the **Kalman form** is then demonstrated. Relations involving the filtered estimates are presented in order to permit comparisons with results from classical least squares adjustments to be made.

The concept of **smoothing** is perhaps the most difficult to grasp. However, it suffices to say that the initial estimate may be somehow improved by future measurements and the equations for smoothing can be obtained in a rigorous manner, as a logical extension of the previous developments. As the parameters at different epochs are related through a kinematic model, the filtered estimate is no longer optimal once subsequent measurements become available. A new estimate that includes the contributions of the later measurements can be computed. Improving previous estimates via a new measurement is therefore smoothing and, as in the case of filtering, is referred to as optimal smoothing if optimal estimation methods are employed. An important point to note is that a filter always has the potential of smoothing, even if this capability is not explicitly stated.

Several derivations for the Kalman filter equations have been given in terms of the least squares principle. KRAKIWSKY (1975) proved the equivalence of Kalman filtering and sequential least squares for the combined case. In a similar manner CROSS (1983) derives the Kalman expressions for the parametric case. The problem is treated as a minimisation under constraints with the Lagrange method of undetermined multipliers (see examples in Appendices A and B). This procedure is easily applicable to the combined case. However, additional unknowns are introduced, the "correlative constants", that need to be eliminated later on. Both authors consider two epochs of measurements, which is rather confusing as the contribution of the observations at the first epoch can be considered as a part of the *a priori* knowledge of the initial state vector. It is hoped that the approach chosen here will make the relationship between filtering and smoothing more apparent.

All the processes described are recursive in nature, and hence only two epochs t_0 and t are considered. This reduces the use of indices and hence contributes to the clarity of the mathematical expressions. An extension to more epochs is, except perhaps for smoothing, obvious.

3.2. PREDICTION

3.2.1. Initial State Vector

The state vector contains some of the parameters needed to describe the system. For kinematic applications its main components are usually position and velocity. Additional parameters can be included, depending on the task at hand and the measurement type available. At time t_0, the state vector and its covariance matrix are represented by $\hat{\mathbf{x}}_0$ and $\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}$. Knowledge of the initial state vector may have been derived from measurements, or as a result of a previous (filter) adjustment.

3.2.2. Kinematic Model

The kinematic behaviour of the system between two epochs is modelled by a relation between the system parameters at two different epochs. The linear form (3-1) is valid for all the kinematic models considered in this thesis.

$$\Phi \mathbf{x}_0 + \mathbf{w} = \mathbf{x} \qquad (3\text{-}1)$$

where
- $\mathbf{x}_0$ is the state vector at time t_0
- $\mathbf{x}$ is the state vector at time t
- Φ is the transition matrix
- $\mathbf{w}$ is the system noise

The derivation of appropriate transition matrices, for various assumptions regarding the GPS receiver motion and system noise, will be carried out in Chapter 4. In most other cases, a linear form can be obtained, at least as an approximation. The noise term $\mathbf{w}$ accounts for unmodelled forces affecting the system behaviour. In other words, the noise term represents the unpredicted (unpredictable?) component of the motion. $\mathbf{w}$ is assumed to have mean zero, and mainly affects the stochastic component of the kinematic model, through its covariance $\mathbf{Q}_{\mathbf{ww}}$. Because (unmodelled) forces are

responsible for the noise, the term "dynamic" model is also commonly used, but "kinematic" is preferred here because the underlying forces governing the behaviour of the system never appear explicitly in the model.

3.2.3. Parametric Expression

Implicit in the kinematic model is the capability to estimate the state vector at future epochs. If $\mathbf{x}$ has been estimated at time t_0, as $\mathbf{w}$ has mean zero, the predicted value at time t is simply:

$$\tilde{\mathbf{x}} = \Phi \hat{\mathbf{x}}_0 \tag{3-2}$$

(3-2) is the **functional** component of the kinematic model. According to the law of propagation of variances, and assuming that $\hat{\mathbf{x}}_0$ and $\mathbf{w}$ are independent, the covariance matrix of the predicted estimate is:

$$\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}} = \Phi \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\Phi^T + \mathbf{Q}_{\mathbf{ww}} \tag{3-3}$$

(3-3) is therefore the **stochastic** component of the kinematic model. The predicted estimate of the state vector can be regarded as an observation of its true value, with random error $\mathbf{v}_x$ and weight $\mathbf{P}_{\tilde{\mathbf{x}}} = \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}$:

$$\begin{aligned} \tilde{\mathbf{x}} - \mathbf{v}_x &= \mathbf{x} \\ \tilde{\mathbf{x}} - \mathbf{v}_x &= \mathring{\mathbf{x}} + \delta\mathbf{x} \qquad \text{with weight matrix} \quad \mathbf{P}_{\tilde{\mathbf{x}}} \end{aligned} \tag{3-4}$$

Although it is very tempting, and usually possible, to choose $\mathring{\mathbf{x}} = \tilde{\mathbf{x}}$, thus reducing relation (3-4) to $\delta\mathbf{x} = -\mathbf{v}_x$, such a substitution will only be considered as a particular case of a more general presentation. Nevertheless, relation (3-4) can be slightly modified:

$$\begin{aligned} \mathring{\mathbf{x}} + \delta\tilde{\mathbf{x}} - \mathbf{v}_x &= \mathring{\mathbf{x}} + \delta\mathbf{x} \\ \delta\tilde{\mathbf{x}} - \mathbf{v}_x &= \delta\mathbf{x} \qquad \text{with weight matrix} \quad \mathbf{P}_{\tilde{\mathbf{x}}} \end{aligned} \tag{3-5}$$

3.3. FILTERING

3.3.1. Measurement Model

Considering the parametric case after linearisation, a vector of observations l, with covariance matrix $\mathbf{Q}_{ll}$, is related to the parameters by:

$$\mathring{\mathbf{v}} - \mathbf{v} = \mathbf{A}\,\delta\mathbf{x} \qquad \text{with weight matrix } \mathbf{P} \qquad (2\text{-}5)$$
$$\mathring{\mathbf{v}} - \mathbf{v} = \mathbf{A}\,(\mathbf{x} - \mathring{\mathbf{x}})$$

The weight matrix depends entirely on the stochastic model of the observations and is defined by $\mathbf{P} = \mathbf{Q}_{ll}^{-1}$. The matrix $\mathbf{P}$ is not necessarily diagonal. When measurements are made at time t , a predicted state vector is computed. This is often the best estimate of the parameters available at this time. Hence, it is logical to linearise the measurement model around the predicted state, and compute the predicted residuals $\tilde{\mathbf{v}}$:

$$\tilde{\mathbf{v}} = l - f(\tilde{\mathbf{x}})$$

Because the measurements and the predicted state are involved, the predicted residuals are of great interest for the monitoring of the filter. Shortcomings in either the measurement or the kinematic models will show up as systematic errors in $\tilde{\mathbf{v}}$. Furthermore, predicted residuals are available prior to the adjustment step, and this enables the testing of the measurements and the rejection of outliers. Ideally, the predicted residuals, also referred to as "innovation sequence", should not be correlated between epochs, and should be normally distributed with mean zero. If this is not true, some measurements or modelling assumptions must be held in suspicion. Some authors, for example TEUNISSEN & SALZMANN (1988), base the performance analysis of Kalman filters entirely on the inspection of the innovation sequence, using a number of statistical tests. The test for zero mean is the most widespread in practice. The relation (2-5) can be modified to include the predicted state, rather than the *a priori* state vector used for linearisation, as the reference for state vector increments.

$$\tilde{\mathbf{v}} - \mathbf{v} = \mathbf{A}\,(\mathbf{x} - \tilde{\mathbf{x}})$$
$$\tilde{\mathbf{v}} - \mathbf{v} = \mathbf{A}\,\delta\mathbf{x} \qquad \text{with weight matrix } \mathbf{P}$$

In this last expression, the increment vector $\delta\mathbf{x}$ is related to the predicted state: $\delta\mathbf{x} = (\mathbf{x} - \tilde{\mathbf{x}})$. This situation is often encountered in practice, but it is preferable to make a clear distinction here between $\tilde{\mathbf{x}}$ and $\mathring{\mathbf{x}}$, because they represent different concepts and have different stochastic behaviours, even if the parameters contained in both state vectors are identical. Relation (2-5) is therefore the more general expression.

3.3.2. Basic Least Squares Filter

The requirements for a least squares filter can now be explicitly stated. Both the predicted state (through the kinematic model) and the measurements can be fully described by their functional and stochastic models, a total of four models. Equations (2-5) and (3-5) can be grouped into a system of parametric equations:

$$\begin{bmatrix} \delta\tilde{\mathbf{x}} \\ \mathring{\mathbf{v}} \end{bmatrix} - \begin{bmatrix} \mathbf{v}_x \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A} \end{bmatrix} \delta\mathbf{x} \qquad \text{with weight matrix: } \begin{bmatrix} \mathbf{P}_{\tilde{x}} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{bmatrix} \tag{3-6}$$

This system can be rewritten in the usual form of parametric equations (2-5), by simply replacing each matrix such that:

$$\mathring{\mathbf{v}}^* - \mathbf{v}^* = \mathbf{A}^* \delta\mathbf{x} \qquad \text{with weight matrix: } \mathbf{P}^*$$

The solution follows immediately from (2-6):

$$\delta\hat{\mathbf{x}} = (\mathbf{A}^{*T}\mathbf{P}^*\mathbf{A}^*)^{-1}\,\mathbf{A}^{*T}\mathbf{P}^*\mathring{\mathbf{v}}^*$$

Back-substituting the matrices $\mathbf{A}^*$, $\mathbf{P}^*$ and $\mathring{\mathbf{v}}^*$ through their original expressions in (3-6) leads to:

$$\delta\hat{\mathbf{x}} = (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,(\mathbf{P}_{\tilde{x}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}) \tag{3-7a}$$

This is now the expression for the filtered estimate. The minimisation of the extended quadratic form $\mathbf{v}^T\mathbf{P}\,\mathbf{v} + \delta\mathbf{x}^T\mathbf{P}_{\tilde{x}}\,\delta\mathbf{x}$ would yield identical results, as demonstrated in Appendix B, and could therefore replace the adjunction of the "pseudo-observation" (3-5) to the measurement (2-5). For most practical applications, the measurement model is linearised around the predicted state.

Explicitly, $\mathring{\mathbf{x}}$ is chosen equal to $\tilde{\mathbf{x}}$, and this implies $\delta\tilde{\mathbf{x}} = \mathbf{0}$ and $\mathring{\mathbf{v}} = \tilde{\mathbf{v}}$. Thus, the most commonly used expression for the computation of the filtered state vector, though not the most general, is:

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} + (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\,\tilde{\mathbf{v}} \qquad (3\text{-}7b)$$

An important restriction on the use of this simplified formula is in the case of solution iterations. It is straightforward to select the set of parameters $\mathring{\mathbf{x}} = \tilde{\mathbf{x}}$ for the initial linearisation. For subsequent iterations however, $\mathring{\mathbf{x}}$ is changed and the complete relation (3-7a) should be used. This considerably reduces the efficiency of the filtering algorithm, hence iterations should be avoided as far as possible. Indeed, if the predicted state vector is so biased that iterations are necessary, the kinematic model itself should be reassessed. Such problems may also be caused by an inappropriate selection of the parameters, as for example when a small change in a measurement leads to a large change in the state vector, which causes large changes in the measurement model (partial derivatives in $\mathbf{A}$ and observation residual $\mathring{\mathbf{v}}$). Such instabilities have always been a problem in surveying (for example, due to poor network geometry), and filtering is not a miracle cure for deficiencies in the measurement design.

The covariance matrix of the filtered estimate can be computed by applying the law of propagation of variances to (3-7a), assuming the stochastic independence of $\tilde{\mathbf{x}}$ and $\boldsymbol{l}$:

$$\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \qquad (3\text{-}8)$$

The covariance matrix of the residuals is obtained by the same method:

$$\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}} = \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\,\mathbf{A}^T \qquad (3\text{-}9)$$

This expression is identical to the standard parametric case, as stated in (2-8), though the definition of $\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$ is somewhat different. All derivations are presented in more detail in Appendix C. In this form, the algorithm is called the **Bayes filter**. The contributions of the predicted and observed components are easily identified. Comparing (3-7a) with the standard solution for the parametric case (2-6):

$$\delta\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}$$

it appears that the filtered estimate is fairly easy to compute. It is sufficient to add the weight of the predicted estimate to the classical normal matrix $\mathbf{A}^T\mathbf{P}\mathbf{A}$, add the weighted predicted estimate to the constant term $\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}$, and proceed as usual. The drawback of this procedure is that it always requires the inversion of a normal matrix of dimension $u \times u$, where u is the size of the state vector, even if only one observation is added. This computational problem makes this filter inappropriate for many applications.

3.3.3. The Kalman Form

The Kalman form is obtained by using a matrix identity described in many textbooks of linear algebra, or more specifically in the estimation literature, for example LIEBELT (1967):

$$(\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} = (\mathbf{P}_{\tilde{\mathbf{x}}}^{-1} - \mathbf{P}_{\tilde{\mathbf{x}}}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{P}_{\tilde{\mathbf{x}}}^{-1}\mathbf{A}^T + \mathbf{P}^{-1})^{-1}\,\mathbf{A}\mathbf{P}_{\tilde{\mathbf{x}}}^{-1}) \quad (3\text{-}10)$$

The validity of this transformation can be proven by showing that the product of the contents of the brackets on each side in the identity (assuming that all inverted matrices exist). Substituting in the Bayes expression (3-7a) and (3-5), the filtered estimate of $\mathbf{x}$ becomes:

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} + \mathbf{K}(\mathring{\mathbf{v}} - \mathbf{A}\,\delta\tilde{\mathbf{x}}) \quad (3\text{-}11a)$$

where $$\mathbf{K} = \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}\,\mathbf{A}^T(\mathbf{A}\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1} \quad (3\text{-}12)$$

Appendix C gives further details of this derivation. At first sight, this new expression for $\hat{\mathbf{x}}$ appears more complicated than the Bayes form. However, the only matrix to invert has dimension $n' \times n'$, where n' is the number of new observations. This is an important computational advantage. In addition, if the observations are linearised around the predicted state, (3-11a) reduces to:

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} + \mathbf{K}\tilde{\mathbf{v}} \quad (3\text{-}11b)$$

$\mathbf{K}$ is referred to as the **Kalman gain matrix**. The term is used to describe the function of feedback loops in electrical engineering, where the output of a system is used to improve the separation of the input signal from the noise

within the system. In the expression (3-11b), **K** represents the propagation of the "unexpected" part of the new measurements into an improvement of the predicted estimate.

Example: n independent measurements l of a distance x are available. The standard deviation of each measurement is σ. The least squares estimate $\hat{x}_n$ is the mean $\bar{l}_n$ of all measurements, with variance σ^2/n. The expected value of a new similar measurement l_{n+1} is $\bar{l}_n$. The new least squares estimate of the distance is the mean of the n+1 measurements:

$$\hat{x}_{n+1} = \frac{\sum_{i=1}^{n+1} l_i}{n+1} = \frac{\sum_{i=1}^{n} l_i + l_{n+1}}{n+1} = \frac{(n\cdot\bar{l}_n) + l_{n+1}}{n+1} =$$

$$= \frac{(n\cdot\bar{l}_n) + \bar{l}_n + l_{n+1} - \bar{l}_n}{n+1} = \frac{(n+1)\bar{l}_n}{n+1} + \frac{l_{n+1} - \bar{l}_n}{n+1} =$$

$$= \bar{l}_n + \frac{1}{n+1}(l_{n+1} - \bar{l}_n) = \tilde{x}_{n+1} + \frac{1}{n+1}(l_{n+1} - \tilde{x}_{n+1})$$

The last expression may be regarded as the Kalman form, where 1/n+1 is the gain matrix. This example is simply a degenerate case, completely described by:

- the measurement model $\mathbf{A} = 1,\ \mathbf{Q}_{ll} = \sigma^2$
- the kinematic model $\Phi = 1,\ \mathbf{Q}_{ww} = 0.$

This can be checked by replacing the terms in (3-2), (3-3), (3-11) and (3-12).

Using (3-11a), the covariance matrix of the Kalman filtered estimate can be computed. According to the law of propagation of variances and the assumed stochastic independence of $\tilde{\mathbf{x}}$ and l:

$$\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{I} - \mathbf{KA})\ \mathbf{Q}_{\tilde{x}\tilde{x}}\ (\mathbf{I} - \mathbf{KA})^T + \mathbf{K}\,\mathbf{Q}_{ll}\,\mathbf{K}^T \qquad (3\text{-}13)$$

This expression can be simplified to:

$$\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{I} - \mathbf{KA})\ \mathbf{Q}_{\tilde{x}\tilde{x}} \qquad (3\text{-}14)$$

See Appendix C for full details. This latter expression is very convenient but according to BUCY & JOSEPH (1968), the use of (3-13) may be preferable because of a reduced sensitivity to numerical instabilities. In the absence of numerical problems, the Kalman and Bayes filters yield identical results.

The covariance matrix of the vector of residuals derived in Appendix C for the Kalman form:

$$\mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T$$

is identical to the expression (3-9) already obtained for the Bayes filter.

3.4. FILTER TYPES AND TERMINOLOGY

During the last few years, a large amount of effort has been devoted to developing and refining parameter estimation methods involving more than one filter. This has resulted in a gradual widening of the filtering concepts, reflected by a steady expansion of the vocabulary associated with Kalman filtering. At present, the trend is to optimise the estimation through a combination of different filters.

3.4.1. Single Filters

Some of the attributes applicable to single filters are discussed below.

The **extended** filter accounts for nonlinearities in the relations between measurements and/or parameters by linearising the functional models of the measurements and motion, usually through a first degree Taylor series expansion. This case presents no difficulty to surveyors, who have always dealt with non-linear relations between directions, distances and Cartesian coordinates. Indeed, all filters used in surveying are "extended", and this attribute is often not explicitly mentioned.

A filter is referred to as being **augmented** when the size of the state vector is increased by the modelling of an additional relation. For example, including a bias common to several measurements in the state vector is one way of dealing with physical correlations between these measurements. This approach is particularly efficient when state parameters or measurements are

correlated between epochs. In GPS data processing, the estimation of satellite clock terms can be regarded as an augmentation of the state vector, permitting a correlation between receiver clocks and positions to be modelled.

When relations involving only parameters are included, the filter is referred to as **constrained**. This is generally achieved by adding pseudo-measurements, although the "decomposition in least squares" provides an alternate and more general procedure (see Section 2.4.5). Some authors consider measurements as constraints, as they limit the freedom of the parameters, by contributing to their determination. However, in this thesis, the term "constraints" is restricted to relations between parameters only.

An interesting feature of sequential estimation procedures is their ability to provide information concerning the quality of the data as they are being processed. This can be used to improve the *a priori* stochastic models of the subsequent measurements and/or movements. Procedures having this capability are called **adaptive**. Typically, the standard deviation of the measurements and/or the system noise can be rescaled according to the size of the adjusted observation residuals. As elegant as this approach may appear, it presents a danger: after several epochs of good agreement between prediction and measurements, the expected discrepancies may become very small and preclude the filter from reacting correctly to an actual change in the movement. Therefore, it is of paramount importance to specify a minimum for the standard deviations of the measurements and system noise, to prevent the filter from becoming over-optimistic. The same problem is also encountered in classical surveying: if the sum of the 3 measured angles of a triangle is exactly 180 degrees, this does not mean that all measurements are perfect, and it is preferable to use precision estimates based on experience to compute the precision of the survey.

3.4.2. Combination of Filters

In modern navigation, several sensors are used to collect information about the movement, for example GPS and an inertial sensor. Of course, all the data collected can be directed to the same filter. However, this results in limited computational efficiency, as the matrices involved become generally large and sparse. Therefore, attempts have been made to divide the computation

burden amongst several smaller filters. Filters can be combined in a number of ways:

- horizontal: a data-set is processed through one filter and the results are sent to another filter. This concept is associated with **cascaded** filtering procedures, where the role of the pre-filter may be to merely compress the measurement data, thus reducing the volume of input for the main filter.

- vertical: different data-sets are processed through different filters. For example, where each filter is associated with a different sensor, the process is known as **distributed** filtering. However, to ensure that a unique solution is finally obtained, the solutions from the different filters must be combined. Therefore, a strictly vertical filter configuration cannot be optimal.

The most promising developments are a combination of both the "horizontal" and "vertical" filter designs. KERR (1985) proposed a decentralised filtering structure in which several smaller filters process data from separate navigation subsystems. The outputs of these local filters are then combined by a "collating" filter. However, no mathematical basis for that filter was given. CARLSON (1987) established a sound basis for this general filtering architecture, known as **federated**, which perfectly illustrates the idea of an

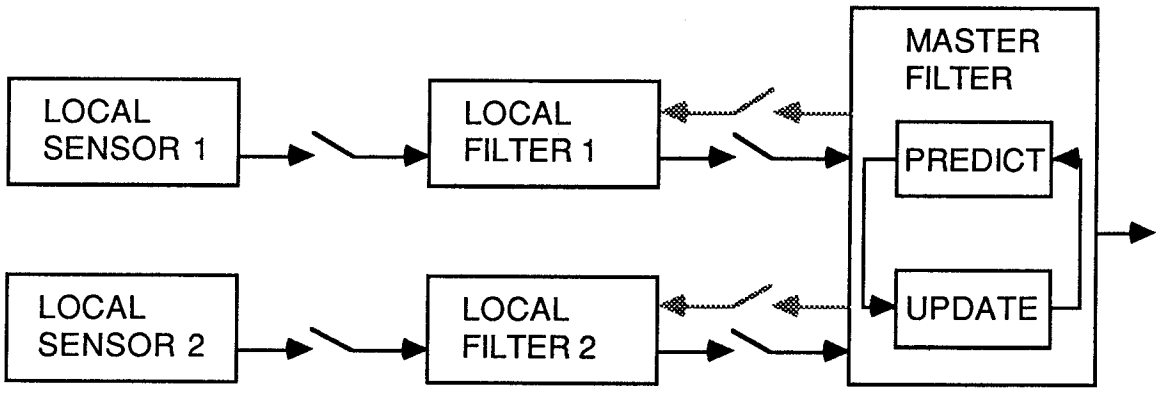


Figure 3.2: Federated Filter Architecture. (adapted from CARLSON, 1987)

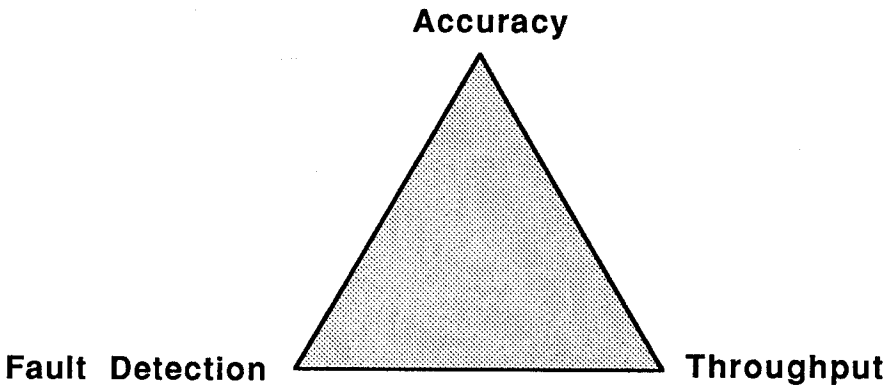
optimal interaction between the **local** filters and the **master** filter. The design presently favoured consists of implementing relative basic models in the local filters and more sophisticated ones in the master filter, to account for long term

effects. In the following, only two sensors are considered. Indeed, the federated filter can be regarded as a combination of cascaded and distributed filters, as shown in Figure 3.3.

	single filter	multi-filter
single sensor	classical	cascaded
multi-sensor	distributed	federated

Figure 3.3: Combination of Filters

The main feature of the federated structure is the feedback from the master filter to the local filters, represented by grey lines in Figure 3.2. The type of feedback will determine whether the "memory" of the filter is contained in the local filters or in the master filter. Roughly, the problem is to select the weight matrix to associate with the predicted parameters sent to the local filters. With all the terms equal to zero, there is a loss of information and precision, but on the other hand, an error in the results from the master filter, due for example to erroneous previous measurements, is not propagated further. The other extreme is to always send the latest weight matrix computed by the master filter, however, the increase in precision is balanced by a higher sensitivity to errors. In some situations, dividing all terms in the weight matrix by a factor of 2, before sending them to the local filter, proved to be a successful compromise (CARLSON, 1988).



3.4: Criteria for an Optimal Filter Architecture

The computational efficiency of the filter, or its "throughput", essentially depends on how many cycles the local filters run before the master filter is activated, and hence on how much data compression is achieved. Various criteria may be used to judge the appropriateness of a combination of filters, including the accuracy of the results, the capacity to detect a fault and the computation load. All criteria cannot be satisfied simultaneously, and the weight given to any of these criteria may depend on the problem at hand. The trade-off can be represented as the choice of a position within the triangle in Figure 3.4.

3.4.3. Parallel Filtering

The simultaneous processing of the same data-set with different filters, each using different models, permits different criteria to be monitored. **Parallel** filtering appeared first in GPS data processing to handle the task of selecting the set of integer ambiguities. Such a multi-hypothesis testing scheme has been implemented in the software package NOVAS, using a **bank** of Kalman filters (WANLESS & LACHAPELLE, 1988), following a proposition by HWANG & BROWN (1985). It is possible that this resolution task will soon be monitored in real-time. Data collection could then stop as soon as a reliable set of integer cycle ambiguities is obtained, resulting in both savings in time and the need for field assessment of the data. Hence, filtering techniques may also prove useful for static GPS surveys. The concept of parallel filtering, although in a different form, is also applied in the **censored** filter proposed for the monitoring of the integrity of the GPS system (McBURNEY & BROWN, 1988). When the measurement residuals from a satellite have a mean consistently and significantly different from zero, the quality of the data from that satellite becomes suspect. Different subsets of the measurement data are filtered in parallel, and the difference in position between the various options is monitored. When the dispersion of the results exceeds a predefined threshhold, action is initiated, generally consisting of the rejection of data from the suspect satellite.

By processing the same data with several federated filters in parallel, each with a different sharing of the filter memory between the local filters and the master filter, different options could be obtained simultaneously. For example, there may be an advantage in using a filter optimal with respect to fault detection and recovery along with a filter designed for optimal accuracy: the

most accurate solution would be output, as long as the other filter doesn't detect any fault. Otherwise, the accurate filter would be re-initialised, using the state vector from the fault tolerant filter.

Theoretically, there are very few restrictions in the combination of these concepts. At this time however, very little practical experience has been gained, and published results are scarce.

Although GPS is the only sensor presently considered in the software package DYNAMO, its overall structure would enable a more complicated filter architecture to be implemented in the future.

3.5. SMOOTHING

3.5.1. Initial State

The true value of the initial state vector is generally not known. However, the smoothed estimate of $\mathbf{x}_0$ should not be too different from the filtered one, which has a known stochastic behaviour described by the covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}$. A certain imprecision in the initial estimate must be tolerated. This error can be seen as having the potential for minimisation. If all the elements of $\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}$ are zero, the notion of smoothing becomes meaningless because later measurements cannot have any effect on $\mathbf{x}_0$.

The filtered estimate $\hat{\mathbf{x}}_0$ can be considered as an *a priori* estimate of the true value of $\mathbf{x}_0$, with a random error $\mathbf{v}_{\mathbf{x}_0}$. This relation can be written in the form of an observation equation, similar to (3-4):

$$\hat{\mathbf{x}}_0 - \mathbf{v}_{\mathbf{x}_0} = \mathbf{x}_0 \qquad \text{with weight matrix } \mathbf{P}_{\hat{\mathbf{x}}_0} \qquad (3\text{-}15)$$

$\mathbf{v}_{\mathbf{x}_0}$ corresponds to the definition of a residual (see Section 2.4.1), as it is the difference between a vector of observations and a functional model of the observations. However, in this particular case of (pseudo-)observations of the parameters, the functional model is the identity, and (3-15) may also be written in terms of state vector increments $\delta\mathbf{x}_0$:

$$\hat{\mathbf{x}}_0 + \delta\mathbf{x}_0 = \mathbf{x}_0 \qquad \text{with weight matrix } \mathbf{P}_{\hat{\mathbf{x}}_0} \qquad (3\text{-}16)$$

Thus, $\mathbf{v}_{x_0}$ and $\delta\mathbf{x}_0$ have the same magnitude, but opposite signs. In many cases, for example when a quadratic form is considered, both terms can be used. Depending on the context however, the conceptual difference between a residual and a state vector increment should not be forgotten. To avoid any confusion with the filtered estimate $\hat{\mathbf{x}}_0$ already obtained, the smoothed estimate of $\mathbf{x}_0$ is denoted as $\hat{\hat{\mathbf{x}}}_0$.

3.5.2. Kinematic Model

The concept of a non-perfect initial estimate also has an influence on the parametric formulation of the kinematic model. Discrepancies between prediction and observations are not only due to errors in the kinematic model, but also to errors in the estimate at time t_0. Thus the kinematic relation is applied to unknown state vectors at both epochs t_0 and t:

$$\Phi\,\mathbf{x}_0 + \mathbf{w} = \mathbf{x} \qquad \text{with weight matrix } \mathbf{P}_w$$

This relation is already linear, however, to be consistent with the measurement model it must be written in terms of state vector increments for the epoch t. For the epoch t_0, the full value of the parameters can be considered, because the state vector $\mathbf{x}_0$ is only involved in linear relations:

$$\begin{aligned} \Phi\,\mathbf{x}_0 + \mathbf{w} &= (\overset{\circ}{\mathbf{x}} + \delta\mathbf{x}) \\ \overset{\circ}{\mathbf{x}} - \mathbf{w} &= \Phi\,\mathbf{x}_0 - \delta\mathbf{x} \qquad \text{with weight matrix } \mathbf{P}_w \end{aligned} \qquad (3\text{-}17)$$

Between epochs t_0 and t, the state vector is affected by the noise $\mathbf{w}$. The weight attributed to the observation (3-17) only depends on the covariance of the system noise acting between the epochs t_0 and t : $\mathbf{P}_w = \mathbf{Q}_{ww}^{-1}$.

3.5.3. Least Squares Filter / Smoother: Statement

According to the least squares principle, a minimum value for several quadratic forms should be obtained. The correspondence of the filtered and smoothed estimate of $\mathbf{x}_0$ requires that $\hat{\hat{\mathbf{v}}}_{x_0}\mathbf{P}_{\hat{\mathbf{x}}_0}\hat{\hat{\mathbf{v}}}_{x_0}$ (or $\delta\hat{\hat{\mathbf{x}}}_0{}^{T}\mathbf{P}_{\hat{\mathbf{x}}_0}\,\delta\hat{\hat{\mathbf{x}}}_0$) be a minimum, see (3-15) or (3-16). On the other hand, for an optimal adjustment of the observations, a solution minimising $\hat{\mathbf{v}}^T\mathbf{P}\hat{\mathbf{v}}$ is desirable. Finally, for an

optimal fit of the change in the state vector with the kinematic model, it is necessary to keep $\hat{\mathbf{w}}^T\mathbf{P}_w\hat{\mathbf{w}}$ as small as possible.

Generally, all three quadratic forms cannot attain their absolute minimum simultaneously, as each of the conditions sets other requirements on the adjustment. The optimal solution can be **defined** as requiring the overall minimisation of the quadratic forms:

$$\Sigma v^T P v = \mathbf{v}_{x_0}{}^T \mathbf{P}_{\hat{x}_0} \mathbf{v}_{x_0} + \mathbf{v}^T \mathbf{P}\, \mathbf{v} + \mathbf{w}^T \mathbf{P}_w\, \mathbf{w} \rightarrow \text{minimum} \qquad (3\text{-}18)$$

Many features of such a filter can be described without actually processing any data, even without defining the mathematical form that the filter will take. The following discussion is brought up already here to make it clear that a filter does not **know** anything about least squares "optimality", but always outputs what the analyst **interprets** as optimal.

From (3-18), it appears that the solution is some form of average between initial conditions, observations and kinematic model. The relative magnitude of the elements in the weight matrices $\mathbf{P}_{\hat{x}_0}$, $\mathbf{P}$ and $\mathbf{P}_w$ is of paramount importance. If a filter is to give realistic estimates, the relative weighting of the components (a process sometimes referred to as "tuning") deserves particular attention. The response of a filter for some extreme cases of weighting can now be described:

(1) $\mathbf{P}_{\hat{x}_0} = \mathbf{0}$, that is, no usable initial state vector. This implies that the state vector cannot be estimated until observations allow for a first determination. This case is of little practical importance since an approximate state vector is generally required for the linearisation of the measurement model. If the trace of $\mathbf{P}_{\hat{x}_0}$ tends towards infinity, more observations will be required to account for possible biases in the initial parameters. Obviously, the effect of $\mathbf{P}_{\hat{x}_0}$ diminishes as observations are added over time. After a "settlement period", the behaviour of a filter is mainly influenced by $\mathbf{P}$ and $\mathbf{P}_w$.

(2) $\mathbf{P} = \mathbf{0}$, a filter could operate just as well without any measurements, and its role is reduced to predicting future states according to the initial states and the assumed kinematics. If the trace of $\mathbf{P}$ tends towards infinity, the

output depends solely on the observations and reflects the full range of their variations.

(3) $\mathbf{P_w} = \mathbf{0}$, renders the kinematic model useless and future states become unpredictable. If the trace of $\mathbf{P_w}$ tends towards infinity, the predicted estimate is not considered free of errors, as the term $\Phi \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} \Phi^T$ remains in $\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}$. However, the filter cannot accommodate unforeseen deviations from the assumed kinematics and tries, no matter what the observations indicate, to bring the estimate back to their predefined trajectory.

There is an obvious analogy here with geodetic networks in which different types of observables are combined, for example directions and distances. To obtain small residuals on the distances, the surveyor can attribute a large weight to them. However, the ratio of the residuals *a posteriori / a priori* may indicate that some observables have been outrageously favoured. The adjustment should then be repeated with rescaled weight matrices, until consistent ratios are obtained for the different types of observables. It must be emphasised that such a procedure is meaningful only if there is an adequate number of redundant observations. In any case, it is preferable to use weights based on experience, if available, rather than try to compensate apparently discrepant variances computed from very few redundant measurements.

Similarly, after some observations have been processed, an imbalance in the weighting of the measurement and kinematic models may become evident. A test of the ratio of the *a posteriori / a priori* mean errors for the adjusted state increments as well as for the measurement residuals can be included within the filter algorithm, and the weights modified accordingly. Filters having this feature are sometimes called "adaptive" (see Section 3.4.1).

3.5.4. Least Squares Filter / Smoother: Solution

Compared to the Basic Least Squares Filter, a new observation is introduced: the observation of the new estimate of the initial state vector. In addition, the relation involving the kinematic model is modified. All these observations can be fully described by their functional and stochastical models and combined in a system of parametric equations, using (3-15), (3-17) and (2-5):

$$\begin{array}{lllll} \hat{\mathbf{x}}_0 - \mathbf{v}_{\mathbf{x}_0} = \mathbf{x}_0 & & & & \text{weight } \mathbf{P}_{\hat{\mathbf{x}}_0} \\ \overset{\circ}{\mathbf{x}} - \mathbf{w} = \Phi\,\mathbf{x}_0 & - \delta\mathbf{x} & & & \text{weight } \mathbf{P}_{\mathbf{w}} \\ \overset{\circ}{\mathbf{v}} - \mathbf{v} = & & \mathbf{A}\delta\mathbf{x} & & \text{weight } \mathbf{P} \end{array}$$

The system can be set up in matrix form:

$$\begin{bmatrix} \hat{\mathbf{x}}_0 \\ \overset{\circ}{\mathbf{x}} \\ \overset{\circ}{\mathbf{v}} \end{bmatrix} - \begin{bmatrix} \mathbf{v}_{\mathbf{x}_0} \\ \mathbf{w} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \Phi & -\mathbf{I} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \delta\mathbf{x} \end{bmatrix} \qquad \text{weight} \begin{bmatrix} \mathbf{P}_{\hat{\mathbf{x}}_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P} \end{bmatrix} \quad (3\text{-}19)$$

The matrices can be replaced by single elements, yielding the standard form of a system of parametric equations, compare with (2-5):

$$\overset{\circ}{\mathbf{v}}^* - \mathbf{v}^* = \mathbf{A}^* \delta\mathbf{x}^* \qquad \text{weight } \mathbf{P}^*$$

The solution follows immediately from (2-6):

$$\delta\hat{\mathbf{x}}^* = (\mathbf{A}^{*\mathrm{T}}\mathbf{P}^*\mathbf{A}^*)^{-1}\,\mathbf{A}^{*\mathrm{T}}\mathbf{P}^*\overset{\circ}{\mathbf{v}}^*$$

After back-substitution of $\delta\hat{\mathbf{x}}^*$, $\mathbf{A}^*$, $\mathbf{P}^*$ and $\overset{\circ}{\mathbf{v}}^*$:

$$\begin{bmatrix} \hat{\hat{\mathbf{x}}}_0 \\ \delta\hat{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\hat{\mathbf{x}}_0} + \Phi^{\mathrm{T}}\mathbf{P}_{\mathbf{w}}\Phi & -\Phi^{\mathrm{T}}\mathbf{P}_{\mathbf{w}} \\ -\mathbf{P}_{\mathbf{w}}\Phi & \mathbf{P}_{\mathbf{w}} + \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{A} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{P}_{\hat{\mathbf{x}}_0}\hat{\mathbf{x}}_0 + \Phi^{\mathrm{T}}\mathbf{P}_{\mathbf{w}}\overset{\circ}{\mathbf{x}} \\ -\mathbf{P}_{\mathbf{w}}\overset{\circ}{\mathbf{x}} + \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\overset{\circ}{\mathbf{v}} \end{bmatrix} \quad (3\text{-}20)$$

The inversion of the (symmetric) normal matrix can be avoided if the normal equations are combined properly. The equivalent derivation based on the explicit minimisation of the quadratic form is given in Appendix D, leading to the new estimates of $\mathbf{x}_0$ and $\mathbf{x}$:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\,(\overset{\circ}{\mathbf{v}} - \mathbf{A}\,\delta\hat{\mathbf{x}}) \qquad (3\text{-}21)$$

$$\hat{\mathbf{x}} = \overset{\circ}{\mathbf{x}} + (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}\,(\mathbf{P}_{\tilde{\mathbf{x}}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\overset{\circ}{\mathbf{v}}) \qquad (3\text{-}7a)$$

This last relation is the filtered estimate already obtained using the Basic Least Squares Filter (Section 3.3.2). The formulae for the covariance of the parameters and residuals derived for the Basic Least Squares Filter still hold, as does the extension to the Kalman form. The only interest in the above

development is therefore the new estimate of the initial state vector (3-21). The expression for the smoothed estimate of $\mathbf{x}_0$ and its covariance can now be written in a simpler fashion:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{H}\,\hat{\mathbf{v}} \tag{3-22}$$

$$\mathbf{Q}_{\hat{\hat{\mathbf{x}}}_0\hat{\hat{\mathbf{x}}}_0} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} - \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T \tag{3-23}$$

where $$\mathbf{H} = \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\Phi^T\mathbf{A}^T\mathbf{Q}_{ll}^{-1} \tag{3-24}$$

The derivation of the expression for the covariance matrix, assuming the stochastic independence of $\hat{\mathbf{x}}_0$ and $\hat{\mathbf{v}}$, is given in Appendix D.

The improvement of previous parameter estimates through the use of later measurement has now been demonstrated. The integration of the concept of smoothing within a least squares algorithm is therefore established. In this presentation only two epochs have been considered. The extension to more epochs presents no additional conceptual difficulty, but nevertheless requires some data manipulations.

3.5.5. Extension to More Epochs, R-T-S Algorithm

The formulae obtained for the smoothed estimate and associated covariance of $\mathbf{x}_0$ depend explicitly on the measurements at time t. This need not be the case, and the difference between the predicted and filtered values of $\mathbf{x}$ -- the so-called improvement -- may have another cause. Appendix E shows that the smoothed estimate and its covariance can be written as:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{J}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \tag{3-25}$$

$$\mathbf{Q}_{\hat{\hat{\mathbf{x}}}_0\hat{\hat{\mathbf{x}}}_0} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} + \mathbf{J}\,(\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} - \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}})\,\mathbf{J}^T \tag{3-26}$$

where $$\mathbf{J} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\Phi^T\,\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1} \tag{3-27}$$

These expressions exhibit an interesting feature: they are completely free of terms explicitly related to the observations; l, $\mathbf{Q}_{ll}$, $\mathbf{P}$ and $\mathbf{A}$ have been completely removed. Indeed, this result is quite intuitive: once the filtered estimate has been obtained, the contribution of the observations at time t is already included in the filtered state vector and its covariance matrix. Hence,

there is no need for the observations to be used again in further computations. In fact, when the smoothed estimate at time t_o and its covariance are computed, only the difference between the predicted $\tilde{\mathbf{x}}$ and the filtered $\hat{\mathbf{x}}$ values at time t matters, and how this improvement of the state vector and its covariance matrix was achieved is irrelevant. Hence, it is easy to generalise the filtering and smoothing processes over more than two epochs. It suffices to consider that the improvement of the state vector at time t includes the contributions of all the measurements realised at subsequent epochs. In fact, an equivalent improvement may also have been obtained with observations at time t exclusively. The smoothed values at time t can replace the filtered ones in the smoothing equations (3-25) and (3-26). This is particularly obvious for the last epoch of measurements, which cannot be improved by further observations. Thus, the smoothed estimate is made equal to the filtered one. From then on, proceeding back in time, the other smoothed estimates are computed recursively, until the first epoch is reached. The process is described by:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{J}\,(\hat{\hat{\mathbf{x}}} - \tilde{\mathbf{x}}) \qquad (3\text{-}28)$$

$$\mathbf{Q}_{\hat{\hat{\mathbf{x}}}_0\hat{\hat{\mathbf{x}}}_0} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} + \mathbf{J}\,(\mathbf{Q}_{\hat{\hat{\mathbf{x}}}\hat{\hat{\mathbf{x}}}} - \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}})\,\mathbf{J}^T \qquad (3\text{-}29)$$

where $$\mathbf{J} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\Phi^T\,\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1} \qquad (3\text{-}27)$$

This recursive smoothing algorithm requires the storage of quite a lot of information during the filtering process:

- the predicted state vector at all epochs,
- the covariance matrix of the predicted state vector,
- the filtered state vector at all epochs,
- the covariance matrix of the filtered state vector,
- the transition matrix between all update points.

This method of smoothing is often called R-T-S, from the names of the authors responsible for its first description, RAUCH, TUNG & STRIEBEL (1965). This smoothing algorithm appears very naturally in the development of the Least Squares Filter/Smoother. However, all the data required for smoothing may be obtained from the Basic Least Squares Filter (Section 3.3) -- whether in the

Kalman form or not -- even though no provision was made for smoothing at that stage.

3.5.6. Modified Bryson-Frazier Algorithm

Other smoothing algorithms have been developed, although their theoretical foundations are generally not as straightforward. The original form of a commonly used procedure was published by BRYSON & FRAZIER (1962). The "Modified Bryson-Frazier" algorithm (MB-F) can be derived by formulating the R-T-S equations with a new set of variables ($\mathbf{z}$, $\mathbf{Q}_{zz}$) where $\mathbf{z}$ is the adjoint state vector and $\mathbf{Q}_{zz}$ is the adjoint covariance matrix. Details of the derivation can be found in BIERMAN (1973). Basically, $\mathbf{z}$ and $\mathbf{Q}_{zz}$ are propagated over epochs in a similar fashion as the state vector and its covariance matrix in a Kalman filter. The recursive formulae are found in WANG (1983), and merely translated here into our notation.

(3-30)

$$\hat{\hat{\mathbf{x}}} = \hat{\mathbf{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\tilde{\mathbf{z}} \qquad \text{smoothing}$$
$$\mathbf{Q}_{\hat{\hat{x}}\hat{\hat{x}}} = \mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{z}\tilde{z}}\,\mathbf{Q}_{\hat{x}\hat{x}}$$

$$\tilde{\mathbf{z}}_0 = \Phi^T \hat{\mathbf{z}} \qquad \text{prediction}$$
$$\mathbf{Q}_{\tilde{z}_0\tilde{z}_0} = \Phi^T \mathbf{Q}_{\hat{z}\hat{z}}\,\Phi$$

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{K}\mathbf{A})^T \tilde{\mathbf{z}} + \mathbf{A}^T\mathbf{D}^{-1}(\mathbf{A}\delta\tilde{\mathbf{x}} - \mathring{\mathbf{v}}) \qquad \text{update}$$
$$\mathbf{Q}_{\hat{z}\hat{z}} = (\mathbf{I} - \mathbf{K}\mathbf{A})^T \mathbf{Q}_{\tilde{z}\tilde{z}}\,(\mathbf{I} - \mathbf{K}\mathbf{A}) + \mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}$$

$$\text{with } \mathbf{D} = (\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T + \mathbf{Q}_{ll})$$

The smoothing starts at the end of the data, by initialising the adjoint state vector for the last epoch of measurement update.

$$\hat{\mathbf{z}} = \mathbf{A}^T\mathbf{D}^{-1}(\mathbf{A}\delta\tilde{\mathbf{x}} - \mathring{\mathbf{v}}) \qquad \text{initialisation}$$
$$\mathbf{Q}_{\hat{z}\hat{z}} = \mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}$$

In this form, the MB-F algorithm is unnecessarily complicated. Therefore, a rearrangement is proposed in the next section.

3.5.7. MB-F Rearranged

The form of the recursive algorithm presented in Section 3.5.6 can be simplified. First of all, the various components can be rearranged to ensure that epoch subscripts are consistent within each step and reflect the true order of the computations. In addition, the predicted adjoint state vector, rather than the filtered one, can be initialised.

(3-31)

$$\begin{aligned}
\tilde{\mathbf{z}} &= \mathbf{0} && \text{initialisation}\\
\mathbf{Q}_{\tilde{z}\tilde{z}} &= \mathbf{0}\\
\\
\hat{\mathring{\mathbf{x}}} &= \hat{\mathbf{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\tilde{\mathbf{z}} && \text{smoothing}\\
\mathbf{Q}_{\hat{\mathring{x}}\hat{\mathring{x}}} &= \mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{z}\tilde{z}}\,\mathbf{Q}_{\hat{x}\hat{x}}\\
\\
\hat{\mathbf{z}} &= (\mathbf{I} - \mathbf{KA})^{\mathsf{T}}\,\tilde{\mathbf{z}} + \mathbf{A}^{\mathsf{T}}\mathbf{D}^{-1}\,(\mathbf{A}\delta\tilde{\mathbf{x}} - \mathring{\mathbf{v}}) && \text{update}\\
\mathbf{Q}_{\hat{z}\hat{z}} &= (\mathbf{I} - \mathbf{KA})^{\mathsf{T}}\,\mathbf{Q}_{\tilde{z}\tilde{z}}\,(\mathbf{I} - \mathbf{KA}) + \mathbf{A}^{\mathsf{T}}\mathbf{D}^{-1}\mathbf{A}\\
\\
\tilde{\mathbf{z}}_0 &= \Phi^{\mathsf{T}}\,\hat{\mathbf{z}} && \text{prediction}\\
\mathbf{Q}_{\tilde{z}_0\tilde{z}_0} &= \Phi^{\mathsf{T}}\,\mathbf{Q}_{\hat{z}\hat{z}}\,\Phi\\
\\
\text{with } \mathbf{D} &= (\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^{\mathsf{T}} + \mathbf{Q}_{ll})
\end{aligned}$$

It is not obvious in this algorithm, but the update step involves predicted residuals and their covariance matrix. This can be shown starting from the definition of the residuals, where l is a vector of observations and $f(\mathbf{x})$ its functional model.

$$\begin{aligned}
\tilde{\mathbf{v}} &= l - f(\tilde{\mathbf{x}})\\
\tilde{\mathbf{v}} &= l - (f(\mathring{\mathbf{x}}) + \mathbf{A}(\tilde{\mathbf{x}} - \mathring{\mathbf{x}}))\\
\tilde{\mathbf{v}} &= (l - f(\mathring{\mathbf{x}})) - \mathbf{A}(\tilde{\mathbf{x}} - \mathring{\mathbf{x}})\\
\tilde{\mathbf{v}} &= \mathring{\mathbf{v}} - \mathbf{A}\delta\tilde{\mathbf{x}}
\end{aligned} \qquad (3\text{-}32)$$

The predicted state vector and the vector of measurements are stochastically independent, hence the covariance marix of the predicted residuals is:

$$\mathbf{Q}_{\tilde{v}\tilde{v}} = \mathbf{Q}_{ll} + \mathbf{A}^{\mathsf{T}}\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A} = \mathbf{D} \qquad (3\text{-}33)$$

These identities can be inserted in the update step.

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{KA})^T \tilde{\mathbf{z}} - \mathbf{A}^T \mathbf{Q}_{\tilde{v}\tilde{v}}^{-1} \tilde{\mathbf{v}} \quad (3\text{-}34)$$

$$\mathbf{Q}_{\hat{z}\hat{z}} = (\mathbf{I} - \mathbf{KA})^T \mathbf{Q}_{\tilde{z}\tilde{z}} (\mathbf{I} - \mathbf{KA}) + \mathbf{A}^T \mathbf{Q}_{\tilde{v}\tilde{v}}^{-1} \mathbf{A} \quad (3\text{-}35)$$

The fundamental and very interesting idea of the MB-F algorithm is now more explicit: a function of the observation residuals is propagated backwards through the filtered states.

3.5.8. The Adjoint State Vector

Minimum attention has been paid to the content of the adjoint state vector. Indeed, it is difficult to find an intuitive interpretation, as its relation with the proper state vector is far from obvious. Nevertheless, an important pitfall should be pointed out, in order to avoid an erroneous simplification: the adjoint state vector cannot be considered as an increment of the state vector. This is easily demonstrated by looking at the units of the quantities contained in each vector. A state vector containing a position, a velocity and an acceleration in one direction is considered.

$$\mathbf{x} = \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} \sim \begin{bmatrix} m \\ m/sec \\ m/sec^2 \end{bmatrix}$$

According to the smoothing step, consistency of the units is achieved with:

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \sim \begin{bmatrix} 1/m \\ sec/m \\ sec^2/m \end{bmatrix}$$

because the adjoint state vector is premultiplied by the covariance matrix of the state vector. This unusual set of units indicates that little insight can be gained from the values contained in **z**. Therefore, the adjoint state vector will be regarded as a mere computational convenience.

3.6. NEW SMOOTHING ALGORITHMS

3.6.1. Using Filtered Residuals

In spite of the elegant concept inherent in the Modified Bryson-Frazier algorithm, its present formulation has an important drawback: the update step requires the predicted state vector and its covariance matrix, like the R-T-S algorithm. Intuitively, it must be possible to deal only with filtered estimates when smoothing. Therefore, a change in the formulation of the update step is proposed. The idea is to replace the predicted observation residuals with the filtered ones. (3-32) is recalled and a similar relation is developed for filtered residuals.

$$\tilde{\mathbf{v}} = \overset{\circ}{\mathbf{v}} - \mathbf{A}\delta\tilde{\mathbf{x}} \qquad (3\text{-}32)$$

$$\hat{\mathbf{v}} = \overset{\circ}{\mathbf{v}} - \mathbf{A}\delta\hat{\mathbf{x}} \qquad (3\text{-}36)$$

Thus $$\tilde{\mathbf{v}} = \hat{\mathbf{v}} + \mathbf{A}\delta\hat{\mathbf{x}} - \mathbf{A}\delta\tilde{\mathbf{x}} = \hat{\mathbf{v}} + \mathbf{A}(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \qquad (3\text{-}37)$$

The following identities are proven in Appendix E, using the relations between predicted and filtered quantities established for the Bayes or Kalman filter.

$$\mathbf{A}^T\mathbf{D}^{-1}\,\tilde{\mathbf{v}} = \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} \qquad (3\text{-}38)$$

$$\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A} = \mathbf{A}^T\mathbf{P}\,\mathbf{Q}_{\hat{v}\hat{v}}\,\mathbf{P}\mathbf{A} \qquad (3\text{-}39)$$

The smoothing algorithm is therefore rewritten:

(3-40)

initialisation

$$\tilde{\mathbf{z}} = \mathbf{0}$$

$$\mathbf{Q}_{\tilde{z}\tilde{z}} = \mathbf{0}$$

smoothing

$$\hat{\tilde{\mathbf{x}}} = \hat{\mathbf{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\tilde{\mathbf{z}}$$

$$\mathbf{Q}_{\hat{\tilde{x}}\hat{\tilde{x}}} = \mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{z}\tilde{z}}\,\mathbf{Q}_{\hat{x}\hat{x}}$$

update

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{K}\mathbf{A})^T\tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}}$$

$$\mathbf{Q}_{\hat{z}\hat{z}} = (\mathbf{I} - \mathbf{K}\mathbf{A})^T\mathbf{Q}_{\tilde{z}\tilde{z}}\,(\mathbf{I} - \mathbf{K}\mathbf{A}) + \mathbf{A}^T\mathbf{P}\,\mathbf{Q}_{\hat{v}\hat{v}}\,\mathbf{P}\mathbf{A}$$

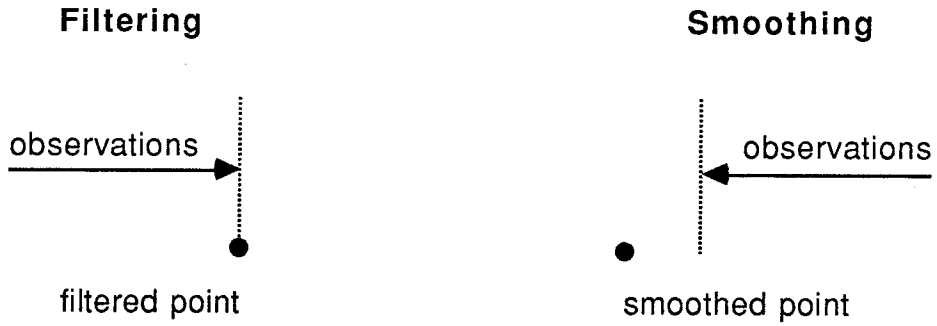
prediction

$$\tilde{\mathbf{z}}_0 = \Phi^T\hat{\mathbf{z}}$$

$$\mathbf{Q}_{\tilde{z}_0\tilde{z}_0} = \Phi^T\mathbf{Q}_{\hat{z}\hat{z}}\,\Phi$$

3.6.2. Using Smoothed Residuals

A close look at the algorithm (3-40) reveals that the smoothing step is performed before the update of the adjoint state vector. This is because only subsequent measurements have an influence on the difference between the filtered and the smoothed estimates, as measurements at epoch t have already been accounted for in the filtered estimate at time t when filtering. The comparison of the two sketches in Figure 3.5 illustrates that concept.



3.5: Observations Included in the Filtering and Smoothing Adjustment Steps

Therefore, smoothed residuals, instead of filtered ones, can be computed and used in the update step. It suffices to proceed with some substitutions. (3-36) is recalled and a similar relation is developed for smoothed residuals.

$$\hat{\mathbf{v}} = \overset{\circ}{\mathbf{v}} - \mathbf{A}\delta\hat{\mathbf{x}} \tag{3-36}$$

$$\hat{\hat{\mathbf{v}}} = \overset{\circ}{\mathbf{v}} - \mathbf{A}\delta\hat{\hat{\mathbf{x}}} \tag{3-41}$$

Thus $$\hat{\mathbf{v}} = \hat{\hat{\mathbf{v}}} + \mathbf{A}\delta\hat{\hat{\mathbf{x}}} - \mathbf{A}\delta\hat{\mathbf{x}} = \hat{\hat{\mathbf{v}}} + \mathbf{A}(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) \tag{3-42}$$

The substitution of the filtered residual in the update step of (3-40) using (3-42) leads to a interesting simplification of the expression for the updated adjoint state vector:

$$\hat{\mathbf{z}} = \tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}} \tag{3-43}$$

Details of the derivation are given in Appendix E. The inclusion of the covariance matrix of the smoothed state vector follows the same principles, and yields:

$$\mathbf{Q}_{\hat{z}\hat{z}} = \mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{A}^T\mathbf{K}^T\mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{Q}_{\tilde{z}\tilde{z}}\mathbf{K}\mathbf{A} + \mathbf{A}^T\mathbf{P}\mathbf{Q}_{\hat{\hat{v}}\hat{\hat{v}}}\mathbf{P}\mathbf{A} \qquad (3\text{-}44)$$

The simplification for the covariance matrix is not as dramatic as for the adjoint state vector. This is due to the correlation between the predicted adjoint state vector and the smoothed state vector. Comparing with the propagation of variances in (3-43):

$$\mathbf{Q}_{\hat{z}\hat{z}} = \mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{A}^T\mathbf{P}\mathbf{Q}_{\hat{v}\tilde{z}} - \mathbf{Q}_{\tilde{z}\hat{v}}\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P}\mathbf{Q}_{\hat{\hat{v}}\hat{\hat{v}}}\mathbf{P}\mathbf{A} \qquad (3\text{-}45)$$

The cross-covariance terms can be set equal:

$$\mathbf{Q}_{\tilde{z}\hat{v}}\mathbf{P}\mathbf{A} = \mathbf{Q}_{\tilde{z}\tilde{z}}\mathbf{K}\mathbf{A}$$

Recalling the alternative expression (C-9) for the Kalman gain matrix:

$$\mathbf{Q}_{\tilde{z}\hat{v}}\mathbf{P}\mathbf{A} = \mathbf{Q}_{\tilde{z}\tilde{z}}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{P}\mathbf{A}$$

Hence $$\mathbf{Q}_{\tilde{z}\hat{v}} = \mathbf{Q}_{\tilde{z}\tilde{z}}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T \qquad (3\text{-}46)$$

There is nothing strange in $\tilde{\mathbf{z}}$ and $\hat{\hat{\mathbf{v}}}$ being correlated, as the predicted adjoint state vector $\tilde{\mathbf{z}}$ is used to compute the smoothed state vector, from which the vector of smoothed residuals $\hat{\hat{\mathbf{v}}}$ is obtained. The covariance matrix of the adjoint state vector can be computed as:

$$\mathbf{Q}_{\hat{z}\hat{z}} = \mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{A}^T\mathbf{P}\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{Q}_{\tilde{z}\tilde{z}}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P}\mathbf{Q}_{\hat{\hat{v}}\hat{\hat{v}}}\mathbf{P}\mathbf{A} \qquad (3\text{-}47)$$

This expression looks more complicated than previous ones, but every single component is readily available. Furthermore, the third term on the right hand side is simply the transpose of the previous term.

The recursive smoothing algorithm becomes:

(3-48)

$\tilde{\mathbf{z}}$	=	$\mathbf{0}$	initialisation
$\mathbf{Q}_{\tilde{z}\tilde{z}}$	=	$\mathbf{0}$	
$\hat{\hat{\mathbf{x}}}$	=	$\hat{\mathbf{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\tilde{\mathbf{z}}$	smoothing
$\mathbf{Q}_{\hat{\hat{x}}\hat{\hat{x}}}$	=	$\mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{z}\tilde{z}}\,\mathbf{Q}_{\hat{x}\hat{x}}$	
$\hat{\mathbf{z}}$	=	$\tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}}$	update
$\mathbf{Q}_{\hat{z}\hat{z}}$	=	$\mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{C} - \mathbf{C}^T + \mathbf{A}^T\mathbf{P}\,\mathbf{Q}_{\hat{\hat{v}}\hat{\hat{v}}}\,\mathbf{P}\mathbf{A}$	
		where $\mathbf{C} = \mathbf{A}^T\mathbf{P}\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{z}\tilde{z}}$	
$\tilde{\mathbf{z}}_0$	=	$\Phi^T\,\hat{\mathbf{z}}$	prediction
$\mathbf{Q}_{\tilde{z}_0\tilde{z}_0}$	=	$\Phi^T\,\mathbf{Q}_{\hat{z}\hat{z}}\,\Phi$	

The use of the smoothed residuals in the smoothing algorithm presents a definite advantage. As in any least squares estimation procedure, the assessment of the results is based on testing the actual residuals against their stochastic model. Hence the calculation of $\hat{\hat{\mathbf{v}}}$ and $\mathbf{Q}_{\hat{\hat{v}}\hat{\hat{v}}}$ required in the proposed update step is by no means a waste of effort: it provides the basis for a sound evaluation of the final results. This recursive smoothing algorithm yields the same results as the more common R-T-S algorithm, as proven in Appendix D.

3.6.3. Storage Requirements

The data required for the R-T-S algorithm has already been listed in Section 3.5.5. For the new recursive smoothing algorithm using smoothed residuals, the following information must be stored during the filtering process:

- the observations at all epochs,
- the filtered state vector at all epochs,
- the covariance of the filtered state vector,
- the transition matrix between all update points,
- the design and weight matrix of the observations at all epochs.

Since optimal smoothing is essentially a post-mission procedure, off-line storage of these quantities is sufficient. In many cases, and this is also valid

for the R-T-S algorithm, the transition matrix can be recomputed during the smoothing process at each epoch. With respect to storage, the main difference between the two algorithms is that the new smoothing algorithm requires the raw data instead of the predicted values of the state vector. The storage of the observation data, rather than the predicted state information, offers an important advantage: the filtering process can be repeated after the completion of the survey, because all the required data are available. In addition, this allows for some changes in the allocation of the weights. Furthermore, the design and the weight matrix of the observations can generally be recomputed when smoothing, using a set of predefined parameters. For a computer implementation, three data sources are necessary:

- the **observation file** -- observations at all epochs,
- the **filtered data** -- filtered state vector and its covariance matrix at all epochs,
- the **definition file** -- a set of constants for the calculation of the design, weight and transition matrices.

It should be noted that the definition file is already required for filtering. Indeed, even the volume of storage compares favourably with the R-T-S algorithm. The data-flows through the filter and the smoother are detailed and compared in Chapter 5. The recursive algorithm using the smoothed residuals was selected for DYNAMO.

4. NAVIGATION MODELS

In this chapter the mathematical expressions for the kinematic and measurement models suitable for navigation applications introduced in the previous chapters are derived. Emphasis is on the presentation of a variety of kinematic models that can be used, because such modelling is less familiar to surveyors. For this reason, some elements of the theory of random processes are introduced. The following presentation is restricted to the navigation concepts actually applied in DYNAMO. For a more in depth treatment of least squares filtering and the theory of random processes, the reader is referred to LIEBELT (1967) or GELB (1974).

4.1. RANDOM SEQUENCES AND FUNCTIONS

4.1.1. Traditional Assumptions

Surveyors are accustomed to dealing with random variables. Traditionally, some assumptions are made regarding the errors affecting the measurements:

- **gross** errors can be eliminated,
- there are no **systematic** errors arising from improperly modelled observations.

Hence, only **random** errors are believed to remain in the data. They affect the measurements at discrete epochs, thus forming a **random sequence**. Additional assumptions are usually made concerning the measurement errors:

- their expectation is zero,
- their distribution is Gaussian,
- their standard deviation is known,
- they are independent.

All these assumptions are reasonable, as long as the functional model of the measurements is accurate. The analysis of the precision of a geodetic network is generally based on these assumptions. Hence a possible error in a particular observation does not affect any other observation. In Figure 4.1, measurement errors relative to their standard deviation (that is, normalised)

are plotted together with their Gaussian probability distribution function p(v). The random sequence is illustrated by the dots at each discrete epoch.

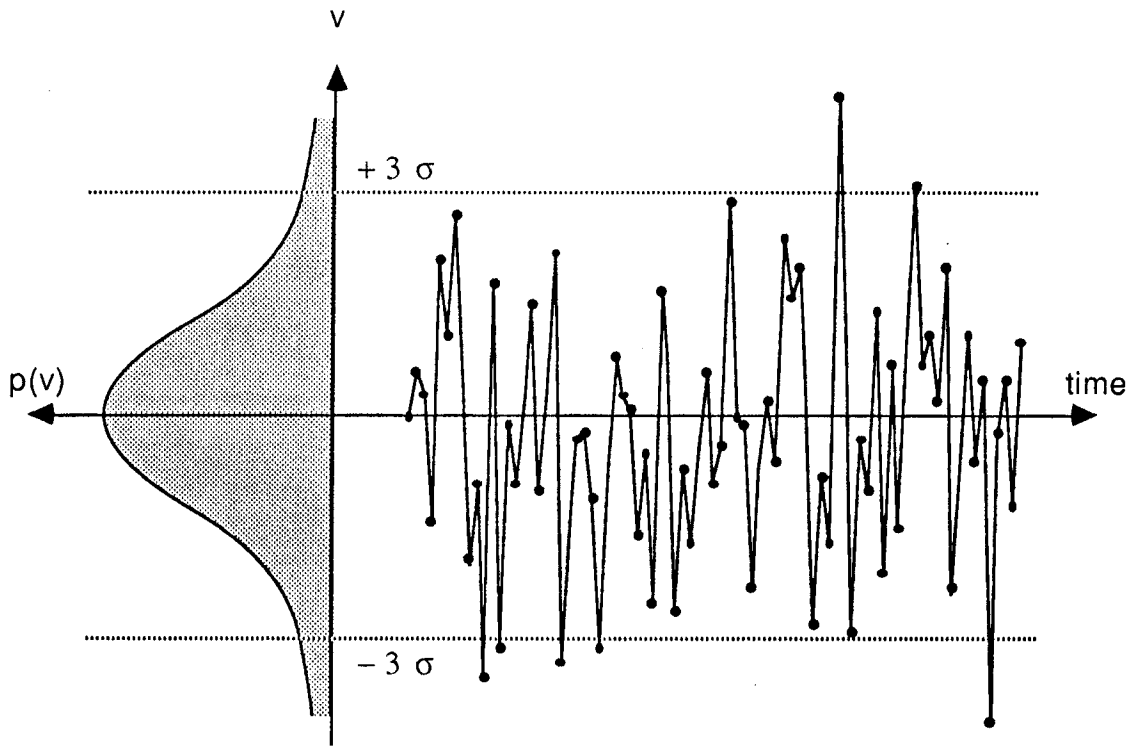


Figure 4.1: Gaussian Random Sequence / Noise

When errors in the measurement data larger than a certain threshold (for example > 3 σ) can be identified, the corresponding observations are removed from the data. In other words, the tails of the bell curve are truncated. However, for the sake of simplicity, the probability distribution function is generally still considered Gaussian.

The problem is more difficult when it comes to kinematic systems. If one over-estimates the speed of his car, he will arrive later than expected at the destination. Unlike errors in the azimuth and distance to a trig station, the errors in the speed and position of a vehicle, for example, are physically correlated. In the framework of Bayesian Least Squares (Section 2.4.6), this correlation between different parameters can be accounted for by introducing off-diagonal elements in the covariance matrix of the *a priori* estimates of the parameters. When filtering, such correlations must be included in the covariance matrix of the predicted parameters.

For a moving object, the uncertainty in position grows with time, unless a new measurement is made and a position fix obtained. This growing uncertainty is expressed as a noise that affects the position parameters, but which may also affect the velocity and acceleration terms. The noise simply reflects the fact that the motion cannot be perfectly modelled. However, unlike measurement errors that only exist at discrete epochs, the position of a moving object is continuously affected by noise, even if the measurements are made at discrete epochs. An example of random noise is illustrated by the continuous function in Figure 4.1. A proper understanding and modelling of this underlying continuous noise requires elements from the theory of **random processes**, which is an extension of the probability theory familiar to surveyors. However, the theory of random processes has tended to be presented from the perspective of electrical engineering. Many applications can be better described in the frequency domain and conversions into the time domain involve an extensive use of Fourier transformations. Such topics are covered in numerous textbooks (for example, PAPOULIS, 1962), and this section introduces some concepts at a rather intuitive level.

4.1.2. Useful Random Functions

A function is referred to as being random, as opposed to "deterministic", when its value is only known in a statistical sense, that is, only with a certain uncertainty. The probability distribution of the noise may be unknown or too complex to express. Thus the characteristics of the noise are defined through some model of how it is generated and how it evolves in the course of time. Often, the underlying probability distribution function could be determined, but is not explicitly required. The large acceptance of the Kalman filter is largely due to the fact that it allows for an expression of the system noise in the time domain, as it is perceived by the widest category of potential users. Amongst all the possible random functions, only a few are relevant for the cases of interest considered here. Essentially, 3 attributes are desirable:

a) **Stationarity** -- By definition, a random variable x is stationary if its probability density function $p(x)$ is time invariant. In other words, the probability of x taking a particular value x_1 is the same at all times, that is:

$$p(x_1,t_1) = p(x_1) \qquad \forall\, t_1$$

The most interesting consequences are reflected by the behaviour of the moments of the distribution:

$$\text{Expectation:} \quad E\{x(t)\} = \int_{-\infty}^{+\infty} x_1 \, p(x_1) \, dx_1 = E\{x\}$$

$$\text{Correlation:} \quad R_{xy}(t_1,t_2) = R_{xy}(t_2 - t_1)$$

The first moment is independent of time and the second moment only depends on the time elapsed between the two epochs. In the particular case where x and y are the same function, the term auto-correlation is used, otherwise it is referred to as cross-correlation. The covariance function is:

$$Q_{xy}(t_1,t_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 - E\{x\}) \, (y_2 - E\{y\}) \, p(x_1,t_1; y_2,t_2) \, dx_1 \, dy_2$$

$$= R_{xy}(t_1,t_2) - E\{x\} \, E\{y\}$$

If the expectation of the noise(s) is zero, the correlation is equal to the covariance:

$$Q_{xy}(t_1,t_2) = R_{xy}(t_1,t_2) - 0 = R_{xy}(t_1,t_2)$$

This is such a common assumption that, in practice, the word auto-correlation is used even when auto-covariance is actually meant.

b) **Linearity** -- It is very difficult to calculate the joint probability density function of samples of a linear combination of random variables from their probability density functions:

$$p(z_1,t_2 - t_1) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (ax_1 + by_2) \, p(x_1,t_1; y_2,t_2) \, dx_1 \, dy_2$$

In fact, it is only simple when the probability density functions of x and y are **Gaussian**. In this case, the distribution of the output z is Gaussian, even if

inputs are correlated. This property is commonly used in geodesy: the coordinates estimated from a linearised least squares adjustment have a Gaussian distribution, with mean and variance provided by the computations. This nice result is indeed a consequence of the Gaussian distribution assumed for the errors in the measurements, because the adjusted parameters are expressed as a linear combination of the measurements, for example through the expression (2-6):

$$\delta\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}$$

c) **Markovian** -- This attribute is used with respect to a random function if its value at a certain epoch depends on the behaviour of the function during a limited preceding interval. Thus, one could say that such a function has a limited memory. In the discrete formulation, the function is replaced by a sequence and the concept is easier to visualise. Markovian sequences are classified according to their order:

- order 0; the value of the variable at any epoch is independent of the value at any other epoch. In other words, even if the variable is estimated at one epoch, its value at a subsequent epoch cannot be predicted. Such a random sequence is called a "white sequence".

- order 1; the variable at one epoch is only related to its value at the last epoch. An example of a first order Markov process is the "random walk", where the variable is modelled as the sum of the values of a white sequence at different epochs.

- order 2; the variable at one epoch is related to its values at the two last epochs. This is the case of a sequence where the variable is modelled as the sum of the values of a random walk variable at different epochs, and generally referred to as "second order random walk".

- and so on ...

In the continuous formulation, the "white sequence" is replaced by a "white noise". For order 1, the variable at the last epoch is replaced by the first derivative of the random function. For order 2, the variable at the two last epochs is replaced by the first and second derivatives of the random function,

and so on. A first order Markov process is defined by the first order differential equation:

$$\dot{x} + \alpha_1(t)\, x = w \qquad (4\text{-}1)$$

where w is white noise. The second order Markov process is defined in a similar fashion:

$$\ddot{x} + 2\,\alpha_2(t)\,\dot{x} + \alpha_2^2(t)\, x = w \qquad (4\text{-}2)$$

and so on, see GELB (1974) for further details. White noise can be defined as: $x = w$, and hence can be thought of as a zero-th order Markov process. The random walk is obtained by setting: $\alpha_1 = 0$ in (4-1), and the second order random walk is obtained by setting: $\alpha_2 = 0$ in (4-2). Note that the derivative of a white noise function is not defined, because changes in the white noise value can occur within an infinitely small time interval, thus yielding infinite values for the derivative.

Each of the attributes presented above -- stationary, Gaussian, Markovian -- is independent of the others. For example, white noise does not necessarily have a Gaussian distribution. The processes described by (4-1) and (4-2) become respectively first and second order Gauss-Markov processes if the probability density function of the white noise w is Gaussian. In such a case, the probability density function of x is also Gaussian. In addition, these processes are stationary if the coefficients of the differential equation (that is, α_1 and α_2) are time invariant.

4.1.3. Examples of Random Processes

Figure 4.2 shows some examples of random processes frequently used for the modelling of time-varying physical phenomena. G-M stands for Gauss-Markov and ss for steady state.

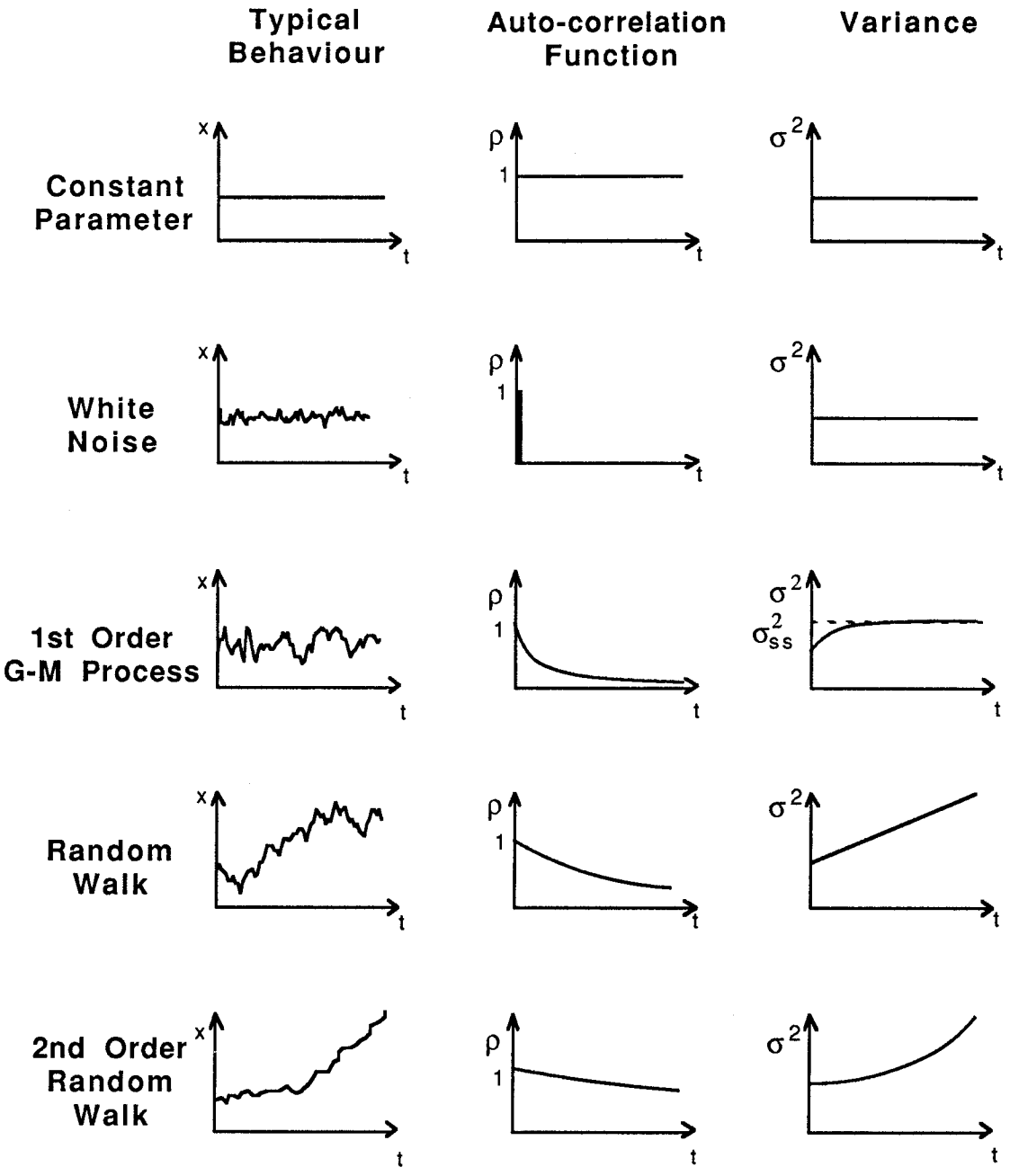


Figure 4.2: Some Random Processes. Adapted from GRANT (1988)

The variance sketches indicate the evolution of the uncertainty in the random variable in the absence of measurements. Three types of variance behaviour can be identified:

- **constant** -- this characteristic is exhibited by both the constant parameter and the white noise process. However, the nature of the variance in these two cases is very different, as indicated by the auto-correlation function. For a constant parameter, no noise is added to the

initial uncertainty. For a white noise process, the values of the parameter are uncorrelated over epochs. The difference appears clearly when a measurement is made: the determination of a constant parameter will reduce its variance at subsequent epochs, but such a reduction in the uncertainty does not take place in the case of a white noise process. However, it must be pointed out that a white noise process is completely unpredictable only if the variance is infinite. In other words, a parameter affected by a very small white noise may, under certain circumstances, be modelled as a constant. For example, repeated theodolite readings of a vertical angle generally form a white noise sequence, but the vertical angle is modelled as a constant in the network adjustment.

- **bounded growth** -- noise is only added until the system has reached a so-called "steady state". The variance level of the steady state and the time necessary to reach it depend on model parameters and must be matched to the problem at hand. This class of models is useful for systems which, although subject to certain fluctuations, are forced to find a new equilibrium. This happens when there is a "recall" effect as soon as the value of the parameter deviates too much from a nominal value, for example the height of an object floating on the ocean.

- **unbounded growth** -- noise is constantly added to the system, irrespective of its state. This model is generally appropriate for sea navigation over limited time spans. If the ship deviates from the intended track, no external force will generally limit the growth of the horizontal deviation.

4.2. KINEMATIC MODEL

4.2.1. Functional and Stochastic Components

In the previous chapter, the transition matrix Φ and the covariance matrix of the system noise $\mathbf{Q}_{ww}$ were introduced in the filter expressions. However, no attention was paid to the content of these matrices. The derivation of both matrices for a general case is given, and the details of the implementation in the software package DYNAMO are presented.

From the prediction model developed in Section 3 for the state vector and its covariance matrix, the role of Φ and $\mathbf{Q}_{ww}$ can be shown more precisely. The superscripts $\tilde{}$ and $\hat{}$ stand for "predicted" and "adjusted" respectively, and the subscript 0 denotes the previous epoch.

- the **functional model** expresses the relationship between position parameters at different epochs, usually adjacent ones. It is represented by the prediction equation (3-2):

$$\tilde{\mathbf{x}} = \Phi \hat{\mathbf{x}}_0$$

The functional model is therefore embodied in the transition matrix Φ. For each dimension, the content of the state vector can be specified by the **degree**, which is the highest derivative of a parameter involved. For example, the functional model of $\mathbf{x}$ has degree 2 when $\mathbf{x}$ contains parameters x, $\dot{x}$ and $\ddot{x}$.

- the **stochastic model** reflects the fact that the prediction cannot be assumed perfect. It is represented by the covariance matrix of the system noise $\mathbf{Q}_{ww}$, which is added to the covariance matrix due to the predicted term at each step, see equation (3-3):

$$\mathbf{Q}_{\tilde{x}\tilde{x}} = \Phi \mathbf{Q}_{\hat{x}_0\hat{x}_0}\Phi^T + \mathbf{Q}_{ww}$$

For each dimension, the stochastic model can be specified by the **order**, which is the derivative affected by white noise. For example, the stochastic model has order 2 when the acceleration is affected by white noise.

In this way, the inclusion of the kinematic model in the least squares adjustment is relatively simple. However, the formulation of the model and the derivation of the matrices are a major difficulty, as they depend on the time behaviour of the system parameters and assumptions made concerning the characteristics of the system noise.

Examples: In GPS data processing, a relationship between estimates at different epochs is often assumed. Three options of modelling a time dependence of a variable are possible in least squares adjustments:

(a) estimating the variable as a constant, that is, time invariant.

(b) treating all measurement epochs separately. This can be realised by considering an independent parameter for each epoch, or a white noise process with infinite variance. In both cases, an approximate value for the parameter is generally required for the linearisation of the measurement model, but no weight is attributed to this approximate value. If an estimation of these parameters is not required, they can be eliminated from the normal equations by various partitioning techniques (KING et al, 1987). The same result can be achieved by differencing observations and eliminating common biases.

(c) expressing the variable as a time dependent function with constant coefficients. The use of polynomials is widespread for modelling GPS satellite clock errors. However, for reasons already mentioned in Section 2.3 when comparing static and kinematic approaches for GPS navigation, the use of polynomials is not considered here.

(a) and (b) are examples of extreme kinematic models: constant and random distribution of the parameters at different epochs. Of course, the reality generally lies somewhere in between. Typically, the offset of a clock may be large, but the clock drift is contained within narrow limits. This knowledge can be expressed as an *a priori* guess of the auto-correlation function of the relevant parameter.

4.2.2 Classification of Kinematic Models

From various investigations, for example MEIJER (1983) and CHISHOLM (1987), it appears that the variety of filters devised for satellite navigation is limited. The mathematical algorithms used -- many of which are unfortunately unpublished -- are based on either the "constant velocity" or the "constant acceleration" assumption. In fact, the velocity or the acceleration are not assumed to be strictly time invariant, but affected by a white noise that has a small variance. In an attempt to classify the kinematic models systematically, the problem of parameter estimation in kinematic mode is presented in the general frame of system analysis.

The moving object is the system, the movement is considered the input and the measurements are regarded as the output. This description is not as abstract as it may seem because the movement of the object is actually perceived through the measurements. The relation between the input and the output is expressed by the measurement model whilst the system reaction to certain inputs, for example the change in position due to a change in acceleration, is expressed by the kinematic model. This can be characterised as follows:

Input = Position parameters
System = Vehicle
Output = Noisy measurements

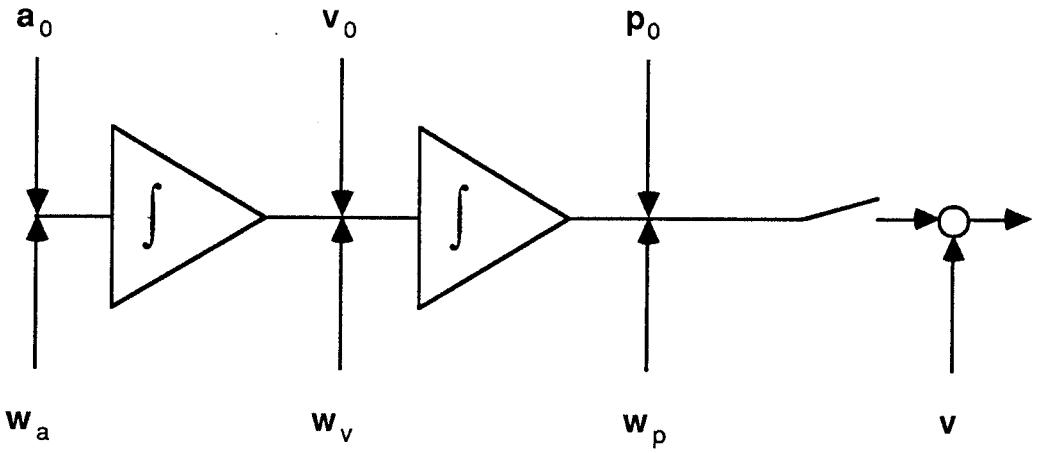


Figure 4.3: Generation of Measurements

Deterministic inputs, for example initial conditions, enter the system from the top: **a**, **v** and **p** are the acceleration, velocity and position respectively. Noise enters the system from the bottom: **w** is the system noise and **v** is the measurement noise. The switch is closed when a measurement is made.

We know:
- how the system reacts to certain inputs
- the measured output

We want:
- the input
- the uncertainty on the input

To estimate the parameters, a reversed system is therefore appropriate:

Input = Noisy measurements
System = Modelled vehicle
Output = Estimated parameters

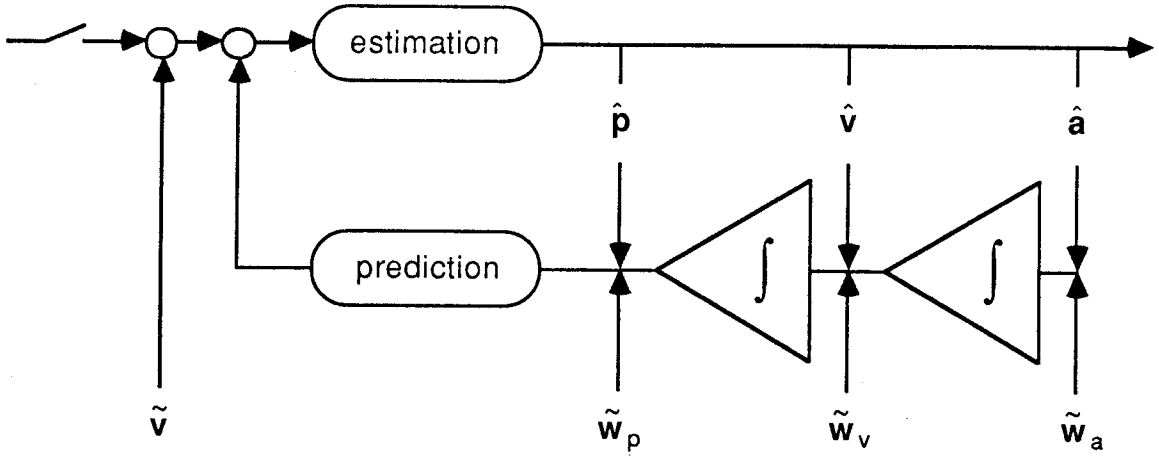


Figure 4.4: Estimation of Parameters

The feedback loop in Figure 4.4 shows how estimated parameters (superscript ^) are used to predict the next position, which is then merged with the new measurements. Predicted system and measurement noises (superscript ~), according to the predefined stochastic models, are added at each step.

A convention regarding terminology is adopted here: the noise entering the system is called **driving** noise. In the models considered in this thesis, the driving noise is always white, Gaussian, has zero mean and enters the system only through one derivative, called the order. Through integration, the driving noise is propagated to the lower derivatives. For example, the noise on the position is regarded as a consequence of the noises on the velocity and accelerations. The noise on the position may however no longer be white. See discussion following Figure 4.5. The total uncertainty in the system, that is, including the noise induced in lower derivatives through integration of the driving noise, is referred to as **system** or **process** noise.

Because the derivative of white noise is not defined, the degree cannot be larger than the order, though it can be smaller (see Figure 4.5). For example, considering that the acceleration is white noise, one may estimate it (order 2

and degree 2) or not (order 2 but degree 1). In both cases, the assumed uncertainty on the acceleration is propagated to the position, but the acceleration itself is not used to predict the next position. In other words, when navigating along a curve, the predicted position will be on the tangent to the curve at the last estimated position, and hence may be biased. For the acceleration to be propagated, the order must be increased (for example, order 3 and degree 2), so that the acceleration is no longer a white noise. In this case, it is possible to predict a position along a curve. However, a possible random error in the estimated acceleration will cause an error in the predicted velocity and position. The **dilemma** underlying all kinematic modelling is therefore: estimating higher position derivatives involves the risk of propagating **wrong** random values, whilst the opposite may introduce systematic errors (unmodelled biases).

The kinematic models considered may be classified according to both their degree and order. For example, considering the vector of parameters **x**:

Degree = 1 means that x and $\dot{x}$ are estimated (but not $\ddot{x}$ or higher derivatives)

Order = 3 means that $\dddot{x}$ is white noise

Degree \ Order	0	1	2	3
0	*	*	*	*
1	-	*	*	*
2	-	-	*	*

Figure 4.5: Possible Combinations of Degree and Order (*)

Another important criteria for the specification of the kinematic model is the intensity of the driving noise, given by the **standard deviation** of the white noise. Of course, a larger value for the driving noise will increase the uncertainty in the lower derivatives.

With a minimum of added complexity, other kinds of noises other than white may be considered. For example, many physical processes can be suitably

modelled using the first order Gauss-Markov process, which has auto-correlation function:

$$R_{xx} = \sigma^2 e^{-|\Delta t| / T} \tag{4-3}$$

where Δt is the time between two epochs and T is the **correlation time**, that is, the time interval for which the noise has an average auto-correlation of $e^{-1} = 0.368$. This exponentially correlated noise is often referred to as "coloured noise" (BIERMAN, 1977). A short correlation time allows for a large change in x between adjacent epochs, whilst a long correlation time indicates a strong correlation over time. The extreme cases are white noise for zero correlation time and random walk for infinite correlation time. Indeed, for any correlation time, coloured noise resembles a random walk for time intervals such that $|\Delta t| << T$, and white noise for time intervals such that $|\Delta t| >> T$. Coloured noise can be generated on x by considering white noise on $\dot{x}$ and the linear relation:

$$\dot{x} = -(1 / T)\, x + w \tag{4-4}$$

where w is the driving noise. (4-4) is a special case of (4-1), where $\alpha_1(t)$ is replaced by the time invariant coefficient $1 / T$. Hence, if white noise on a position parameter is to be replaced by coloured noise, it suffices to set the order to one and include the appropriate correlation time in the system equation. In summary, four parameters have been retained to specify the kinematic models used in DYNAMO:

- the degree,
- the order,
- the standard deviation of the driving noise,
- the correlation time.

Using this classification, the particular models presented in Section 4.2.1 are described by:

(a) constant bias: degree 0, order 1, standard deviation 0.
(b) random bias: degree 0, order 0, standard deviation ∞.

The correlation time is irrelevant in both cases, because there is no noise in (a) and the (white) driving noise in (b) is never integrated (order 0). A wavelike movement however may be approximated by a first order G-M process, for example:

coloured noise: degree 1, order 1, st. dev. 0.1m/sec, corr. time 10sec.

The choice of degree 1 implies that the velocity, although it is affected by white noise, will be estimated. The principles used to derive the corresponding transition and system noise covariance matrices are discussed in Section 4.2.3.

4.2.3 Transition and System Noise Matrices

The state vector **x** usually contains various parameters pertaining to a number of dimensions, such as coordinates, satellite and receiver clock offsets together with some of their time derivatives. The time behaviour in each dimension is usually expressed as a differential equation, and to illustrate this, a one-dimensional, third order model is considered. The coefficients of the time derivatives are assumed constant. For a situation that does not meet this criteria, it is generally possible to obtain another relation with constant coefficients, at least in an approximate manner, through further differentiation:

$$\dddot{x} + a_2\ddot{x} + a_1\dot{x} + a_0x - w = 0 \qquad (4\text{-}5)$$

One n-th order linear differential equation may be written in the form of n first order linear differential equations (GELB, 1974):

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \\ \dddot{x} \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ a_0 & a_1 & a_2 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \\ \dddot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix} \qquad (4\text{-}6)$$

This explicit form shows the content of the state vector **x**, and can be compressed using matrix notation:

$$\dot{\mathbf{x}} = \mathbf{F}\,\mathbf{x} + \mathbf{w} \tag{4-7}$$

In this matrix notation, the following development is also valid for multi-dimensional problems, see Appendix F for more details.

The matrix **F** describes the relations between the derivatives of a parameter, that is, the equations of motion. For many systems of interest, **w** can be described as a random forcing function, and expressed as a linear combination of white noise random processes **Gu**, where each element of the vector **u** is an independent random function and the matrix **G** describes the effect of **u** on the system parameters. In other words, **u** is the driving noise and **G** specifies where the driving noise enters the system. Hence, (4-7) can be expressed as:

$$\dot{\mathbf{x}} = \mathbf{F}\,\mathbf{x} + \mathbf{G}\mathbf{u} \tag{4-8}$$

An explicit form of (4-8), as well as the following derivation, are presented in full in Appendix F. Two assumptions are stated for the cases considered:

- **F** and **G** are time invariant, that is, the system is time invariant
- **u** is a vector of independent, white, Gaussian variables, with mean zero.

These assumptions are not as restrictive as they may appear at first. The condition that **F** and **G** are time invariant can always be achieved by breaking up the total time interval into pieces for which this condition is satisfied. For the adjustment of GPS observations, the system behaviour between two measurement epochs is relevant. In this context, **F** and **G** are considered as step functions that are constant between measurement epochs. The assumption that the forcing function **Gu** is random with mean zero over a certain time span is reasonable if all major error sources have been properly modelled. The solution of (4-8) is (LIEBELT, 1967):

$$\mathbf{x}(t) = \Phi(t,t_0)\,\mathbf{x}(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\,\mathbf{G}\mathbf{u}(\tau)\,d\tau \tag{4-9}$$

and the transition matrix satisfies the equations:

$$\Phi(t_0,t_0) = \mathbf{I}$$

and $$\dot{\Phi}(t,t_0) = \mathbf{F}\,\Phi(t,t_0)$$

LIEBELT (1967) presents a solution based on numerical integration:

$$\Phi(t,t_0) = \mathbf{I} + \mathbf{F}(t - t_0) + \frac{1}{2!}\mathbf{FF}(t - t_0)^2 + \frac{1}{3!}\mathbf{FFF}(t - t_0)^3 + \ldots \quad (4\text{-}10)$$

Some complex models require purely numerical solutions, but for those discussed here, an analytical expression can be derived for the transition matrix, and this results in considerable computer time savings. Often, the Taylor series expansions obtained by numerical integration can be replaced by elementary analytical functions, for example exponential or trigonometric forms (see Appendix F). An alternative analytical method, using the inverse Laplace transform, is outlined in SCHWARZ (1983).

Assuming the stochastic independence of $\mathbf{x}(t_0)$ and $\mathbf{u}(\tau)$, $\forall\ t_0 \le \tau \le t$, the propagation of the covariance matrix of the state vector is given by:

(4-11)

$$\mathbf{Q_{xx}}(t) = \Phi(t,t_0)\,\mathbf{Q_{xx}}(t_0)\,\Phi^T(t,t_0) + \int_{t_0}^{t}\int_{t_0}^{t} \Phi(t,\tau_1)\,\mathbf{G}\,E\{\mathbf{u}(\tau_1)\mathbf{u}(\tau_2)\}\,\mathbf{G}^T\,\Phi^T(t,\tau_2)\,d\tau_1\,d\tau_2$$

As all elements of $\mathbf{u}$ are white noise random processes with zero mean, the expectation of the product of the random values of $\mathbf{u}$ at different epochs τ_1 and τ_2 is (see Section 4.1):

$$E\{\mathbf{u}(\tau_1)\cdot\mathbf{u}^T(\tau_2)\} = \mathbf{Q'_{uu}}\,\delta(\tau_2 - \tau_1) \quad (4\text{-}12)$$

where $\mathbf{Q'_{uu}}$ is the constant spectral density matrix of $\mathbf{u}(\tau)$ and $\delta(\tau_2 - \tau_1)$ is the Dirac delta function (GELB, 1974). Considering the scalar white noise process u, there is a simple relation between the standard deviation of the noise σ_u and the spectral density q'_u:

$$\sigma_u^2 = q'_{uu}\,\delta(\tau_2 - \tau_1) \quad (4\text{-}13)$$

However, the dimensions need particular attention: for example, a standard deviation of 5m/sec implies a spectral density of $25m^2/sec$. With the white noise assumption for all elements of $\mathbf{u}(\tau)$, and substituting in (4-11), the covariance matrix of the state vector can be obtained with a single integration:

$$\mathbf{Q_{xx}}(t) = \Phi(t,t_0)\,\mathbf{Q_{xx}}(t_0)\,\Phi^T(t,t_0) + \int_{t_0}^{t} \Phi(t,\tau)\,\mathbf{G}\,\mathbf{Q'_{uu}}\,\mathbf{G}^T\,\Phi^T(t,\tau)\,d\tau \qquad (4\text{-}14)$$

In the discrete case where $\mathbf{v}$ is a white random sequence, the expectation of a product of variables becomes:

$$E\{\mathbf{v}(\tau_1)\cdot\mathbf{v}^T(\tau_2)\} = \mathbf{Q_{uu}}\,\delta(\tau_2 - \tau_1) \qquad (4\text{-}15)$$

where $\mathbf{Q_{uu}}$ is the constant covariance matrix of $\mathbf{v}(\tau)$ and $\delta(\tau_2 - \tau_1)$ is the Kronecker delta function. Consider now a linear function of the random sequence $\mathbf{v}$, the expectation of a product becomes:

$$E\{\Gamma(t,\tau_1)\mathbf{v}(\tau_1)\cdot\mathbf{v}^T(\tau_2)\Gamma^T(t,\tau_2)\} = \Gamma(t,\tau_1)\,\mathbf{Q_{uu}}\,\Gamma^T(t,\tau_2)\,\delta(\tau_2 - \tau_1) \qquad (4\text{-}16)$$

Thus, the expression for the covariance matrix of the state vector simplifies to:

$$\mathbf{Q_{xx}}(t) = \Phi(t,t_0)\,\mathbf{Q_{xx}}(t_0)\,\Phi^T(t,t_0) + \Gamma(t,t_0)\,\mathbf{Q_{uu}}\,\Gamma^T(t,t_0) \qquad (4\text{-}17)$$

Equation (4-17) indicates that the predicted covariance has two parts: the propagation of the covariance at a previous epoch and a term due to the propagation of the driving noise. In the continuous case, by analogy and for convenience, expression (4-14) is sometimes also found in the literature. However, an expression for $\Gamma(t,t_0)$ is not readily obtained, and the covariance of the system noise must be computed directly from the integration:

$$\mathbf{Q_{ww}}(t) = \Gamma(t,t_0)\,\mathbf{Q_{uu}}\,\Gamma^T(t,t_0) = \int_{t_0}^{t} \Phi(t,\tau)\,\mathbf{G}\,\mathbf{Q'_{uu}}\,\mathbf{G}^T\,\Phi^T(t,\tau)\,d\tau \qquad (4\text{-}18)$$

In Appendix F, the method presented here is applicable for kinematic models of degree up to 2 and order up to 3.

4.2.4 The GPS Receiver Clock

Very few results based on practical experiences have been published on the kinematic modelling of GPS clocks. Different models are considered here, which include an initial synchronisation error and/or a drift of the clock:

1) **constant bias**: only the initial synchronisation error is considered. This model is adequate if an external frequency standard is used (for example an atomic clock) or if a very short period of time is considered. It is also frequently used when less than 4 satellites are simultaneously visible. The drawback of this model is that even a small clock offset will cause large systematic errors in the position parameters by introducing a bias in all ranges.

2) **white noise**: receiver clock offsets at different epochs are uncorrelated. If an infinite variance is considered for the white noise, the receiver clock offset becomes unpredictable. Indeed this model is no model, its choice means that no attempt is made to relate clock offsets over time. This model can be justified with different arguments. The receiver clock is reset on GPS time periodically through a navigation fix. As modelling residuals are scaled by the speed of light, no kinematic model is better than an approximate one. This is the standard approach when there is enough redundancy to solve for a receiver clock parameter at each epoch. However, as the clock offset need not be explicitly estimated (nuisance parameter), it can be eliminated from the computations. This is generally achieved by differencing the observations between satellites (KING et al, 1987).

3) **random walk**: the deviation of the frequency of the oscillator with respect to its nominal value is assumed to be a white noise random process. A quartz crystal oscillator, as implemented in most receivers, has a stability better than 10^{-10} over a few hours, that is, a deviation smaller than 1μsec is expected after 10^4sec (≈ 3hrs). Hence, the spectral density of the driving noise is 10^{-16}sec^2/sec (=(1μsec)^2/10^4sec, see Appendix F), which corresponds to a standard deviation of 10^{-8} sec/sec for the white noise on the clock rate. Scaled by the speed of light, this represents 3m/sec. Indeed, the movement of the object can generally be predicted much better than that. If the receiver clock offset is not as well

behaved as the position, that is less predictable, the contribution of a kinematic model for the receiver clock to positioning may just as well be neglected. For a reasonably good satellite geometry, simulations have shown that the random walk model with a driving noise of 3m/sec for the receiver clock offset can be replaced by a white noise model with infinite variance: estimated positions remain almost unaffected. To obtain a significant improvement in the position of a GPS station by modelling the receiver clock offset, it is necessary to assume a smaller standard deviation for the driving noise. This in turn may introduce major biases in the position if the kinematic model proves to be over-optimistic, that is, the frequency of the oscillator is not as stable as assumed.

4) **second order random walk**: the frequency drift of the oscillator is assumed to be a white noise random process. In this case, a stability of 10^{-10} over a few hours corresponds to a spectral density of $3 \cdot 10^{-24}$ sec^2/sec^3, which corresponds to a standard deviation of $1.7 \cdot 10^{-12}$ sec/sec^2 $(=(1\mu sec)^2 \cdot 3 / (10^4 sec)^3$, see Appendix F) for the white noise on the frequency drift. Scaled by the speed of light, this represents an acceleration of $5 \cdot 10^{-4}$ m/sec^2. Over a period of 10^4 sec, this model is equivalent to the random walk model presented above, but closer epochs are more strongly correlated.

The comparison of the models is summarised in Table 4.1.

Table 4.1: Comparison of Kinematic Models for the GPS Receiver Clock

Oscillator Status	White Noise on	Equivalent Coordinate	Model	Variance Growth
perfect	none	static	constant	none
unstable	offset	random	white noise	none (var=∞)
freq. offset	frequency	"const. velocity"	random walk	linear
frequency drift	frequency drift	"constant acceleration"	2nd order random walk	3rd order polynomial

All these models are illustrated in Figure 4.2. Two types of models also shown in Figure 4.2 are considered inappropriate for the GPS receiver clock because of the bounded growth of their variance:

- white noise with finite variance
- first order Gauss-Markov (coloured noise)

If an oscillator is not perfect, nothing will maintain or "recall" its frequency to a nominal value during a GPS tracking session. Hence the offset may become very large (when scaled into distance !).

A sound understanding of the behaviour of the system is therefore required of the filter designer. The choice of the models must be adapted to both the data type and the motion. In other words, there is no one general purpose optimal filter.

4.2.5. Reference Coordinate System

Theoretically, the modelling of the motion can be done in any coordinate system, as long as the correlations between the noises affecting the various directions are properly accounted for. However, there is an obvious advantage in choosing a reference system such that stochastic independence can be assumed for the noises affecting the motion along different coordinate axes.

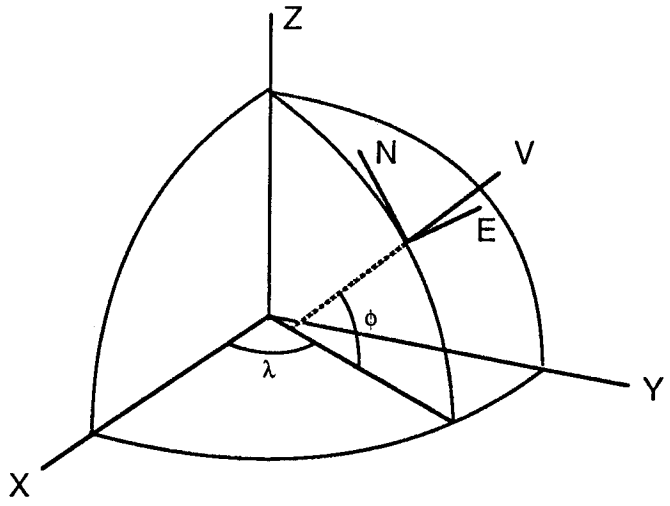


Figure 4.6: Earth Centred and Local Topocentric Coordinate Systems

In the case of a ship, the deviation of the latitude and longitude from the intended route can be significant, but will grow slowly. For the height determination, a different behaviour is expected: sudden changes are possible, but the variations are bounded within a very limited span defined by

the wave motion. Only two horizontal dimensions are desired for many navigation applications. With GPS however, the height must be considered, at least in the measurement model, even if only as an unadjusted or nuisance parameter. In this example, clearly the kinematic model is easier to describe in terms of Easting, Northing and Vertical (ENV) rather than in an Earth Centred, Earth Fixed (ECEF) system such as WGS 84 (XYZ), as illustrated in Figure 4.6. Indeed, it is more convenient to transform the measurement model into the local topocentric system than transform the kinematic model into the ECEF system in which the satellite ephemerides are expressed. Both coordinate systems can be selected in the DYNAMO program, for the input of the models and the output of the coordinates.

4.3. MEASUREMENT MODEL

This component of the estimation procedure is specific to GPS. The principal types of GPS observables used for positioning were introduced in Chapter 1. For the functional models of the pseudo-range and phase-rate, one receiver and one satellite are considered in all the derivations. Operating simultaneously with several receivers may allow for simplifications in the measurement model. Strategies for filtering in the differential mode will be presented in Section 4.4. However, a small discussion of the intricate role of time in GPS measurements is first necessary.

4.3.1. Time and GPS

Unlike most distance measurements used in geodesy, including EDM and lunar or satellite laser ranging, where the signals are transmitted and received by the same instrument with the aid of a reflector at the target, GPS measurements are "one-way" ranges. This means that a possible offset between time systems maintained by the transmitting satellite and the receiver must be accounted for. However, time is present in GPS positioning under a number of other guises and is by far the most difficult concept to grasp. Several time scales are explicitly or implicitly involved in GPS, including:

- The **orbital** time scale -- although this is often overlooked, time exists even without oscillators: a time scale can be defined simply by the orbits of the satellites, according to Newton's laws of motion.

- **GPS Time** (GPST) -- defined by the Control Segment of the GPS system. The orbital elements contained in the Broadcast Ephemeris are functions of GPST. Indeed, GPST itself may not perfectly match the orbital time frame, and this may require modelling when the ultimate accuracy is required (for example, for some crustal deformation surveys). For GPS navigation however, possible differences between GPST and the orbital time scale can be neglected.

- The **satellite clock** -- time-tags are generated initially by the satellite clock, then encrypted in the transmitted message. Each satellite clock, together with the correction parameters contained in the Navigation Message, is an (imperfect) realisation of GPST.

- The **receiver clock** -- When tracking commences, the GPS receiver clock is (imperfectly) synchronised with GPST using the time-tags from a satellite (together with the broadcast correction parameters) and the estimated transit time of the coded message. Then the receiver maintains its own time scale.

- The **time datum** -- chosen as reference for the modelling of the observations. For this purpose, a satellite clock, a receiver clock, an average of different clocks or even an external frequency standard may be used. In other words, a time datum must be defined. This is necessary to ensure the consistency of the models for observations involving a common satellite or a common receiver.

A detailed discussion of the problems caused by possible discrepancies between the various time scales implicit in GPS positioning can be found in GRANT et al (1989). For GPS navigation, two levels of sensitivity to time offsets may be distinguished:

(1) time datum versus GPST -- an offset simply results in the computation of wrong positions for the satellites in their orbits. As the maximum range rate is about 760m/sec (GRANT, 1988), even an error of 1msec will affect the modelled pseudo-ranges by less than a metre. From this point of view, even a relatively unstable receiver clock can provide the time datum for most applications.

(2) satellite or receiver clock versus time datum -- the transit time of the coded message measured in the receiver depends on the times kept by both the satellite and receiver. The measured pseudo-range is only affected by an offset between the two clocks involved. However, in the functional model of the pseudo-range, each clock offset is defined with respect to the time datum. An offset in any of the clocks strongly affects the modelled pseudo-range, as it must be scaled by the speed of light, and even an offset of only 100nsec will affect the pseudo-range by 30m.

Conventionally, in the case of 4 satellites and 1 receiver, the 4 estimated parameters are the 3 station coordinates and the receiver clock offset. If all epochs are considered separately, it is impossible to estimate any additional parameter, thus all satellite clocks are held fixed. In other words, all the satellite clocks are assumed to be perfectly synchronised. In fact, in each modelled pseudo-range, the receiver clock offset is defined with respect to a different time datum. Of course, discrepancies between satellite clocks have an adverse effect on the positioning results.

To overcome the weakness inherent in instantaneous single point positioning, for example to estimate the satellite clocks, more redundancy is required. This can be achieved by modelling the motion (filtering) and/or operating other receivers simultaneously (DGPS). Increasing the redundancy will open some choice for the selection of the datum. This issue will be addressed in more detail in Section 4.4.

In summary, the choice of the time datum for GPS navigation is rather influenced by: "how well can the satellite and receiver clocks be modelled with respect to the time datum" than: "how well does the time datum match GPST".

4.3.2. Modelling the Pseudo-range

The measured pseudo-range r is modelled as:

$$r = \rho + c_R - c^S + \rho_{system} + \rho_{noise} \qquad (4\text{-}19)$$

where: ρ is the range computed from the *a priori* coordinates of the satellite and receiver. It can be expressed as $\sqrt{\mathbf{p}^T\mathbf{p}}$ where $\mathbf{p}$ is the vector joining the satellite and the antenna,

c_R and c^S are respectively the receiver and satellite clock offsets, both scaled into distance by the speed of light,

ρ_{system} is a systematic range error caused by unmodelled effects,

ρ_{noise} is a white, Gaussian, zero mean measurement noise.

GPS positioning is particularly influenced by systematic errors, caused by the unmodelled effects of errors in the orbits, atmospheric delay and selective availability. ρ_{system} is generally larger than ρ_{noise}. If all the physical processes influencing the measurements could be perfectly modelled, ρ_{system} could be computed explicitly, but this is not feasible at present. However, a dramatic simplification arises in the differential mode: for any satellite, the term ρ_{system} is almost identical for all receivers in a small network. Therefore, ρ_{system} is independent of the receiver site and, in effect, becomes part of the satellite clock offset. Hence, ρ_{system} can be removed from the model. To reflect this lumped nature of the satellite clock offset, it will frequently be referred to as the satellite range bias (SRB).

The model (4-19) is linearised using a first order Taylor expansion around predicted values $\tilde{\rho}$, $\tilde{c}_R$ and $\tilde{c}^S$ of the range and clock offsets. The differentiation of the range with respect to the range vector yields:

$$\frac{\partial \rho}{\partial \mathbf{p}} = \frac{\partial \sqrt{\mathbf{p}^T\mathbf{p}}}{\partial \mathbf{p}} = \frac{\mathbf{p}^T}{\sqrt{\mathbf{p}^T\mathbf{p}}} = \frac{\mathbf{p}^T}{\rho} = \mathbf{p}_u{}^T \qquad (4\text{-}20)$$

where $\mathbf{p}_u$ is the unit vector of the direction from the satellite to the antenna. The other derivatives are straightforward and the linearised functional model becomes:

$$r = \tilde{\rho} + \tilde{c}_R - \tilde{c}^S + \begin{bmatrix} \tilde{\rho}_{uE} & \tilde{\rho}_{uN} & \tilde{\rho}_{uV} & 1 & -1 \end{bmatrix} \begin{bmatrix} \delta E \\ \delta N \\ \delta V \\ \delta c_R \\ \delta c^S \end{bmatrix} + \rho_{noise} \qquad (4\text{-}21)$$

where: $\tilde{\mathbf{p}}_u$ is the predicted range unit-vector, expressed here in a local topocentric system, see Section 4.2.5,

δE, δN and δV are increments of the predicted East, North and Vertical coordinates,

δc_R and δc^S are increments of the predicted clock offsets.

For more sites and satellites, the elements of the design matrix must simply be inserted according to the position of the corresponding parameter in the state vector. For certain applications, some parameters of the state vector may not be estimated, and the corresponding partial derivatives can be left out.

4.3.3. Modelling the Phase-rate

The measured phase-rate $\dot{r}$ is modelled as:

$$\dot{r} = \dot{\rho} + \dot{c}_R - \dot{c}^S + \dot{\rho}_{system} + \dot{\rho}_{noise} \qquad (4\text{-}22)$$

where: $\dot{\rho}$ is the range-rate computed from the velocities of the satellite and receiver. It can be expressed as $\dot{\rho} = \mathbf{v}^T \mathbf{p}_u$ where $\mathbf{v}$ is the difference in velocity between the satellite and the receiver and $\mathbf{p}_u$ is the unit vector of the direction from the satellite to the antenna,

$\dot{c}_R$ and $\dot{c}^S$ are respectively the receiver and satellite clock drifts (that is frequency offsets) both scaled into velocities,

$\dot{\rho}_{system}$ is a systematic range-rate error caused by unmodelled effects,

$\dot{\rho}_{noise}$ is a zero mean, normally distributed measurement noise.

Under similar circumstances and for the same reasons already given in Section 4.3.3, $\dot{\rho}_{system}$ can be removed. The functional model (4-22) must be linearised. Therefore, all partial derivatives must be determined. First, the range-rate is modelled using solely the differences in position and velocity:

$$\dot{\rho} = \mathbf{v}^T \frac{\mathbf{p}}{\sqrt{\mathbf{p}^T\mathbf{p}}} \qquad (4\text{-}23)$$

The derivative with respect to position is therefore:

(4-24)

$$\frac{\partial \dot{\rho}}{\partial \mathbf{p}} = \mathbf{v}^T \left(\frac{\sqrt{\mathbf{p}^T\mathbf{p}} \; - \; \mathbf{p} \; \frac{\mathbf{p}^T}{\sqrt{\mathbf{p}^T\mathbf{p}}}}{\mathbf{p}^T\mathbf{p}} \right) = \frac{\mathbf{v}^T \rho \; - \; \frac{\mathbf{v}^T\mathbf{p}}{\rho} \mathbf{p}^T}{\rho^2} = \frac{\mathbf{v}^T \; - \; \dot{\rho} \; \mathbf{p}_u^T}{\rho} = \mathbf{p}'^T$$

which is a linear combination of the position and velocity vectors. The geometric significance of the vector **p'** is discussed in MERMINOD (1988). The derivative with respect to velocity is simply:

$$\frac{\partial \dot{\rho}}{\partial \mathbf{v}} = \frac{\partial (\mathbf{p}_u^T \mathbf{v})}{\partial \mathbf{v}} = \mathbf{p}_u^T \qquad (4\text{-}25)$$

The functional model, linearised around the predicted values $\tilde{\dot{\rho}}$, $\tilde{\dot{c}}_R$ and $\tilde{\dot{c}}^S$ of the range-rate and clock drifts, becomes:

(4-26)

$$\dot{r} = \tilde{\dot{\rho}} + \tilde{\dot{c}}_R - \tilde{\dot{c}}^S + \begin{bmatrix} \tilde{\rho}'_E & \tilde{\rho}'_N & \tilde{\rho}'_V & 0 & 0 & \tilde{\rho}_{uE} & \tilde{\rho}_{uN} & \tilde{\rho}_{uV} & 1 & -1 \end{bmatrix} \begin{bmatrix} \delta E \\ \delta N \\ \delta V \\ \delta c_R \\ \delta c^S \\ \delta \dot{E} \\ \delta \dot{N} \\ \delta \dot{V} \\ \delta \dot{c}_R \\ \delta \dot{c}^S \end{bmatrix} + \dot{\rho}_{noise}$$

where: **p'** is expressed in a local topocentric system, see Section 4.2.5,
$\delta\dot{E}$, $\delta\dot{N}$ and $\delta\dot{V}$ are increments of the predicted East, North and Vertical velocities,
$\delta\dot{c}_R$ and $\delta\dot{c}^S$ are increments of the predicted clock drifts.

The other terms have been defined in Section 4.3.3.

4.4. DIFFERENTIAL GPS

4.4.1. Principle and Application

The rationale of Differential GPS (DGPS) was introduced in Section 1.3, and mathematical implications were discussed in Section 4.3. It is the standard method used to reduce the effect of unmodelled phenomena influencing the range measurements. The **physical** correlation between the measurements recorded at each receiver can be expressed mathematically in two different ways:

1) as a **functional** correlation, by including a common bias in the functional model of the measurements, that is, in the state vector.

2) as a **stochastic** correlation, by including off-diagonal terms in the covariance matrix of the measurements recorded simultaneously at both sites.

If only one epoch is considered, both methods can be made equivalent. For example, holding the reference site fixed and completely correlating all measurements involving the same satellite or estimating one range bias per satellite (with infinite variance) yield identical coordinates for the remote site. However, when data is collected over a certain time, the first approach is preferable, because it matches the time behaviour of the physical correlation much better. An analysis of positioning results obtained using static differential pseudo-range data reveals a clear distinction between the nature of the correlation in the short and in the long term:

- In the **short term** (say over 30sec): the variations in the difference in position between two close sites is larger than the variations in the individual single point positions, and the ratio is approximately $\sqrt{2}$. This phenomenon has already been noted by CLYNCH & HARPER (1987). Hence, the variations are mainly produced within the receiver. Furthermore, the ratio of $\sqrt{2}$ indicates a simple addition of the variances, and measurements at two distinct receivers should therefore rather be considered as being uncorrelated. At first, this seems to contradict the principles of DGPS, however even with stronger random variations, the difference in positions is more accurate than single point positions,

because biases generally larger than the random variations are removed.

- In the **long term**: the mathematical nature of the systematic errors in the pseudo-range measurement appears to be a slow drift. Both (static) single point positions drift regularly by several metres over say 30 minutes. This indicates that the satellite clock offset, the deviations from the modelled orbits and the atmospheric delays do not change quickly, but may attain large values. In spite of these changes in the single point positions, the mean difference in position remains constant.

From this comparison, it is clear that noise and biases can be distinguished. The problem of reducing the effect of systematic errors for each individual epoch is addressed by DGPS, whilst the problem of separating biases from noise using their different time behaviour is addressed by filtering.

As introduced in Section 1.4.1, two techniques have been suggested for the transfer of the differential corrections from a reference site to other sites. The choice can be illustrated by an example with 2 receivers, 4 satellites and 8 pseudo-range measurements at an epoch:

(a) **Correction in the measurement domain**: the coordinates and the receiver clock of the reference site are held fixed, that is, the space datum is provided by the satellites and the reference site, whilst the time datum is provided exclusively by the reference site. The ranges between the reference site and the satellites are known and the discrepancies with the actual measurements are computed and applied to the ranges measured at the remote site. In effect, the remote site is computed as a single point, but using corrected pseudo-ranges.

(b) **Correction in the solution domain**: both coordinates and receiver clock offsets are estimated, unlike the satellite range biases (~ satellite clock offsets). That is, both the space and time datums are provided by the satellites. No estimated parameter is common to both sites and, as long as no stochastic correlation is introduced for the measurements, the positions obtained are uncorrelated, that is, the covariance matrix of the parameters of both sites is block diagonal. Roughly, both positions are systematically biased by the same amount. The difference between the

computed and fixed positions of the reference site is then applied to the computed position of the remote site.

Both methods involve 8 measurements and 8 unknowns and yield identical results. However, the choice of the unadjusted parameters -- the datum -- is not completely free. At this stage, it suffices to say that at least one clock must be held fixed. GPS datum problems are discussed in LINDLOHR & WELLS (1985) and GRANT (1988).

4.4.2. Extension to Filtering

When a simultaneous data processing scheme is considered (see Section 1.4.1), the computation of differential corrections is replaced by the use of common parameters in the functional model of observations involving common sites or satellites. In other words, bias terms are introduced in the state vector. It is still possible to consider corrections in either the measurement or the solution domain. The difference is in the assignment of the biases:

(a) in the measurement domain -- one bias per satellite,
(b) in the solution domain -- one bias per coordinate.

Again, for a single epoch, both approaches yield identical results. However, this may change when more epochs are processed together, even if it is sequentially. The systematic errors are generally associated with the satellites, and this can be illustrated by the case of a setting satellite: with approach (a), the bias of the setting satellite is not used any more and the other satellite biases remain unaffected, whilst with approach (b), several coordinate biases would be affected by this occurrence, quite suddenly. Hence approach (a) is more appropriate, because satellite biases can be modelled in a manner that is closer to the time variation of the systematic errors. For the applications considered in this thesis, a satellite bias such as required in approach (a) corresponds to the term ρ_{system} in (4-19), and cannot be mathematically distinguished from the satellite clock offset, see Section 4.3.2. Therefore, such a satellite (range) bias will absorb the satellite clock offset as well as errors in the orbit and in the atmospheric delay. To define a kinematic model for the satellite range bias, the effect of various error sources must be taken into account.

When filtering, the datum problem remains basically the same as when only one epoch is considered. However, an extension is rendered possible: not only does the choice of the correction domain remain, but filtering can be done in the measurement and solution domains simultaneously. A kinematic model is assigned to the coordinates of the remote site and another kinematic model is assigned to the satellite range biases. The extreme cases are:

(a) **Filtering in the measurement domain**: with a "hard" model (a small system noise) for the satellite biases and a "soft" model (a large system noise) for the coordinates of the remote site, the adjusted range corrections strongly depend on those obtained at previous epochs, and the position of the remote site, with its weakly predicted values, is in effect computed as for a single point, but using filtered range corrections.

(b) **Filtering in the solution domain**: with a "soft" model for the satellite biases and a "hard" model for the coordinates of the remote site, the predicted range corrections have a low weight whilst changes in the position of the remote site between adjacent epochs can be predicted precisely. Consequently, the adjusted range corrections largely depend on the assumed motion of the remote site. In other words, possible discrepancies between the modelled and actual range measurements will be absorbed by the range corrections.

Any other combination of models results in a filter acting simultaneously in both domains. Therefore, the problem is to know whether the range corrections or the coordinates expressing the motion are better behaved, and hence more predictable, and assign the appropriate kinematic models to the various parameters.

In the context of filtering, the choice of the time datum is important. Satellite clocks are generally more stable with respect to each other than with respect to a receiver clock. Also, a receiver clock is generally better behaved with respect to a satellite clock than with respect to another receiver clock. Consequently, the choice of a satellite clock for the time datum permits more precise kinematic models (smaller system noises) to be used for the other clocks, thus improving the predicted clock offsets and increasing the redundancy.

4.5. CONSTRAINTS

4.5.1. Principle and Application

The basis for the application of constraints within a least squares estimation procedure was presented in Section 2.4.4. The approach consists of treating measurements and constraints in two consecutive adjustments, rather than introducing constraints as (pseudo-)measurements and carrying out a single adjustment. Any relationship between the elements of the state vector can be considered, and treated in an optimal manner, as long as the full state covariance matrix is available before constraints are applied. The inclusion of constraints always follows the same pattern, that is, linearising the relation so that it takes the form (2-13):

$$\mathbf{U}\delta\mathbf{x} = \mathbf{t}$$

and (2-14) can then be applied. 3 types of constraints are considered here:

- the distance between two points,
- the difference between the azimuth and the velocity of two points,
- the vertical angle (pitch) between two points.

An important extension here is that the constraints need not be deterministic, that is, the condition need not be completely fulfilled. The approach is illustrated by considering the constraint on the distance, and some of the possibilities that may be encountered:

1) the distance is exactly known and does not change with time. This is the deterministic case most frequently encountered in geodesy.
2) the initial distance and its rate of change are exactly known. This is still a deterministic case, even though the *a priori* distance is different at each epoch.
3) the distance is not known precisely, but is invariant. This is no longer a deterministic constraint, and is indeed a constraint only if measurements at several epochs contribute to the determination of the distance.
4) the distance may change with time, however the change can only be slow and regular. This information can be accounted for by assigning an appropriate kinematic model to the distance.

To ensure maximum flexibility, the constraints will be modelled stochastically. The distance and some of its time derivatives are added to the state vector, and suitably modelled. For the cases above, the solutions are:

1) the distance is introduced in the state vector, with initial variance zero. The kinematic model has degree -1, that is the distance is not estimated (see Section 4.2). Alternatively, the distance is not introduced in the state vector, but stored separately and recalled whenever the constraint must be applied. To reduce the size of the matrices involved, the latter approach is implemented in DYNAMO, as in the case of any other unadjusted and invariant parameter.
2) the initial distance and its rate of change are included in the state vector, with initial variance zero, a kinematic model with degree -1 and no driving noise. The advantage of this procedure is that the *a priori* distance is automatically updated during the prediction step, without any further action.
3) the kinematic model for the distance has degree 0, order 1 and no driving noise. This is the normal procedure for the estimation of a constant.
4) the kinematic model for the distance has degree 0, order 1 and some driving noise. By increasing the degree and the order, the rate of change of the distance may also be estimated, and propagated in the prediction step.

The functional model is developed for each of the constraints in the following sections.

4.5.2. Distance

All computations are performed in a local topocentric system (East, North, Vertical) and the two points are identified by subscripts. The constraint has the form:

$$d = \sqrt{(E_2 - E_1)^2 + (N_2 - N_1)^2 + (V_2 - V_1)^2} = D \qquad (4\text{-}27)$$

In general, the value of the distance contained in the state vector (lowercase) does not match the value computed from the coordinates contained in the

same state vector (uppercase). Before the constraints are applied, the state vector is considered here "predicted" and its elements are noted with the superscript "~".

$$\tilde{d} \neq \sqrt{\Delta\tilde{E}^2 + \Delta\tilde{N}^2 + \Delta\tilde{V}^2} = \tilde{D} \qquad (4\text{-}28)$$

After the constraint is applied, the relation must be satisfied and the problem is therefore to compute increments for the predicted values such that:

$$\tilde{d} + \delta d = \tilde{D} + \delta D$$

or:

$$\delta D - \delta d = \tilde{d} - \tilde{D} \qquad (4\text{-}29)$$

δD is replaced by its coordinate components:

$$\delta D = \frac{\partial D}{\partial E_1}\delta E_1 + \frac{\partial D}{\partial N_1}\delta N_1 + \frac{\partial D}{\partial V_1}\delta V_1 + \frac{\partial D}{\partial E_2}\delta E_2 + \frac{\partial D}{\partial N_2}\delta N_2 + \frac{\partial D}{\partial V_2}\delta V_2 \qquad (4\text{-}30)$$

Using classical calculus, the partial derivatives are established (they have the same form as those explicitly derived for the pseudo-ranges, see Section 4.3.2), and the linearised constraint becomes:

$$\begin{bmatrix} -\frac{\Delta\tilde{E}}{\tilde{D}} & -\frac{\Delta\tilde{N}}{\tilde{D}} & -\frac{\Delta\tilde{V}}{\tilde{D}} & \frac{\Delta\tilde{E}}{\tilde{D}} & \frac{\Delta\tilde{N}}{\tilde{D}} & \frac{\Delta\tilde{V}}{\tilde{D}} & -1 \end{bmatrix} \begin{bmatrix} \delta E_1 \\ \delta N_1 \\ \delta V_1 \\ \delta E_2 \\ \delta N_2 \\ \delta V_2 \\ \delta d \end{bmatrix} = \tilde{d} - \tilde{D} \qquad (4\text{-}31)$$

This is a linear model of the form (2-13). Note that this constraint cannot be applied if the predicted distance between the points tends to zero, because some partial derivatives tend to infinity.

4.5.3. Azimuth and Heading

The usefulness of this type of constraint is not as obvious as for the distance. Figure 4.7 illustrates the situation of a ship. Although the azimuth need not be equal to the heading, the difference between them may be expected to change slowly.

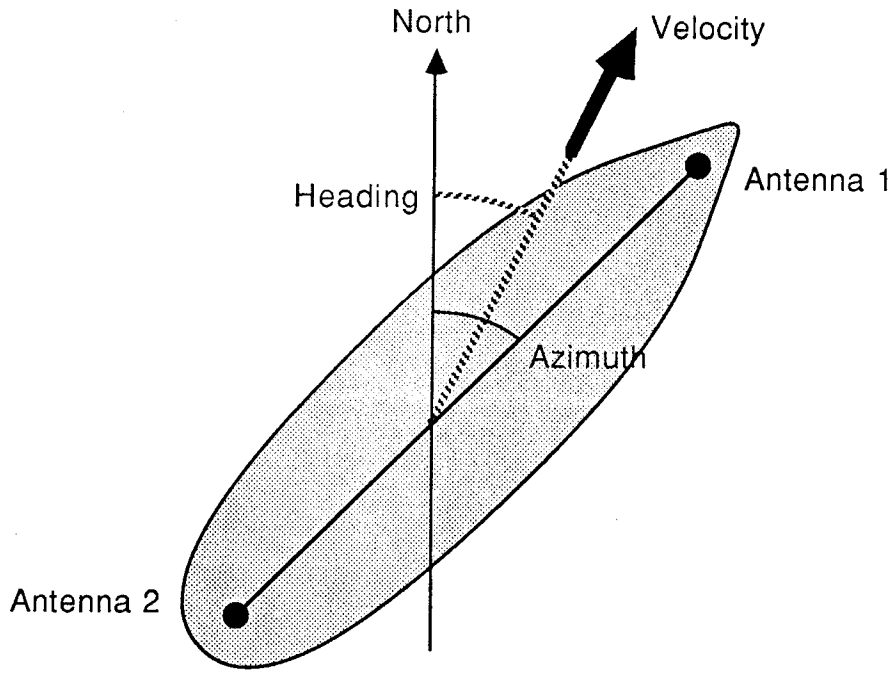


Figure 4.7: Azimuth and Heading of a Ship

The linearisation follows exactly the same pattern as for the distance in Section 4.5.2. The relation between the azimuth and the heading has the form:

$$\omega = \text{Azimuth} - \text{Heading} = \arctan\frac{E_2 - E_1}{N_2 - N_1} - \arctan\frac{\dot{E}_1 + \dot{E}_2}{\dot{N}_1 + \dot{N}_2} = \Omega \qquad (4\text{-}32)$$

The above relation is generally not satisfied by the actual elements of the state vector:

$$\tilde{\omega} \neq \arctan\frac{\Delta\tilde{E}}{\Delta\tilde{N}} - \arctan\frac{\Sigma\dot{\tilde{E}}}{\Sigma\dot{\tilde{N}}} = \tilde{\Omega} \qquad (4\text{-}33)$$

Setting: $$\delta\Omega - \delta\omega = \tilde{\omega} - \tilde{\Omega} \tag{4-34}$$

and replacing $\delta\Omega$ by its coordinate components:

(4-35)

$$\begin{bmatrix} -\frac{\Delta\tilde{N}}{\tilde{D}_H^2} & \frac{\Delta\tilde{E}}{\tilde{D}_H^2} & \frac{\Delta\tilde{N}}{\tilde{D}_H^2} & -\frac{\Delta\tilde{E}}{\tilde{D}_H^2} & -\frac{\Sigma\tilde{\dot{N}}}{\tilde{\dot{S}}_H^2} & \frac{\Sigma\tilde{\dot{E}}}{\tilde{\dot{S}}_H^2} & -\frac{\Sigma\tilde{\dot{N}}}{\tilde{\dot{S}}_H^2} & \frac{\Sigma\tilde{\dot{E}}}{\tilde{\dot{S}}_H^2} & -1 \end{bmatrix} \begin{bmatrix} \delta E_1 \\ \delta N_1 \\ \delta E_2 \\ \delta N_2 \\ \delta \dot{E}_1 \\ \delta \dot{N}_1 \\ \delta \dot{E}_2 \\ \delta \dot{N}_2 \\ \delta\omega \end{bmatrix} = \tilde{\omega} - \tilde{\Omega}$$

where: D_H is the horizontal distance between the two points, that is:

$$D_H^2 = \Delta E^2 + \Delta N^2$$

$\dot{S}_H^2$ is the sum of the squared horizontal velocities of both points:

$$\dot{S}_H^2 = \Sigma\dot{E}^2 + \Sigma\dot{N}^2$$

a linear model of the form (2-13) is obtained. However, some pitfalls deserve particular attention:

- The partial derivatives of the position parameters tend to infinity if the predicted distance between the points tends to zero, as the azimuth is not defined in this case.

- The partial derivatives of the velocity parameters tend to infinity if the predicted sum of the velocities of the points tends to zero, that is $\Sigma\dot{E}$ and $\Sigma\dot{N}$ tend to zero, as the heading is not defined. Such a situation arises when the points are rotating around their centrepoint. Indeed, it is wise not to apply the constraint when the velocity is too small or not significant, for example when the magnitude of the velocity is smaller than three times its mean error.

- All angles must be expressed in consistent units. In many computers, the use of the function "arctan" yields Ω in radians. If the parameter ω contained in the state vector is in degrees, which is more convenient, some partial derivatives must be scaled by a conversion factor.

- The function "arctan" used for the computation of $\tilde{\Omega}$ is not continuous. However, discontinuities can be avoided by adding or subtracting multiple of 360 degrees so as to ensure that $\tilde{\omega} - \tilde{\Omega}$ is always comprised between -180 and +180 degrees.

4.5.4. Vertical Angle

The linearisation follows the same pattern as for the distance in Section 4.5.2. The constraint on the vertical angle has the form:

$$\theta = \arctan \frac{(V_2 - V_1)}{\sqrt{(E_2 - E_1)^2 + (N_2 - N_1)^2}} = \Theta \qquad (4\text{-}36)$$

which is generally not satisfied by the actual elements of the state vector:

$$\tilde{\theta} \neq \arctan \frac{\Delta\tilde{V}}{\sqrt{\Delta\tilde{E}^2 + \Delta\tilde{N}^2}} = \tilde{\Theta} \qquad (4\text{-}37)$$

Setting:
$$\delta\Theta - \delta\theta = \tilde{\theta} - \tilde{\Theta} \qquad (4\text{-}38)$$

and replacing $\delta\Theta$ by its coordinate components:

(4-39)

$$\begin{bmatrix} \frac{\Delta\tilde{E}\Delta\tilde{V}}{\tilde{D}_H\tilde{D}^2} & \frac{\Delta\tilde{N}\Delta\tilde{V}}{\tilde{D}_H\tilde{D}^2} & -\frac{\tilde{D}_H}{\tilde{D}^2} & -\frac{\Delta\tilde{E}\Delta\tilde{V}}{\tilde{D}_H\tilde{D}^2} & -\frac{\Delta\tilde{N}\Delta\tilde{V}}{\tilde{D}_H\tilde{D}^2} & \frac{\tilde{D}_H}{\tilde{D}^2} & -1 \end{bmatrix} \begin{bmatrix} \delta E_1 \\ \delta N_1 \\ \delta V_1 \\ \delta E_2 \\ \delta N_2 \\ \delta V_2 \\ \delta\theta \end{bmatrix} = \tilde{\theta} - \tilde{\Theta}$$

where D_H and D are respectively the horizontal and slope distances between the two points, that is:

$$D_H^2 = \Delta E^2 + \Delta N^2$$

and $$D^2 = \Delta E^2 + \Delta N^2 + \Delta V^2$$

a linear model of the form (2-13) is obtained. Note that this constraint cannot be applied if the predicted horizontal distance between the points tends to zero, as some partial derivatives tend to infinity. In addition, the remark regarding the consistency of the angle units in Section 4.5.3 also applies to θ, Θ and the corresponding partial derivatives.

5. THE SOFTWARE PACKAGE "DYNAMO"

5.1. STRUCTURE AND FEATURES

5.1.1. The Software Components

Very few software packages are available for GPS navigation. Some receiver manufacturers provide a package that can be readily used for real-time differential positioning. However, these packages are designed to be used with particular equipment and can only treat standard cases of positioning, such as single point positioning or differential positioning with only one static and one mobile receiver. The choice of kinematic models for the various parameters, if any, is very limited. Therefore, an important effort has been dedicated to the development of the software required to support the present research.

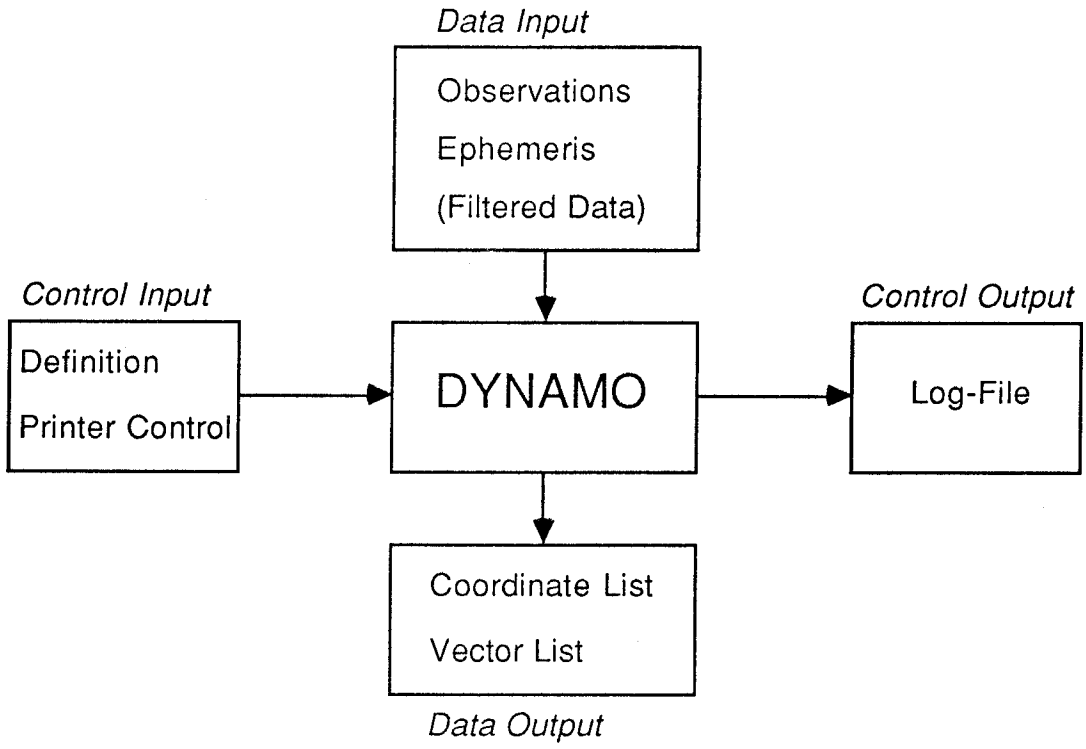


Figure 5.1: Input and Output of DYNAMO

The main characteristics required for a software package aimed at supporting GPS kinematic research is its **flexibility**: it must be possible to consider new

applications by simply replacing subroutines or including new ones, without having to change the structure of the program. From the start, it was clear that an optimal smoothing algorithm and the capability of including constraints on the parameters had to be incorporated.

The software is written in FORTRAN 77 and, at present, is installed on a mainframe VAX computer. An overview of the data flow through the estimation procedure is shown in Figure 5.1. As a prerequisite to the use of DYNAMO, the GPS observations and ephemeris must be stored into the predefined "Archive" formats. This is realised with the preprocessors available at the School of Surveying, University of New South Wales, for different GPS receiver brands (TRIMBLE, TI4100, WM101). After the processing, the resulting coordinates and trajectories can be directed to general purpose plotting routines.

Once the input has been correctly formatted, the main menu can be called by running the command file KIN.COM. The components of the package are listed on the screen and all subsequent operations are performed in the interactive mode.

PROCESSING GPS KINEMATIC DATA

1. Create the measurement file
2. Modify the measurement file
3. Create the definition
4. Modify the definition
5. Create the printer control file
6. Modify the printer control file
7. Estimation Procedure
8. Select coordinate output
9. Select vector output
10. Compute coordinate differences
11. Compute coordinate statistics
12. Exit

Enter the code [1-12]: ...

Figure 5.2: The Main Menu of DYNAMO

Code 7 activates the principal program, described in Section 5.2. Codes 1 to 6 and 8 to 11 activate auxiliary programs to facilitate the preparation of the input and the further manipulation of the output. These are described in

Section 5.3. For all programs, the main design criteria were user-friendliness and easy maintenance.

5.1.2. User Interface

User-friendliness is achieved by a sophisticated organisation of the data within the program, so that the user need not be concerned about such things as the storage of the data. For example, the location of a particular parameter within the state vector is irrelevant to the user, who only has to indicate which parameters must be considered, in any sequence. Sites and satellites are only known to the user by an identification that is independent of the actual storage location. This principle can be illustrated by the search of information related to the satellite called SAT, in an array of identifiers SATS with dimension MSAT (the number of satellites available) and counter NSAT.

```
100   READ (keyboard,*) SAT
      DO NSAT=1, MSAT
           IF (SAT .EQ. SATS(NSAT)) GOTO 200
      ENDDO
      WRITE (screen,10) SAT
 10   FORMAT (' Satellite ',I2,' not found !')
      GOTO 100

200   ...
```

Using NSAT, the storage location is then derived (for example, within an array), and the relevant information can be stored or extracted by the code. For the state vector, the procedure is similar except that the identification of the parameters is an alphanumeric string. Some examples demonstrate the systematic classification used:

"P E 1"	⇔	Position	Easting	Site 1
"V DIS 3-2"	⇔	Velocity	Distance	Site 3 minus Site 2
"A SRB 13"	⇔	Acceleration	Sat. Ran. Bias	Satellite 13

The second example refers to the rate of change of a distance, whilst the third example refers to the frequency drift of a satellite clock and second time derivatives of unmodelled range errors. Each string corresponds to a

parameter number, which is used internally within the program, but never shown. Conversions are executed in small subroutines CODEX and BACODEX. Obviously, the price of such convenience is an increase in the size of the program: a friendly user interface requires more effort than the coding of the mathematics of the problem at hand.

5.1.3. Maintenance

Software maintenance is closely related to the flexibility. To ensure consistency between the different programs in the DYNAMO package, the array storage parameters are contained in a separate file, and then included in different programs (code 3, 4 and 7). Thus, if only maximum array dimensions are to change, only the "include" file SIZE.INC requires editing, and the programs must simply be recompiled. The same principle is applied to the allocation of the units, defined in the "include" file UNIT.INC. As an important feature, all subroutines are valid for any maximum dimensions selected in the programs. This is achieved by passing the required arrays through the argument list of the "call" statement. Consequently, subroutines need not even be recompiled if the storage parameters included in the main program are modified.

All the theoretical background for the software development has been presented in Chapters 2 to 4. However, many comments have been inserted in the code, in particular in the subroutines dealing with the transition matrix, the system noise, the design matrix and the constraints.

5.2. MAIN PROGRAM

5.2.1. Characteristics

The estimation procedure is the heart of the DYNAMO package, but the user can exercise very little control at this stage. The only option is whether filtering or smoothing should occur. In the second case, the file containing the filtered data must be input.

All least squares subroutines are written for general purpose: they are not specific to any sensor and are valid for any state vector configuration. The same principle applies to the modelling of the movement of the receiver(s).

Both the transition matrix and the covariance matrix of the system noise can treat any combination of stochastic variables.

Operations specific to GPS are required to build the design matrix of the measurements, and consist mainly of orbit interpolation.

All subroutines can be used for computations in either an Earth Centred, Earth Fixed (ECEF) coordinate system such as WGS 84 or a local topocentric system (E,N,V see section 4.2.5). The choice is indicated by a flag in the argument list. All the subroutines used in the estimation procedure are grouped in the library FOREST (FOR ESTimation).

The time required to run through a cycle of the filter, for example with 8 measurements and 16 state parameters, is less than 1 second. The speed of the computation has not been given high priority. This may seem strange, as navigation problems often must be solved using an onboard computer with limited capacity. However, a refinement of the computational efficiency would render subsequent modifications more tedious. Therefore, full matrices are stored, even if they are symmetric or sparse, and double precision arithmetic is used throughout. There exist a number of methods to improve the computational efficiency of the Kalman filter, such as the square root formulation (BIERMAN, 1977). Although this issue was essential in the early computer age, the dramatic increase in the capacity of even small machines makes it difficult to justify spending much effort on algorithm efficiency. Nevertheless, the problem of storage cannot be completely neglected: with well over one hundred different types of state parameters available, dimensioning the matrices for the maximum size would be senseless. Dynamic memory allocation has been resorted to: storage is reserved and arrays are dimensioned at run-time, once the size of the state vector has been defined. In a multi-user environment, by reducing the storage requirements to the amount actually used, the execution of the program is accelerated. Should the throughput be further improved for a particular application, some matrix manipulations could be optimised without changing the structure of the package.

5.2.2. The Filter

The task of DYNAMO is to allocate the input to the appropriate units, read the definition file, name the output and reserve memory for the estimation process (according to the actual size of the state vector). The process is then started by FILTER, which is primarily a succession of calls to the subroutines contained in FOREST. This organisation is apparent in Figure 5.3.

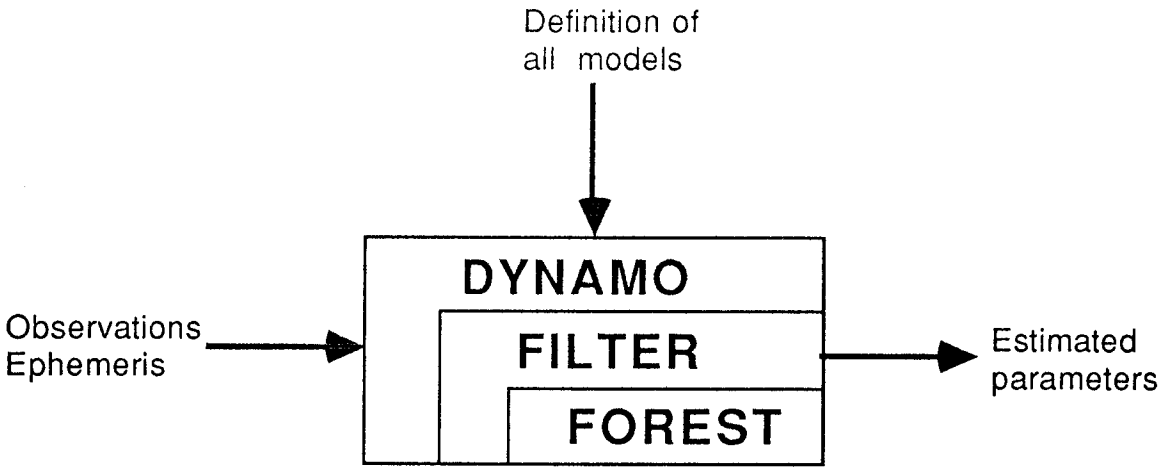


Figure 5.3: Data-Flow of the Filter

The Kalman and the Bayes formulations of the filtering update are implemented. Both can be used in all cases and yield identical results. The Kalman formulation is the normal choice, however the Bayes formulation is faster if the number of observations processed simultaneously is larger than the number of parameters. To set the importance of the update step of the procedure in perspective, both formulations consist of about 30 lines of FORTRAN code, whilst the entire DYNAMO package requires over 2000 lines of code, without comments and general purpose subroutines such as orbit interpolation or matrix inversion.

The stability of the algorithms has been tested with various datasets. Two problems did arise:

- In some cases, discrepancies were found in the filtered state covariance matrix, using the simplified Kalman update formula (3-14). With the complete update formula (3-13) however, the problem disappeared

completely and the results have always been identical to those obtained with the Bayes formulation (3-8).

- With large gaps between adjacent epochs, the symmetry of the predicted state covariance matrix (3-3) was not always warranted. The degradation however was always progressive and started very slowly. This problem was removed by forcing the symmetry at each step, that is, by taking the mean of symmetric elements.

5.2.3. The Smoother

The smoothing estimation procedure is organised like the filtering one, as the comparison of Figures 5.3 and 5.4 shows. The filtered data consists of the state vector and its full covariance matrix at all epochs. In order to limit the size of the file, only the variances on the diagonal are stored with double precision, and all correlations are stored as integers r such that: correlation = r / 1000.

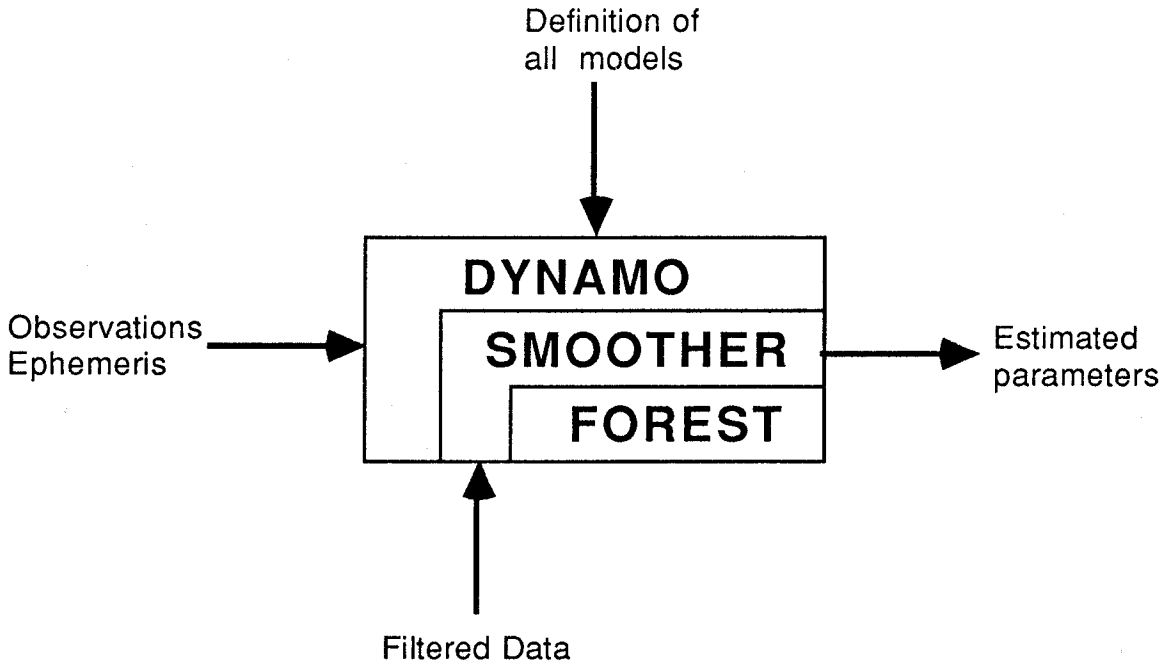


Figure 5.4: Data Flow of the Smoother

The formulation using smoothed residuals has been implemented, see section 3.6.2. The validity of both the concept and the implementation have been tested with a number of simulated and real data-sets. A formal test is possible with coordinates modelled as constants: after the last epoch has been

processed by the filter, the solution is the same as that obtained by the simultaneous processing of all observations, and the smoothed solution for all epochs must be identical to the last filtered solution, because smoothing is equivalent to processing all the data (past, present and subsequent) at each epoch. Figure 5.5 is a sketch of results obtained using real GPS pseudo-range and phase rate data. The trend may have different causes, for example an unmodelled satellite clock drift, but this is irrelevant in the following discussion. As the number of processed observations increases, it is more difficult for the filtered solution to adapt to the trend, because the positioning information is accumulated and the predicted coordinates become stronger. Despite the noise and the biases in the data, all the smoothed coordinates were contained within a range of 2 mm, over the entire half-hour session.

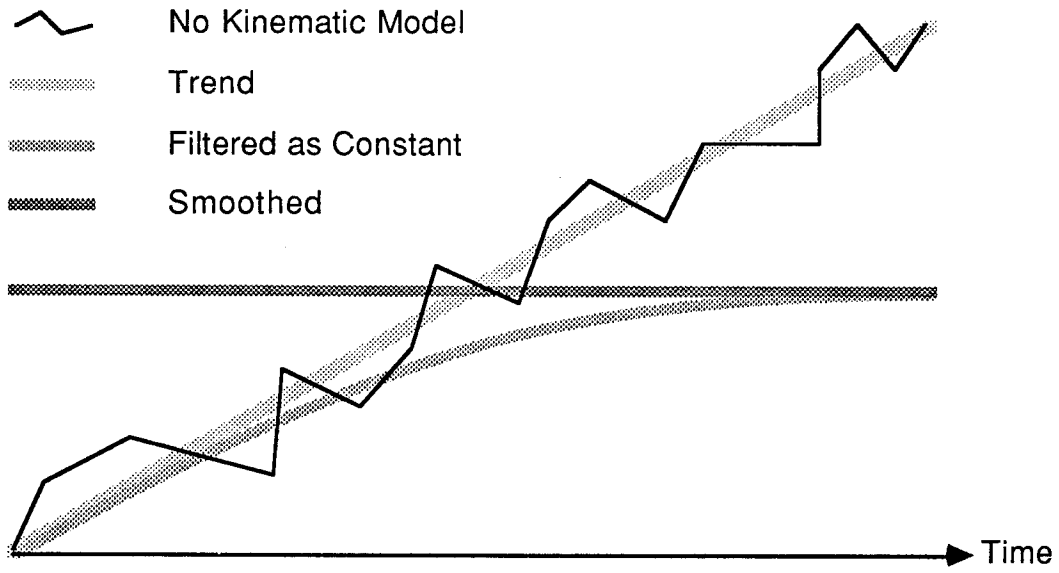


Figure 5.5: Formal Test of the Smoother

5.2.4. Comparison of Filter and Smoother

Smoothing is often said to proceed in the same way as filtering, only backwards in time, and this may be a statement good enough for a presentation of the principles. However, when it comes to an actual implementation of the algorithms, the differences between the two procedures cannot be glossed over. Although the filtering and smoothing algorithms may share some subroutines, each procedure also requires some specific components. This is illustrated by a comparison of flow charts in Figure 5.6. For the sake of simplicity, the restriction is imposed on both sides that the state

vector is updated once per observation epoch. When a new epoch of measurement is encountered, the previous epoch is processed and a new step is initialised. The structure remains the same for applications involving other sensors than GPS. **x** is the state vector and **z** is the adjoint state vector, see Section 3.5.8.

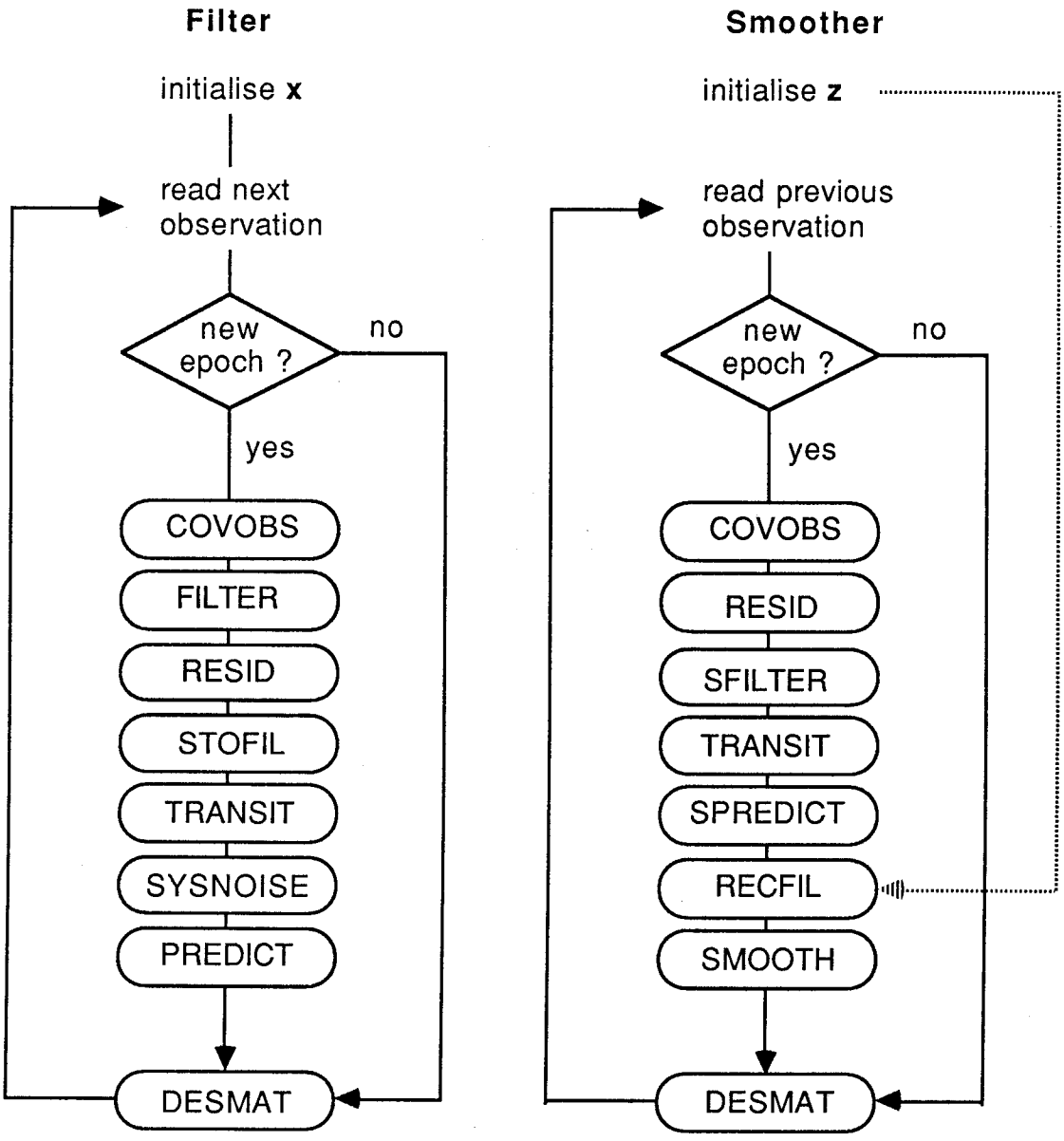


Figure 5.6: Compared Data Flow of Filter and Smoother

The role of the different subroutines is outlined below.

a) Common routines

COVOBS	Build the covariance (and the weight) matrix of the observations.
RESID	Compute the adjusted observation residuals and their covariance matrix, then test the results.
TRANSIT	Build the transition matrix.
DESMAT	Add one line to the design matrix and compute the approximate observation residual.

b) Routines specific to the filter

UPDATE	Update the state vector, using Bayes or Kalman formulation.
SYSNOISE	Build the covariance matrix of the system noise.
PREDICT	Propagate the state vector, its covariance and add system noise.
STOFIL	Store the filtered data.

c) Routines specific to the smoother

RECFIL	Recall the filtered data.
SUPDATE	Update the adjoint state vector, using smoothed residuals.
SPREDICT	Propagate the adjoint state vector and its covariance.
SSTATE	Compute the smoothed increments of the filtered state vector.

When smoothing using the smoothed residuals, the computation of adjusted observation residuals in RESID is superfluous, because they are equal to the approximate residuals already computed in DESMAT.

5.3. AUXILIARY PROGRAMS

5.3.1. Prepare the Measurement File

The estimation program requires all measurements to be contained in the same file. In addition, the measurement file is intended for direct access, rather than sequential, mainly to facilitate the backwards reading by the smoother. The task consists therefore of merging and time-sorting the data collected by several receivers, as shown in Figure 5.7. If desired, the satellite clock corrections contained in the broadcast message are applied. For a real-

time capability, the program SORTOBS presently used could be replaced by an appropriate data-logger. No other task would require any modification.

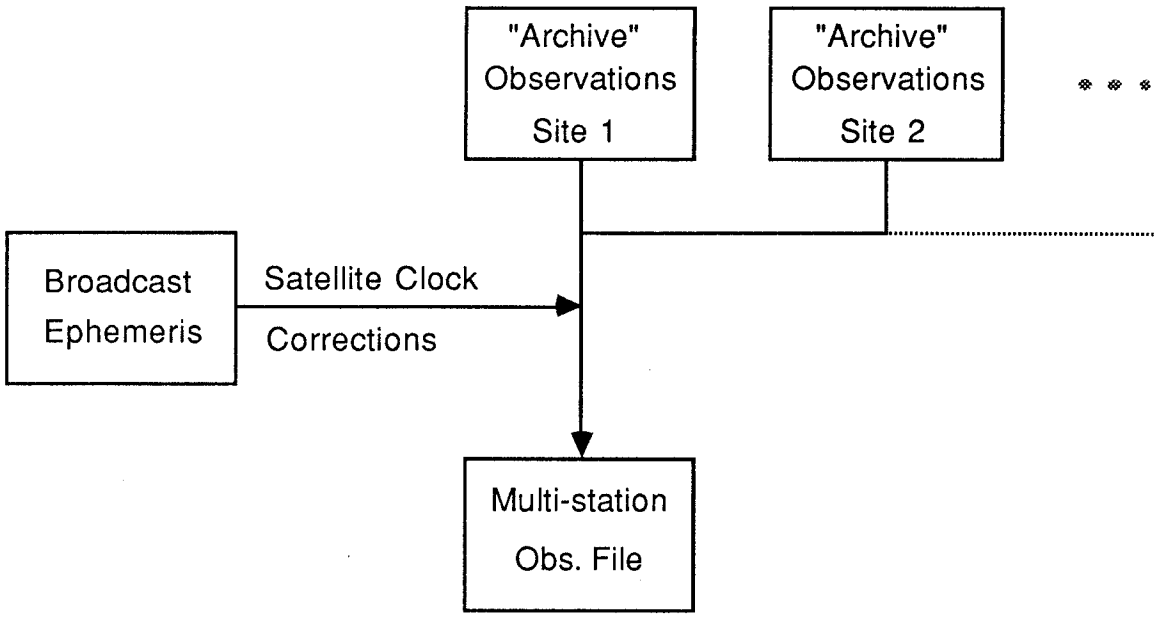


Figure 5.7: Preparation of the Measurement File

5.3.2. Define the Estimation Procedure

The number of available options for the estimation is quite large, and it would be very inconvenient to have to redefine them each time a run is made. Therefore, most options are contained in a definition file, created interactively in the module CREDEF. This operation is by far the most fastidious part of the processing, since all stochastic models must be expressed. Six parts can be distinguished in the definition:

1) Measurements and Ephemeris -- filenames and addresses.

2) Data Selection:
   - start and stop times,
   - rejected sites and satellites,
   - cut-off elevation angle.

3) Initial Elements -- value and standard deviation of all parameters in the state vector:
   - receiver coordinates and clocks,
   - constraints,

- satellite range biases,
- any correlation between initial parameters.

4) Kinematic Model -- for each dimension considered:
   - degree,
   - order,
   - standard deviation of the driving noise,
   - correlation time.

5) Measurement Model: - standard deviation of each type of measurement at each individual site,
   - correlations between sites, satellites or types of measurements.

6) Update Step: - Type of step (each epoch or each satellite at each epoch),
   - maximum number of measurements per step,
   - Bayes or Kalman filter formulation.

Once completed, the definition can be edited using the program TESDEF. This enables very easy changes in any of the definition files created. The consistency of the options is tested before returning to the main menu. The same definition file is used for smoothing, even though some components are irrelevant, for example the initial values of the parameters, as they are extracted from the last epoch in the filtered data.

5.3.3. Control the Printer

Generally, the site coordinates are the only parameters of interest, and a default output is accordingly defined. However, there are many options. As shown in Figure 5.1, the output of the estimation procedure consists of three files:

- The **log-file** lists the selected options and reports about any particular event during the run, such as rejection of observations and large adjustment residuals. For a normal run, only the number of measurements processed at each epoch is reported, together with the ratio *a posteriori / a priori* for the mean error of the residuals. In addition,

most intermediate results can be listed, such as satellite positions, full state vector, design matrix and covariance matrix of the system noise. These options are mainly useful for program development.

- The **coordinate file** is strictly formatted: each group of two lines contains the time, the four components associated with each site (3-D position and receiver clock parameter) and the corresponding mean errors. For each site, positions and/or velocities and/or accelerations can be output, at each selected stage of the computations (predicted, filtered and constrained). Unadjusted parameters are also listed, for example a fixed height. In addition, baselines from a selected reference site can be output. A code at the end of the line indicates to which stage, derivative and site (or baseline) the information is related. For example, "F+C POS 2-1" stands for the filtered and constrained position of the baseline Site 2 minus Site 1, that is, the difference in coordinates from Site 1 to Site 2.

- The **vector file** contains 2-D positioning information, in the form of magnitude and azimuth, with their mean errors. When a local topocentric system is used for the computations, this format is useful to express a horizontal baseline (for example "FIL POS 3-1") or the velocity of a single site (for example "SMO VEL 2 "). Like the coordinate file, this output list may contain information pertaining to different stages, derivatives and sites.

5.3.4. Select Output

The raw output files, whether in coordinate or vector form, contains information that is too heterogeneous to be analysed or directed to a plotter. Therefore, the desired information is extracted from the main output and directed to a new file, by selecting the appropriate stage, derivative and site (or baseline).

5.3.5. Test the Coordinates

This testing consists of two small programs, generally applied after the desired information has been extracted from the raw output. DELTA computes differences between two coordinate files and SIGMA computes a linear regression on all coordinates, or coordinate differences, between selected epochs. This is useful to detect variations and trends in the results.

6. TESTING FILTER PERFORMANCE

6.1. THE ROLE OF SIMULATIONS

There are a number of criteria that may be considered for testing a filter. Three of them are discussed here:

- Optimality in the least squares sense: the evaluation is based on the ratio mean error *a posteriori* / mean error *a priori* of the measurement residuals. This ratio can be output after each adjustment step.

- Minimum number of parameters: this results in computational savings.

- Sensitivity to systematic errors: the neglection of some nuisance parameters (for example, atmospheric delay) leads to systematic biases in the parameters of interest. This also happens if the motion of the object is improperly modelled (for example, drift neglected).

All criteria cannot be satisfied simultaneously, this is particularly true for the second and the third criteria. Therefore, several models should be tested and a compromise found. The standard approach is to look first at the least squares criteria. However, such an optimal filter is often not viable, because for most real-time applications, a small onboard computer is the rule rather than the exception. Nevertheless, it does provide a reference against which other filter designs can be tested. Simplifications of the models allow for a reduction in the time required for a computation cycle. Finally, the effect of systematic errors must be estimated, in order to ascertain whether the proposed simplification is reasonable.

Testing is full of trade-offs and all filters used in practice are sub-optimal. The complexity of the evaluation of filters renders simulation studies necessary. Their principal utility in least squares analysis is in the testing of different observation and computation schemes before actual measurements are collected.

Geodetic networks have been adjusted using least squares estimation techniques for over a century, but simulations have only become common practice since the dramatic increase in the power of computers. However,

experience with triangulation networks has led to an almost intuitive feel for the strengths and weaknesses of such networks and it is now relatively easy to come up with a sound design without explicitly simulating the adjustment of such a network. When electronic distance measurements became available, simulations were used on a wider scale to select the best lines to measure in order to upgrade a triangulation network.

With GPS of course, the situation is much more complex. Because of the nature of the observables and the changing satellite geometry, it is very difficult to evaluate an estimation procedure when guided only by intuition. In addition, the error budget of GPS is often dominated by systematic errors, caused for example by imperfectly modelled clocks, orbit errors and atmospheric delays. Such effects can be studied in different ways:

- **Deterministic simulations** give the change in the parameters of interest induced by a change in the measurements or unadjusted parameters.

- **Stochastic simulations** analyse the propagation of mean errors and correlations, that is, the precision of the results, and their sensitivity to systematic errors can be investigated. Such programs are based on the techniques of covariance analysis and require no observations.

The complexity of GPS is not the only reason for performing simulation studies, the novelty of the system means that users cannot rely on a lot of experience. Rigorous simulations may assist in the understanding of complicated phenomena, but on the other hand, simple simulations can also prove very useful: the well known DOP factors are nothing more than the outcome of a simulation of a resection of instantaneous pseudo-ranges. Indeed, simulations have become an essential tool in the design of complex filters, as they are now in GPS survey design.

6.2. GPS SIMULATION CAPABILITY

Results obtained using real data constitute a strong basis for developing more elaborate simulations based on realistic assumptions. As the intention is to investigate particular applications of kinematic GPS, the need for an efficient deterministic simulation capability quickly appeared when designing the filter.

The programs TRAJ and SIMOBS have been written for this purpose. Simulated observations are produced as follows:

1) a 3-dimensional geometric trajectory is defined, with lines, circles and transition curves.
2) a time trajectory is created by defining the chainage and its time derivatives (up to degree 4) for each time section. Sinusoidal movements can be added.
3) from the two files described above, a time-tagged trajectory is computed, that is, a file of the positions (and velocities) at selected epochs is generated.
4) a tabulated ephemeris file is generated using the Broadcast Message transmitted by the satellites, or obtained from any other source (for example, to simulate future satellite constellations).
5) time-tagged observations are then generated from the time-tagged positions of the receiver and the interpolated positions of the satellites.

The tasks 1) to 3) are executed in the system of programs TRAJ and the two remaining tasks in SIMOBS. Many subroutines, in particular those dealing with orbits are already used in DYNAMO. This is apparent in Figure 6.1 when comparing the production of simulated observations and the processing of real observations.

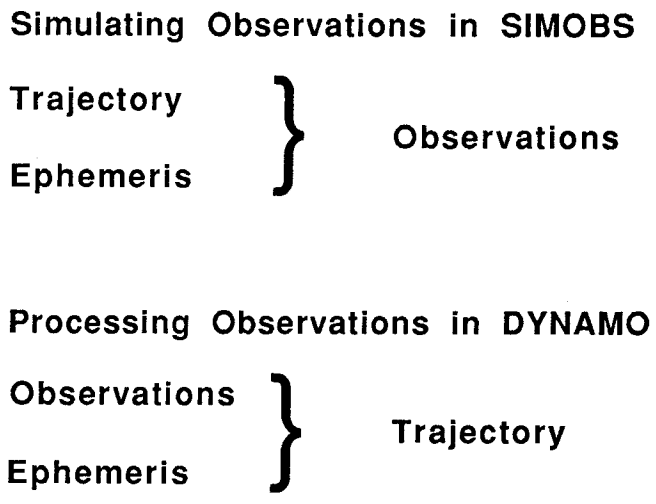


Figure 6.1: SIMOBS versus DYNAMO

The simulated measurements are stored in the "Archive" format, and their treatment by the estimation program DYNAMO is identical to that of real data, as shown in Figure 6.2.

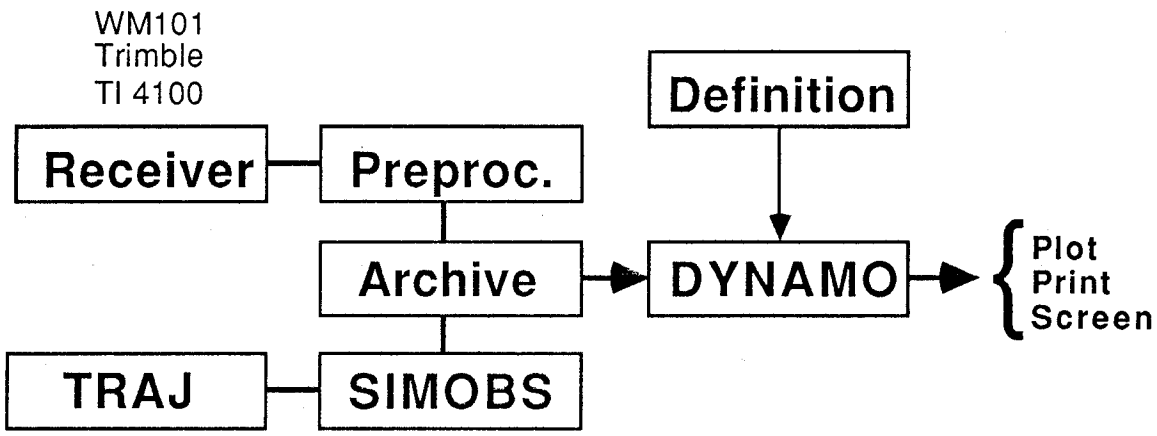


Figure 6.2: Real and Simulated Data Processing

In many navigation problems, the true track is not known precisely. By processing simulated data, a reference route is provided, against which the filtered results can be compared. Noise could be added to the simulated measurements, but this possibility has not been implemented yet. Thus, the comparisons mainly show how a kinematic model may affect the resulting trajectory. In the absence of measurement noise, the original trajectory must be fully recovered when instantaneous position fixes are computed. All kinematic models introduce some inertia in the system and the filter reacts to changes in the trajectory with a delay. The delay is a source of systematic error, but this is the price to pay for a more regular, better behaved, trajectory estimate. Indeed, the problem of the filter designer can be expressed as: **" what systematic errors can be afforded to justify a reduction in the random noise ? "** Simulations permit one to gain a feeling for the behaviour of the filter, by showing explicitly the delays in the system.

Another GPS simulation software package, called DASH, has been completed recently at UNSW (GRANT, 1988). Its purpose is to investigate the propagation of systematic errors, and its strength is the capacity to simulate the adjustment of multi-station GPS networks. Systematic errors are regarded as the consequence of errors in unadjusted parameters, and their effect on the adjusted parameters is computed. Not only deterministic, but also stochastically modelled error sources can be considered, such as changing

characteristics of the ionosphere. Although this program was designed for static studies, it has already proven useful in the preliminary design of models for kinematic applications. DASH is used mainly to test which nuisance parameters require modelling for a desired precision of the results, thus assisting in the design of sub-optimal filters.

6.3. THE SYDNEY-MANLY FERRY

6.3.1. Assumptions

The Sydney-Manly ferry provides an interesting test case for kinematic models. Its route is illustrated in Figure 6.3. Starting from the Manly pier, the ferry accelerates to 9m/sec in 3min, then travels with a constant speed until the middle of the second straight line. Then it decelerates to 3m/sec before the left curve. A final deceleration, markedly stronger for the last 30sec, precedes the arrival in Sydney Cove. It takes a half hour to travel this 11.3km route. The speed profile is given in Figure 6.4.

The future constellation of 18 satellites is considered and the time origin is chosen arbitrarily. The elevation cut-off angle is set at 20 degrees. Two observation periods are considered, one with the best geometry of 5 satellites and a PDOP value of 3.5, the other with the worst geometry of only 3 satellites and PDOP undefined.

A receiver is located on a control station onshore and the position of the ferry is computed by translocation. The observables are singly-differenced pseudo-ranges and the difference between the receiver clock offsets is the only nuisance parameter introduced in the measurement model. The interval between measurement epochs is 15sec. The stochastic model assumes a standard deviation of 5m for each observation.

The initial conditions are given in Table 6.1, the assumed knowledge of the horizontal position of the Manly pier is only vague. The three kinematic models investigated are given in Table 6.2. In Model 1, the motion is assumed almost unpredictable in E, N and T. The mean error selected for the horizontal accelerations in Model 2 is much smaller than the changes defined in Figure 6.4. In this respect, Model 3 is based on a more realistic assumption.

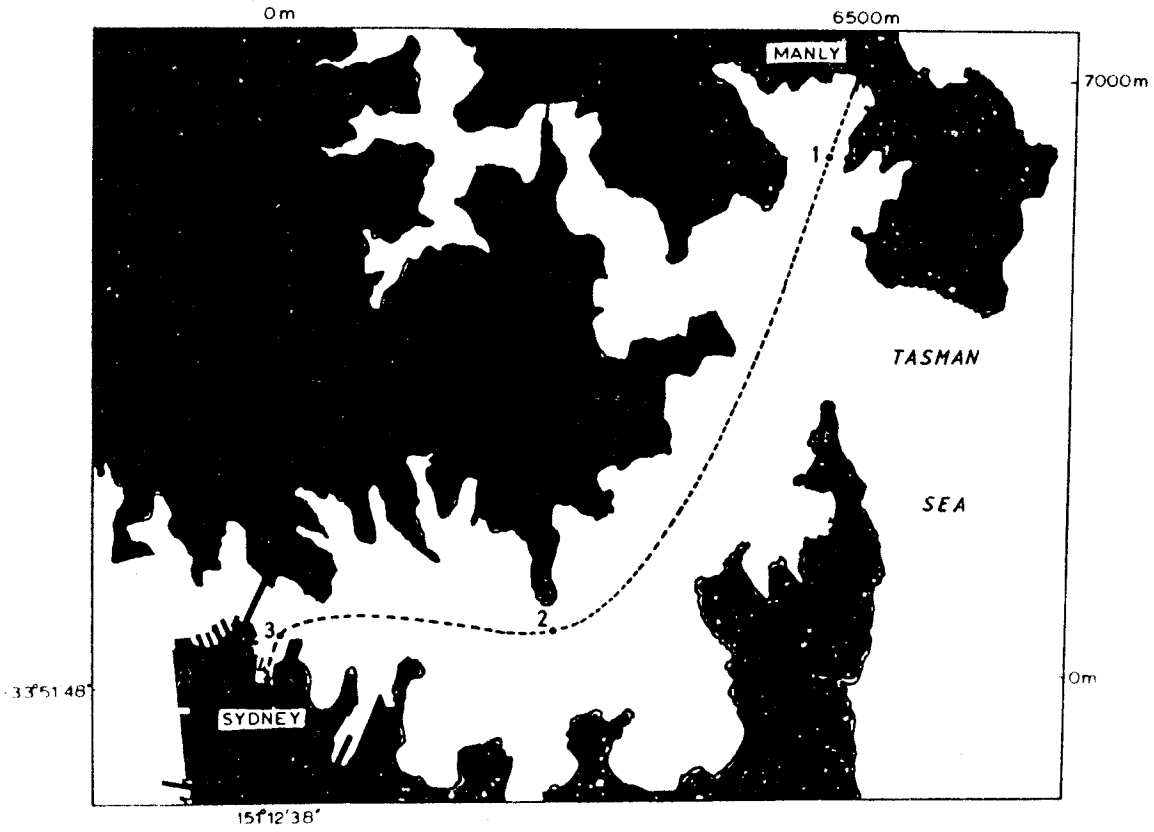


Figure 6.3: Manly-Sydney Route

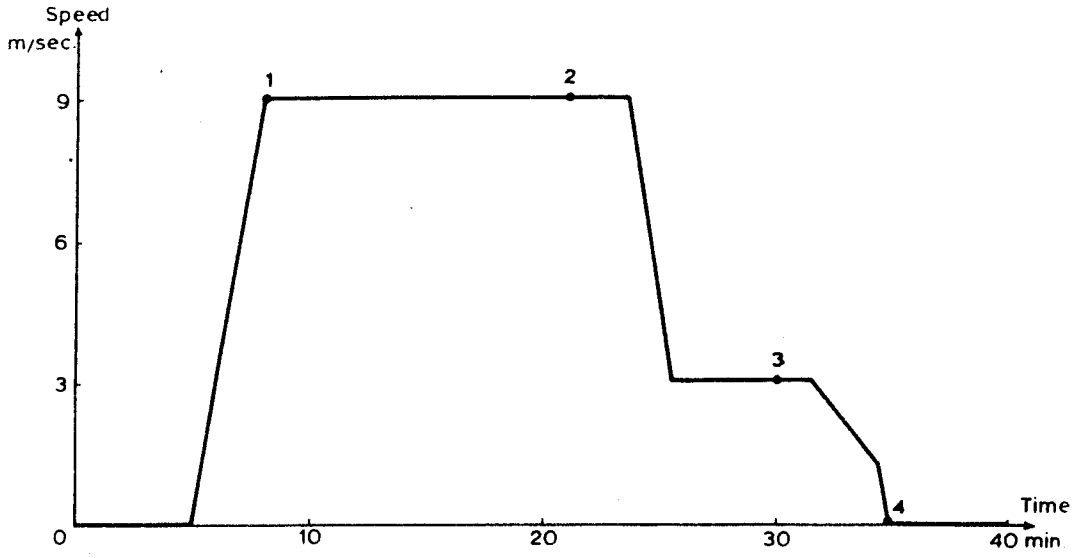


Figure 6.4: Speed Profile

The heading and its mean error can be estimated, even if only position measurements are available. This is achieved without having to explicitly divide the change in position by the time increment, because the estimates of velocity and position are correlated through the kinematic model and estimated together. In the first model however, the heading cannot be computed as the velocities are not estimated.

Table 6.1

INITIAL CONDITIONS

		Mean Error on:	
Coord.	Position [m]	Velocity [m/sec]	Acceleration [m/sec^2]
E	1000	0.3	0.1
N	1000	0.3	0.1
V	3	0	0
T	1000	0	0

Table 6.2

SELECTED KINEMATIC MODELS

Model 1	Coord.	Degree	Order	Mean Error
	E	0	1	100 m/sec
	N	0	1	100 m/sec
	V	0	0	3 m
	T	0	0	1000 m
Model 2	Coord.	Degree	Order	Mean Error
	E	2	2	0.01 m/sec^2
	N	2	2	0.01 m/sec^2
	V	0	0	3 m
	T	0	0	1000 m
Model 3	Coord.	Degree	Order	Mean Error
	E	2	2	0.1 m/sec^2
	N	2	2	0.1 m/sec^2
	V	0	0	3 m
	T	0	0	1000 m

6.3.2. Results and Comments

Four points have been selected, where either the heading or the speed undergoes sudden changes (Figures 6.3 and 6.4). The deviations from the original trajectory and their mean error are given in Tables 6.3 and 6.4. There

are no deviations for the first model because perfect observations have been used and no inertia is introduced by the extremely "soft" kinematic model. However, the larger mean errors indicate a higher sensitivity to random measurement noise. Although the vertical motion is considered white noise, the mean error of 3m sets a strong limitation to possible variations. With a standard deviation of 1000m on the vertical motion, the mean error in the adjusted coordinates E, N and V at the time of worst PDOP would be 152, 414 and 688m respectively. This clearly shows how only approximate knowledge of a position parameter may affect the other parameters. Because of the critical location of the points, all the deviations shown for Models 2 and 3 are rather large. Between these points, even the second filter gives a much better response.

Table 6.3

DEVIATIONS AND THEIR MEAN ERROR FOR BEST PDOP [m]

Model 1	Point	East	North	Vertical
	1	0.0 ± 3.7	0.0 ± 4.8	0.0 ± 2.9
	2	0.0 ± 3.7	0.0 ± 5.0	0.0 ± 2.9
	3	0.0 ± 3.7	0.0 ± 5.1	0.0 ± 2.9
	4	0.0 ± 3.7	0.0 ± 5.1	0.0 ± 2.9

Model 2	Point	East	North	Vertical
	1	15.0 ± 2.4	69.3 ± 2.9	-4.9 ± 2.9
	2	22.2 ± 2.4	-84.1 ± 3.1	7.3 ± 2.9
	3	-18.5 ± 2.4	26.6 ± 3.1	-5.9 ± 2.9
	4	-2.8 ± 2.4	-24.9 ± 3.1	1.2 ± 2.9

Model 3	Point	East	North	Vertical
	1	0.7 ± 3.4	3.9 ± 4.3	-0.3 ± 2.9
	2	0.7 ± 3.4	-7.5 ± 4.4	0.7 ± 2.9
	3	-1.2 ± 3.4	1.1 ± 4.5	-0.4 ± 2.9
	4	-0.3 ± 3.4	-3.1 ± 4.5	0.2 ± 2.9

DEVIATIONS OF THE HEADING, WITH MEAN ERROR [deg]

Point	Model 2	Model 3
1	2.0 ± 0.6	0.4 ± 2.4
2	-26.8 ± 0.5	-7.4 ± 2.4
3	23.8 ± 1.5	5.3 ± 6.8
4	-1.4 ± 3.0	3.0 ±42.4

Table 6.4

DEVIATIONS AND THEIR MEAN ERROR FOR WORST PDOP [m]

Model 1	Point	East	North	Vertical
	1	0.0 ± 4.8	0.0 ± 6.7	0.0 ± 3.0
	2	0.0 ± 7.2	0.0 ± 9.7	0.0 ± 3.0
	3	0.0 ± 7.9	0.0 ±12.1	0.0 ± 3.0
	4	0.0 ± 8.6	0.0 ±14.1	0.0 ± 3.0

Model 2	Point	East	North	Vertical
	1	21.7 ± 3.0	97.7 ± 3.9	29.5 ± 2.9
	2	15.7 ± 4.1	-128.7 ± 5.2	41.5 ± 3.0
	3	-18.2 ± 4.3	25.8 ± 6.1	59.5 ± 3.0
	4	-32.2 ± 4.5	-90.2 ± 6.8	59.7 ± 3.0

Model 3	Point	East	North	Vertical
	1	1.2 ± 4.2	6.1 ± 5.7	2.2 ± 2.9
	2	-3.3 ± 6.1	-17.1 ± 7.9	3.7 ± 3.0
	3	-2.3 ± 6.4	0.0 ± 9.5	3.2 ± 3.0
	4	-4.1 ± 6.8	-10.8 ±10.7	4.5 ± 3.0

DEVIATIONS OF THE HEADING, WITH MEAN ERROR [deg]

Point	Model 2	Model 3
1	3.2 ± 0.8	0.6 ± 2.7
2	-30.9 ± 0.5	-11.0 ± 2.8
3	29.5 ± 1.9	6.4 ± 7.4
4	5.1 ± 2.3	7.5 ±26.4

It must be pointed out that the mean errors obtained would be the same if real data were processed, because they depend solely on the stochastic models defined for both the measurement and kinematic models. However, the agreement of the functional and stochastic models can be monitored using the ratio mean error *a posteriori* / mean error *a priori*, output after each adjustment step. This ratio sometimes exceeds 10 for Model 2 and indicates its inadequacy: it is too "hard". For Model 3, the ratio rarely exceeds 2, hence, the errors introduced by the kinematic model remain at all times smaller than those expected from random measurement errors (factor 3). The large uncertainty in the heading at the arrival in Sydney Cove (pt 4) for Model 3 is due to the quick response of this filter to the final deceleration, and the very small estimated velocity makes the heading irrelevant, as expected.

Using the ratio mean error *a posteriori* / mean error *a priori*, the system noise could be rescaled, in effect implementing an adaptive filter. The risk inherent in this procedure must not be overlooked and a minimum error must be defined for the standard deviation of the driving noise on any of the parameters. Otherwise, after several epochs of good agreement between prediction and measurements, the expected discrepancies may become very small and preclude the filter from reacting to an actual change in the movement. This typically would occur at the end of the first straight line (pt 2): starting at Manly with Model 3, the small size of the residuals on the line would cause a reduction in the mean error of the driving noise and the curve would be approached with a kinematic model close to Model 2.

Tables 6.3 and 6.4 show that the results depend on the geometry of the satellites, even when the future constellation of 18 GPS satellites is considered. Without the stringent limitation imposed on the vertical motion, the results obtained using different satellite configurations would present a much wider range of variation.

7. PROCESSING OF REAL DATA

While designing a software package like DYNAMO is already a demanding task, processing real GPS kinematic data is almost as demanding. This is due to a number of reasons:

- The **nature of the GPS data** -- different biases must be taken into account for each of the various observables, such as satellite and receiver clock offsets and drifts.

- The **volume of the data** -- a half hour session involving 5 satellites, 2 receivers and recording 2 types of measurements every 2 seconds generates 18000 measurements !

- The **number of options** -- this is both the strength and the difficulty in the use of DYNAMO. The filtered and, to a lesser extent, the smoothed solution can be changed through a large range of settings (a process known as "tuning") resulting in discernable changes in the trajectory. Obtaining the "best" solution therefore requires skill and experience.

Three differential GPS experiments are presented in this chapter. The principal aim is to demonstrate the validity of the concepts discussed in the previous chapters.

7.1. SYDNEY AIRPORT EXPERIMENT

7.1.1. Data Collection

This was a joint experiment involving staff and resources of the UNSW, the University of the Federal Armed Forces (UFAF), Munich, West Germany, and the Royal Australian Survey Corps (RaSVY). RaSVY provided the TI-4100 GPS receivers. On the 18th October 1987 two TI-4100 GPS receivers were deployed: one mounted in an army jeep, while the other was set up at a fixed reference site about 10 kilometres away. Observations commenced with 25 minutes of static tracking by the two receivers. Then the jeep was driven at a speed up to 60 kph around the perimeter of the main north-south runway of Sydney's Kingsford Smith Airport (Figure 7.1), stopping at each of 4 survey marks, where the antenna was removed from the vehicle and placed on a

prepositioned tripod for a period of about one minute. Once this traverse was completed the jeep was driven back over the same approximate route halting at the same points. Upon return to the starting point, a further 10 minutes of static observations were recorded. Because the TI-4100 GPS receiver can only track up to 4 satellites at any one time, the observation window was selected to ensure the best geometry (PDOP was approximately 4). Pseudo-range, phase and phase-rate data were collected on both L1 and L2 frequencies, at 3 second intervals. The data were recorded on cassette tapes, which were then transcribed onto 5.25" diskettes.

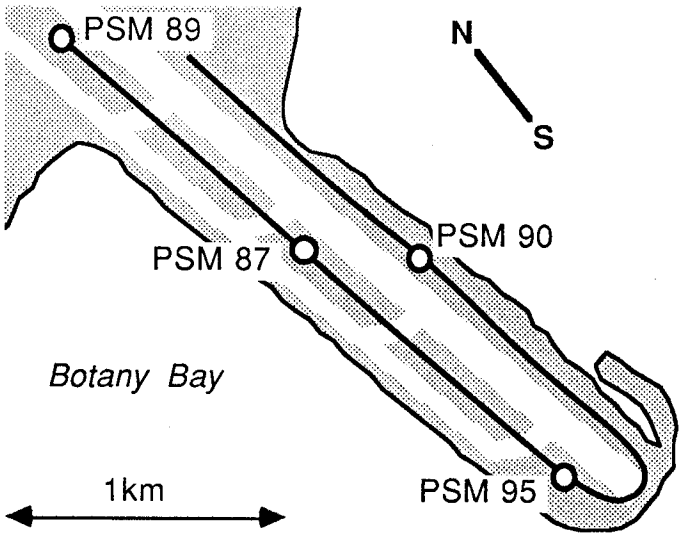


Figure 7.1: Sydney Airport TI4100 GPS Experiment

7.1.2. Data Processing

The data has first been processed using software developed at UFAF, and centimetre level positioning accuracies have been obtained from the combined processing of pseudo-range and carrier phase data from both L-band frequencies (STOLZ & HEIN, 1988). Although data from both frequencies could be considered with DYNAMO, the analysis at UNSW concentrated on using only the L1 pseudo-range and phase-rate data.

The position and velocity (=zero) at the reference station were held fixed (to define the space datum). The bias and bias drift (both equal to zero) were held fixed for one of the satellites (to define the time datum). The following parameters were estimated:

- East, North and Vertical position components at the mobile site (3 parameters). If relevant, also East and North velocities (that is, a total of 5 parameters).
- Receiver clock offset and drift at both the reference and mobile sites (4 parameters).
- Satellite range bias and drift for 3 of the 4 satellites (6 parameters).

Thus, a Kalman filter with up to 15 states was used, and three solutions were attempted:

(1) **Differential Navigation Solution** -- only the pseudo-ranges are used, and independent positions or fixes were computed at each epoch. This is a "no filter" solution. In practice, the filter is made ineffective by considering a very large process noise on the coordinates of the mobile antenna, see Section 6.3.2. No direct velocity estimation can be obtained from this approach.

(2) **Differential Filtered Solution** -- both the pseudo-range and the phase-rate measurements are used. The following parameters or "settings" have been selected for the various stochastic models.

Mean error for the different types of measurements:

- pseudo-range: 1m
- phase-rate: 0.05m/sec

Driving noise for the different components:

- East and North parameters: white noise on the acceleration term of 0.5m/sec^2 (second order random walk for position).
- Vertical component: white noise on the velocity term of 0.2m/sec (random walk for position).
- Receiver clock bias parameters: white noise on the acceleration term of 1m/sec^2 (second order random walk for clock offset).
- Satellite bias parameters: white noise on the acceleration term of 0.05m/sec^2 (second order random walk for "clock" offset).

The choice of these values is based on a thorough inspection of the measurement residuals and coordinate increments resulting from the adjustment steps. This is an iterative process repeated until residuals for the adjusted measurements and parameters are consistent with their modelling, that is, the ratio *a posteriori / a priori* for all types of residuals is approximately equal to 1. Hence, the data themselves indicate which models are appropriate. In practice, many solutions were inspected until confidence was gained in selecting the appropriate noise models for the various parameters.

(3) **Differential Smoothed Solution** -- using all the data processed *backwards* in time. See Sections 3.6.2 and 5.2.3 for details.

The same filter settings were used throughout the observing session. In particular, the static mode was not assumed during the short stays on the survey marks. However, the results obtained for the static positions and the trajectory between the survey marks are discussed separately, in Sections 7.1.3 and 7.1.4 respectively.

7.1.3. Survey Marks

Repeatability: as each of the 4 Permanent Survey Marks (PSM) was visited twice during the experiment, a test on repeatability was possible. Figure 7.2 shows the discrepancies (E, N & V) for the 3 solutions attempted.

Accuracy: From the data processing executed at UFAF, the coordinates of the reference station and the 4 PSMs were known. As these coordinates can be assumed to have precisions at the centimetre level, they are regarded as ground truth, against which DYNAMO results can be tested. For each stop at the marks, about 20 sets of coordinates are available. The epoch closest to the middle of the short static session was selected for the tests and no averaging took place. Table 7.1 shows the comparison between DYNAMO coordinates and the ground truth, for the 3 solutions attempted. The residuals are expressed in a local topocentric system oriented East, North and Vertical. The standard deviation is not centred around the mean (that is, σ is computed as $\sum x^2/n$). Note that the mean is always much smaller than the standard deviation, indicating the absence of systematic errors affecting the entire session.

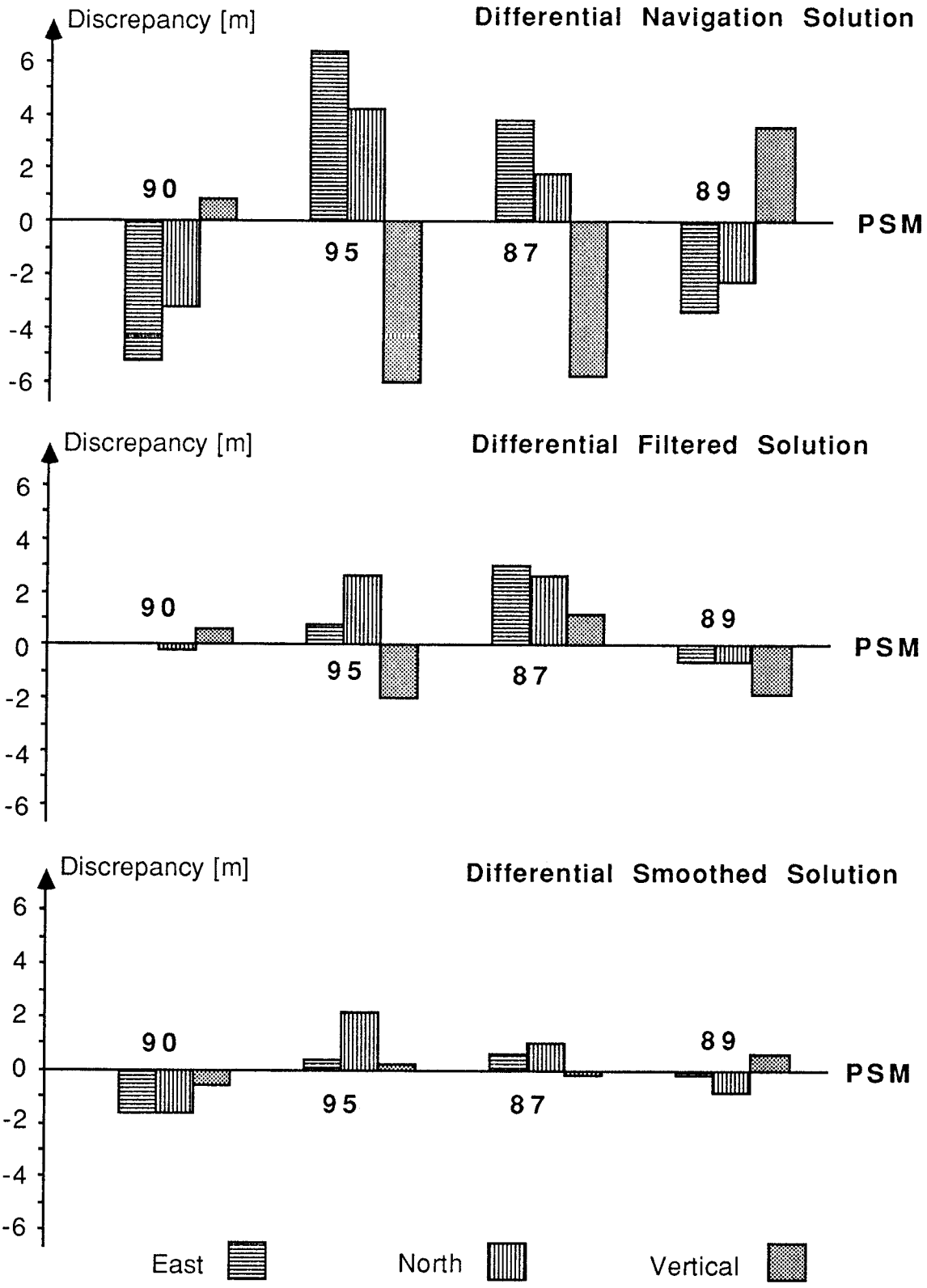


Figure 7.2: Discrepancies in Repeatability Test

Table 7.1: Discrepancies between coordinates from DYNAMO and Ground Truth, in metres

(1) **Differential Navigation Solution**:

Station	East	North	Vertical
PSM 90 A	0.86	1.07	-2.71
PSM 95 A	-1.37	1.71	5.87
PSM 87 A	-1.71	-0.92	3.50
PSM 89 A	2.82	2.65	-2.64
PSM 89 B	-0.55	0.41	0.88
PSM 87 B	2.19	0.87	-2.34
PSM 95 B	4.99	5.82	-0.11
PSM 90 B	-4.28	-2.08	-1.98
Mean	0.37	1.19	0.06
Mean error (1σ)	2.82	2.55	3.02

(2) **Differential Filtered Solution**:

Station	East	North	Vertical
PSM 90 A	-0.21	1.22	0.43
PSM 95 A	0.34	0.78	2.95
PSM 87 A	-0.16	-0.76	0.01
PSM 89 A	0.74	1.10	-1.97
PSM 89 B	0.15	0.57	-3.74
PSM 87 B	2.78	1.88	1.11
PSM 95 B	1.15	3.32	0.97
PSM 90 B	-0.14	1.03	0.91
Mean	0.58	1.14	0.08
Mean error (1σ)	1.12	1.59	1.95

(3) **Differential Smoothed Solution**:

Station	East	North	Vertical
PSM 90 A	0.53	1.44	0.91
PSM 95 A	0.57	0.68	1.94
PSM 87 A	0.74	-0.07	1.33
PSM 89 A	0.55	0.90	-2.73
PSM 89 B	0.36	0.14	-2.04
PSM 87 B	1.40	0.97	1.10
PSM 95 B	0.86	2.89	2.19
PSM 90 B	-1.00	-0.26	0.27
Mean	0.50	0.84	0.37
Mean error (1σ)	0.82	1.27	1.76

7.1.4. Trajectory

To illustrate the effect on the trajectory of the mobile receiver (as against the tests related to the PSMs, during which the roving receiver was stationary), a portion between PSM 87 and PSM 95, on the backward run, is shown in Figure 7.3. The plot of relative height change of the trajectory for the UFAF and the 3 DYNAMO solutions illustrates the behaviour of the different solutions. There were some losses of lock during the movement of the antenna and the sensitivity of the UFAF solution to cycle slips is illustrated by the jumps in the relative height. These jumps cannot be real as the ground surface was very flat. Therefore, the UFAF trajectory can only be regarded as "ground truth" outside these troubled periods.

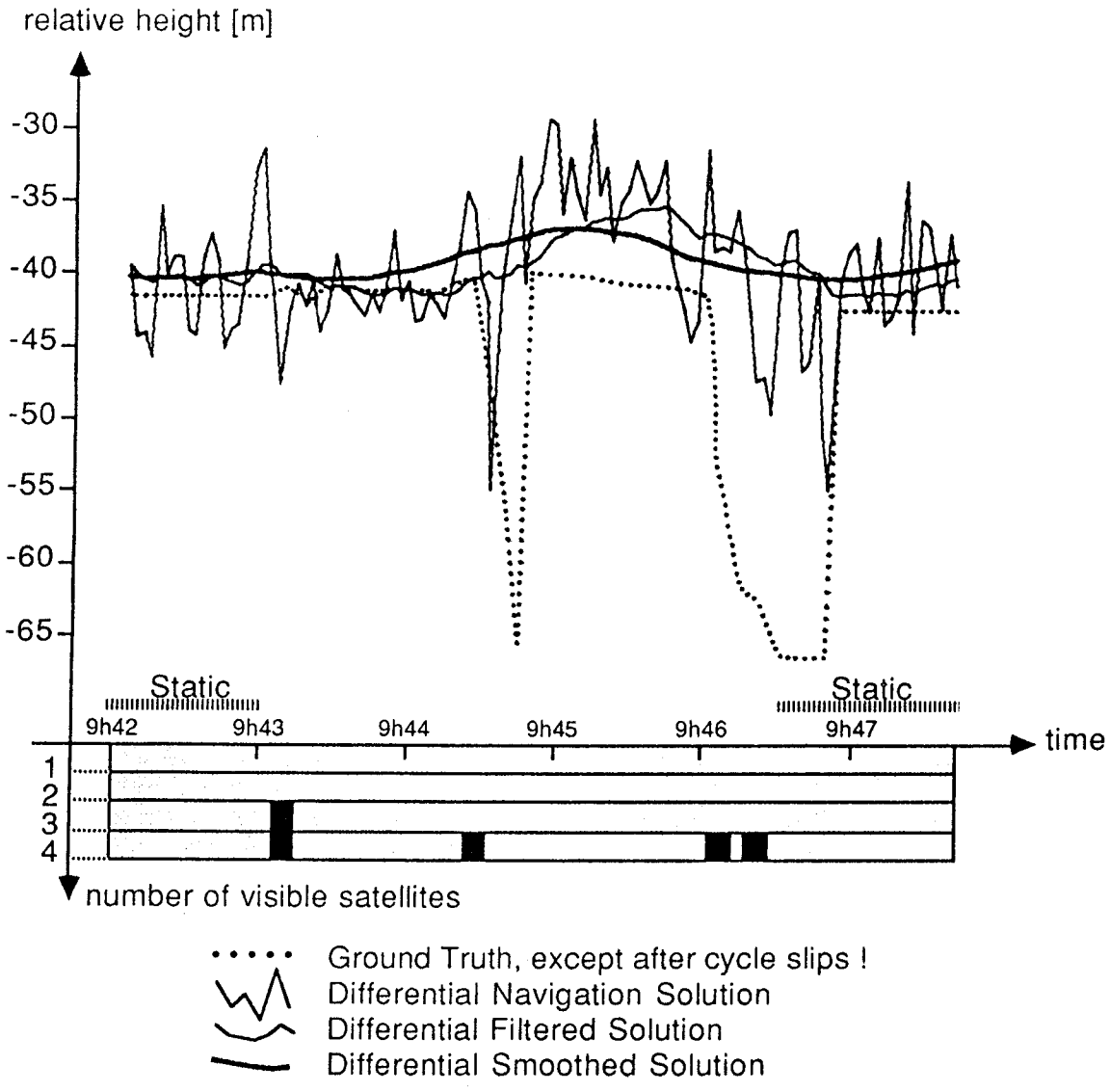


Figure 7.3: Relative Height Between PSM 87 and PSM 95

The differences of the three DYNAMO solutions versus the ground truth were computed over one hour, that is, 1200 measurement epochs. 120 epochs, or 10%, had to be eliminated because of obvious errors in the "ground truth", all associated with losses on lock on one or more satellites. The instantaneous differential navigation solution yields a relatively noisy trajectory, as the relative height jumps by well over 10 metres from epoch to epoch. The filtered and smoothed solutions are much better behaved. Note the delay in the filtered solution with respect to the smoothed solution. The statistics of the comparison are given in Table 7.2. The short static sessions are included in the comparison.

Table 7.2: Statistics of the Differences in Coordinates Between DYNAMO Solutions and Ground Truth, in metres

(1) Diff. Navigation Solution:	E	N	V
Mean	0.40	0.25	0.66
Mean Error (1σ)	3.48	2.36	4.62
(2) Diff. Filtered Solution:	E	N	V
Mean	0.34	0.21	0.35
Mean Error (1σ)	1.99	1.20	2.69
(3) Diff. Smoothed Solution:	E	N	V
Mean	0.34	0.24	0.58
Mean Error (1σ)	1.65	1.07	2.03

Again, the mean is always much smaller than the mean error. However, with respect to time, the deviations from the ground truth tend to exhibit a "wavelike" behaviour. This is especially pronounced for the smoothed solution, and reflects the stronger correlation between solutions for adjacent epochs. For all three solutions the mean error is in good agreement with that obtained for the static locations. Hence the motion does not reduce the effectiveness of the filter. A further comparison can be made with the precision estimated by DYNAMO. When all 4 satellites have been tracked for some time, the steady state is reached and the estimated mean error in the coordinates is stable. However, the precision in the steady state depends on the strength of the satellite geometry, and hence changes during the tracking session. The precision estimates in the steady state at the start (09:00), the middle (09:30)

and the end (10:00) of the considered session are given in Table 7.3. For the different solutions, the mean is computed using all epochs with 4 satellites tracked. These precision estimates are rather optimistic, as they are associated with optimal tracking conditions. However, it must be pointed out that missing satellites cause a prompt increase in the mean error of the estimated coordinates, stronger than the actual deterioration of the estimated coordinates. Therefore, the user is warned when problems are encountered.

Table 7.3: Estimated Precision of Coordinates in the Steady State for Various DYNAMO Solutions, in metres

(1) Diff. Navigation Solution:	E	N	V
Start	3.50	1.32	4.51
Middle	3.10	1.93	4.51
End	6.70	5.12	5.06
Session Mean (1σ)	3.73	2.30	4.59
(2) Diff. Filtered Solution:	**E**	**N**	**V**
Start	1.46	0.76	1.48
Middle	1.43	0.97	1.21
End	2.33	1.86	1.25
Session Mean (1σ)	1.56	1.07	1.21
(3) Diff. Smoothed Solution:	**E**	**N**	**V**
Start	1.04	0.61	0.94
Middle	1.07	0.75	0.87
End	1.75	1.40	1.03
Session Mean (1σ)	1.16	0.82	0.88

From a comparison with the accuracy (Table 7.1), these precision estimates can be considered realistic, although the precision in the height determination tends to be over optimistic for the filtered and smoothed solutions. This is illustrated in Figure 7.4, where the results of Tables 7.1, 7.2 and 7.3 have been combined. Overall the Kalman Filter results are encouraging. There appears to be a two-fold improvement in accuracy and repeatability between the standard Differential Navigation solution based on pseudo-range data alone, and the best solution using the Differential Filter technique. Furthermore, the smoothed solution is, as expected, the best solution.

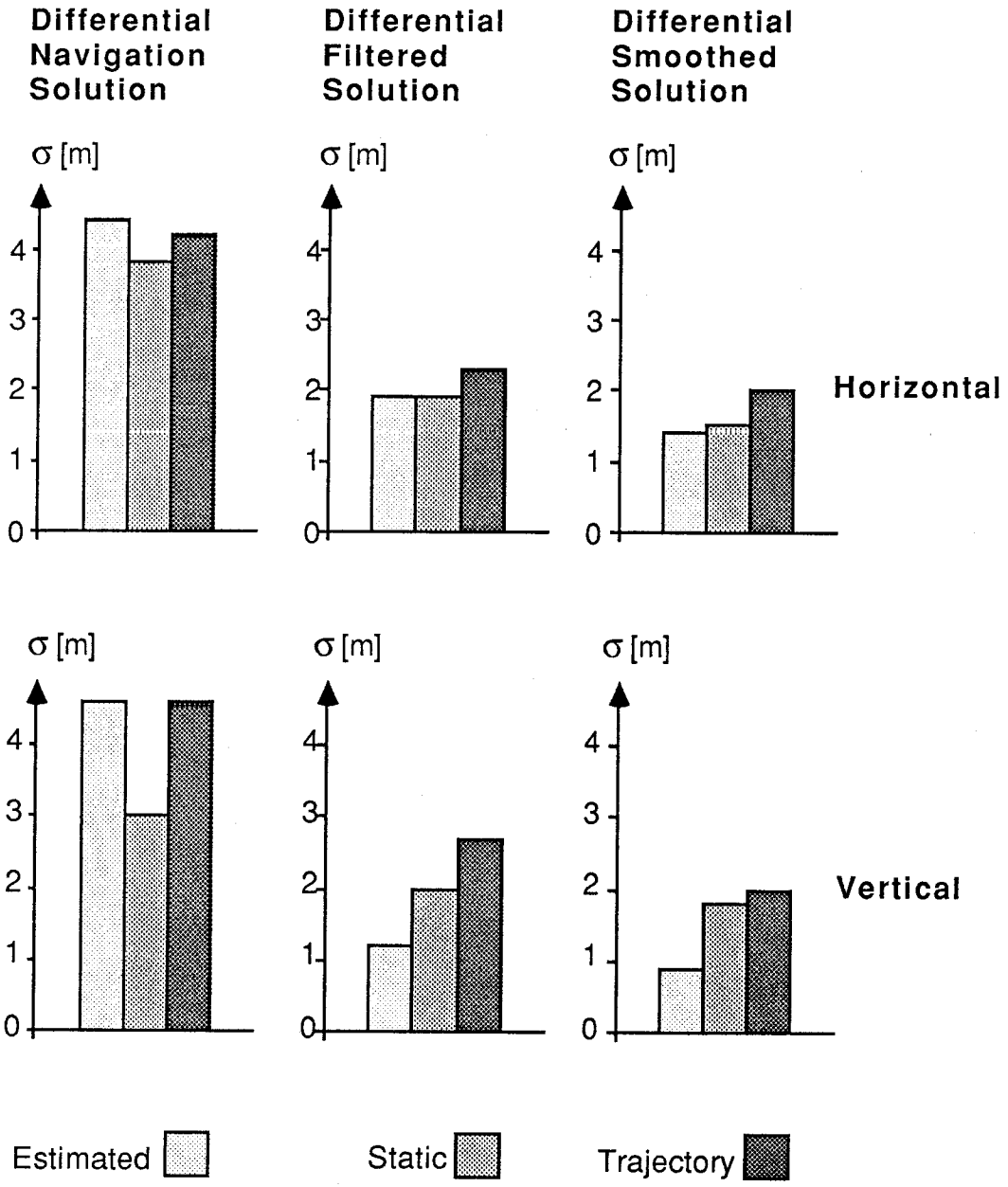


Figure 7.4: Accuracy of Various Processing Methods

7.1.5. Velocity

Figure 7.5 illustrates a filtered and a smoothed solution, with respect to the velocity profile between PSM 87 and PSM 95. In this particular case, the phase-rate data is used in the smoothed solution, but not in the filtered solution. Both profiles are very similar: this shows that the filter reacts quickly to a change in position indicated by the pseudo-range data alone, and updates the velocity estimates with only a small delay. However in the steep

sections of the graphs (large accelerations) a difference of say 1m/sec between the velocity estimates is hardly noticeable, although such a difference has a significant effect on coordinates predicted over a few seconds. In such cases the precise velocity information provided by the phase-rate data as well as the smoothing process gain in importance. A filtered solution including the phase-rate data was also attempted: the resulting profile is closer to the smoothed solution, and therefore cannot be shown together with the other profiles in Figure 7.5.

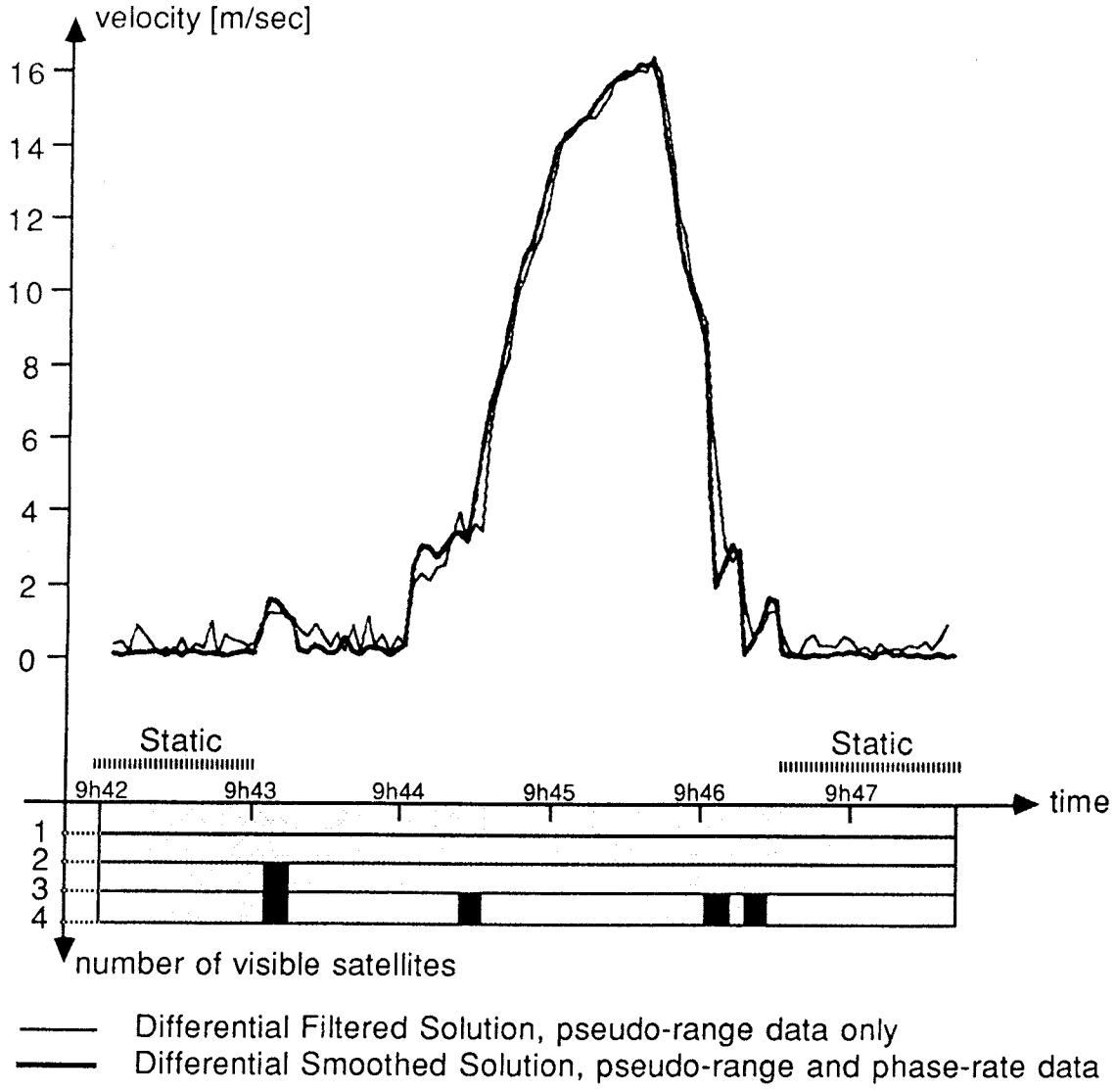


Figure 7.5: Velocity Profile Between PSM 87 and PSM 95

7.2. RIG EXPERIMENT

7.2.1. Campaign Design and Data Collection

On the 17th September 1987, one WM101 GPS receiver was mounted on the SEDCO 600 Drilling Rig operating off Sydney, while two others were set up at trig stations onshore. This experiment involved a different kinematic behaviour to that of the experiment reported in Section 7.1. The rig moved with the tides and swell, with the vertical and horizontal motion being of the order of 1 metre, with a period of about 20 seconds. This joint WILD/UNSW experiment was conducted to determine whether:

1. The WM101 was able to successfully make measurements on a moving platform.
2. The PoPS post processing software was able to process carrier phase data from a moving offshore rig.
3. The differential pseudo-range position of the drill stem would compare with that determined by the acoustic transponders presently used for positioning the rig.

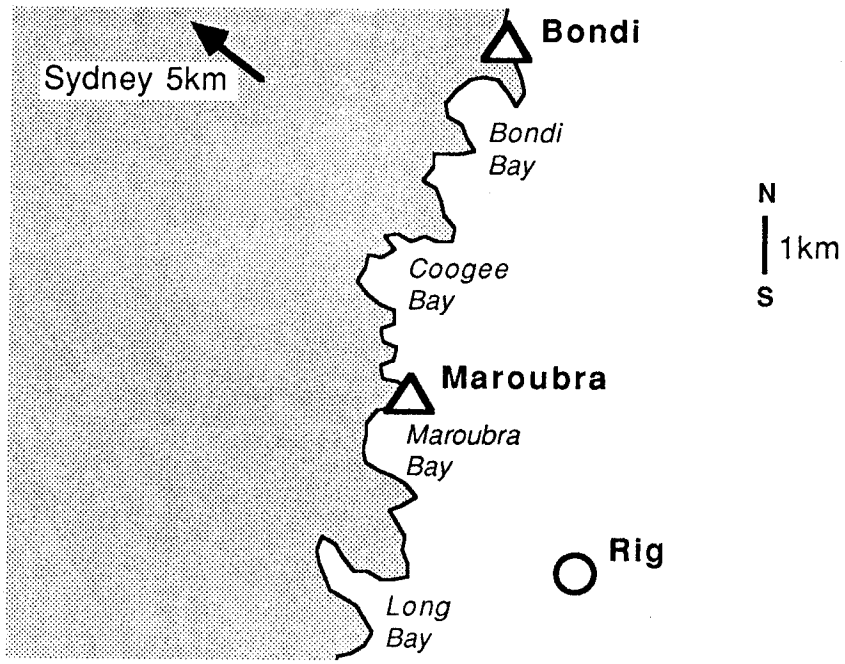


Figure 7.6: Rig Survey

The challenge of this experiment was in using the WM101 GPS receivers in a kinematic environment, as this instrument was designed specifically for use as a (static) surveying tool. There were 3 major problems:

- The 60 second compressed data normally output for subsequent processing in GPS surveying software such as PoPS is not suitable for kinematic processing, as the period of the swell is approximately 20 seconds.
- The 2 second compressed data, referred to as "raw", is only intended for factory check purposes. No use of the raw data for positioning had been reported.
- Because the WM101 receiver has only 3 hardware channels, they must switch between two satellites when more than 3 satellites are observed.

During the evening, three tracking sessions took place, each using a different observation mode:

(1) In the first session, 75 minutes of data were collected to all available satellites (up to 6) for a 60 second compression interval. Many receiver warnings were given to indicate that oscillator noise problems and data noise on each of the observing channels were being experienced. Subsequent processing with a special "data assessment" software package indicated that the receiver had difficulty in keeping track on the satellites. Losses of lock frequently occurred and this was considered to be a switching problem encountered when two satellites shared one observing channel.

(2) Then, the receiver mounted on the rig was forced to record data from only 3 satellites, to avoid switching. 45 minutes of such data were recorded. Again, many receiver warnings indicated oscillator and data noise problems. However, losses on lock on the satellites were very scarce.

(3) In the third session, raw satellite data was observed by the three receivers to all available satellites (4), for 30 minutes. Very few receiver warnings were given.

7.2.2. Processing with PoPS

a) Carrier Phase

For the trig stations Bondi and Maroubra, separated by a distance of about 6km (see Figure 7.6), three hours of 60 second carrier phase data from all visible satellites were processed using the PoPS GPS surveying software package, using the standard procedures. The resulting difference in coordinates are in excellent agreement with solutions from previous GPS campaigns and official coordinates. With an accuracy of a few mm, these baseline results were considered as ground truth data.

Although the movements of the rig were small, processing of the carrier phase observations with PoPS was not possible. The problem with a software package specifically developed for surveying applications is that the antennae are assumed to be static. As a result, any motion is interpreted as cycle slips. Typically, the mean error of the double difference residuals was approximately 5 cycles -- which matches the order of magnitude of the expected rig movement -- and cycle slips smaller than about 10 cycles could not be repaired. This was the case even when each hardware channel was dedicated to one satellite (session 2). With carrier phase tracking divided into many short sessions, unambiguous ranges could not be derived and no acceptable position could be obtained. The receiver warnings given during the recording of 60 second compressed data suggest that the software installed in the WM101 also assumes a static behaviour of the antenna for the data compression, hence interpreting any antenna movement as data noise.

b) Pseudo-ranges

The processing of pseudo-ranges was more successful. Prior to the carrier phase adjustment and for each site, PoPS computes a static single point position using the 60 second pseudo-ranges (in order to derive a model of the receiver clock offset as a polynomial). The static assumption is not a problem in this case, because the noise on the navigation solution is much larger than the swell. Relative positioning results were achieved by computing the differences between single point position solutions.

Baseline: with respect to the ground truth, the errors in the determination of the baseline Bondi Maroubra for the first two sessions are:

Error in metres	East	North	Vertical
Session 1	+9.71	-5.00	-26.99
Session 2	-25.90	-0.23	-9.02

These results are not very satisfactory, but it must be emphasised that the WM101 receiver and the PoPS software have not been designed for precise positioning using pseudo-range data.

Rig: the position of the drill stem is determined using acoustic transponders with a metre level accuracy. Considering the results obtained for the baseline with GPS, the "acoustic" solution can be regarded as ground truth. However, as the height was not determined with the transponders, a comparison is only possible in the horizontal plane. Both Bondi and Maroubra were used as reference sites, and the solutions were averaged.

Error in metres	East	North
Session 1	-16.84	-14.79
Session 2	+7.53	+25.44

The accuracy of these results can be compared with that obtained for the (static) baseline onshore.

7.2.3. Asynchronous Measurements

An additional and unexpected problem arose during the processing with DYNAMO: namely the time tags of the raw measurements varied in an apparently random fashion. Whilst measurements involving one particuliar receiver and satellite pair were recorded at exactly the nominal 2 second output rate (4 seconds when the channel was switching), measurements from different satellites were not recorded at the same time by any receiver, and measurements from one particular satellite were not recorded at the same time by the different receivers. In a filtering process, this means that measurements must be processed in distinct adjustment steps. This is easily achieved with DYNAMO, but the asynchronism does affect the results and their precision. This can be illustrated with a simple, one-dimensional example.

The initial position of an object moving along a line between A and B is known with a precision of 2m. The movement is modelled as a random walk with the

velocity being a zero mean white noise with standard deviation 0.707m/sec, that is, a spectral density of $0.5m^2$/sec. In the first case, distances to the object are measured from both A and B at each even second. In the second case, distances from A are measured at each even second and distances from B are measured at each odd second. The evolution of the variance is compared in Table 7.4. The epochs are denoted by the number of seconds from the initial time and the sign indicate whether the adjustment has been performed or not. Minus is for predicted values and plus for adjusted values.

Table 7.4: Variance of the Position in m^2

Epoch	Simultaneous Measurements	Alternate Measurements
0-	2.000	2.000
0+	0.400	0.667
1-	0.900	1.167
1+		0.538
2-	1.400	1.038
2+	0.368	0.509
3-	0.868	1.009
3+		0.502
4-	1.368	1.002
4+	0.366	0.501
5-	0.866	1.001
5+		0.500
6-	1.366	1.000
6+	0.366	0.500

The steady state is quickly reached in both cases and the situation is illustrated in Figure 7.7. All other things remaining equal, asynchronous measurements often yield better predicted values, as the time span without new incoming measurements is shorter. However, adjusted values can never be as good as in the case of simultaneous measurements, due to the process noise affecting the position between each pair of measurements. As we are generally interested in adjusted values, rather than predicted ones, asynchronous measurements are a disadvantage. Note that in the absence of process noise, asynchronous measurements yield exactly the same precision as simultaneous ones. However, this case is rarely encountered with GPS:

clocks are generally not perfectly predictable, even if coordinates are, as for example in the static case.

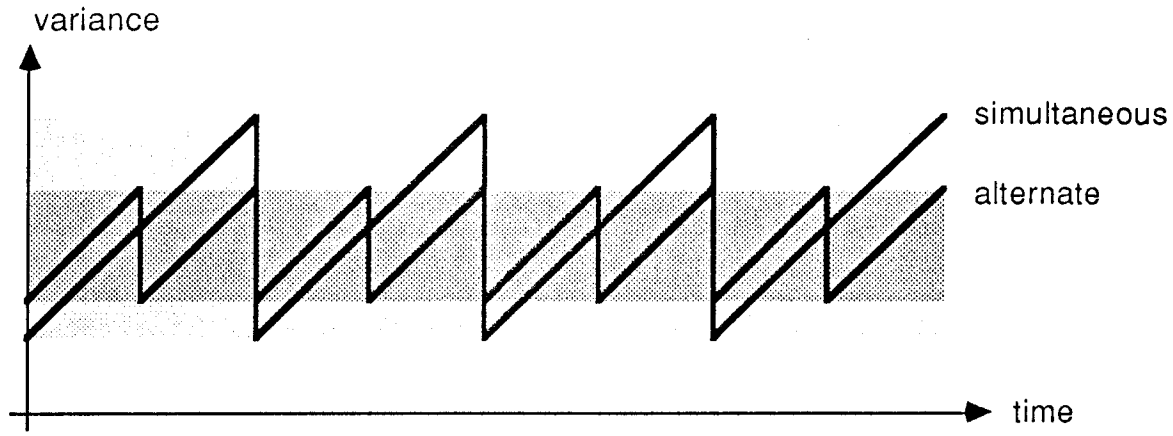


Figure 7.7: Alternate versus Simultaneous Measurements

Asynchronous data can only be processed successfully with a filter. Because only little information is available at once, all biases can never be resolved with only one epoch of data. A proper kinematic modelling of the parameters over time is therefore essential. For one of the WM101 receivers involved in this experiment, the (well-behaved) clock drift amounted to 60m/sec (that is, a relative drift of $2 \cdot 10^{-7}$ or 0.02sec/day). Assuming the simultaneity of measurements actually recorded 1/100sec apart causes an error of 60cm. DYNAMO offers the option to process each measurement individually, and this possibility was resorted to for all subsequent processing.

7.2.4. Results Using DYNAMO

At the time when raw data was recorded, only 4 satellites could be tracked, and the PDOP was about 5. Table 7.5 presents the evolution of the static differential pseudo-range solution for the baseline onshore, as compared with the ground truth, and along with their 1σ internal precision. One satellite clock bias is used as time datum (offset = drift = 0) and the other clock offsets are modelled as second order random walks with driving white noises of respectively $0.05 m/sec^2$ and $1 m/sec^2$ for satellites and receivers. In this static positioning, due to the accumulation of data pertaining to the same parameters, the effect of incorporating phase-rate is negligible after a few epochs have been processed.

Table 7.5: Ground Truth and Discrepancies for the 6km Baseline Onshore (in m)

	East	North	Vertical
Ground Truth	-1950.312	-5957.185	-26.517
after 1 min	-2.42 ± 0.87	+6.49 ± 0.70	+7.51 ± 3.05
after 3 min	-0.65 ± 0.50	+7.28 ± 0.41	+9.32 ± 1.80
after 5 min	-0.50 ± 0.39	+7.56 ± 0.32	+9.71 ± 1.40
after 10 min	-0.86 ± 0.27	+7.61 ± 0.23	+9.62 ± 0.99
after 20 min	-0.76 ± 0.19	+7.54 ± 0.16	+9.95 ± 0.67

The results from Table 7.5 clearly indicate that GPS positioning inaccuracies are dominated by systematic biases, rather than random measurement noise. When enforcing the ground truth as solution, that is holding both trig stations fixed, time series of range residuals involving one receiver and satellite pair are systematically biased by up to a few metres. According to CHEZELLES (1988), see Section 1.3.3, such a range bias magnitude should be expected. As compared with the pseudo-range positioning capability of PoPS, better results have been obtained with a shorter session involving less satellites. The main merit of DYNAMO is its capability to process the 2 second measurement data.

The raw data collected on the rig could be processed without any additional complication. However, the noise in the solution, even when the pseudo-range and phase-rate data are processed together, by far exceeds the magnitude of ocean swell. When processing the data in the static mode, the internal precision of the height is still at the metre level after 10 minutes of data have been processed. Therefore, no kinematic model can strengthen the solution so that the internal precision would become smaller than the expected movement, and consequently, no monitoring of the motion of the rig could be carried out. Perhaps a stronger satellite geometry would solve this problem. The comparison of the static solution with the position from the acoustic transponders is presented in Table 7.6. The convergence of the solution compares with the results shown in Table 7.5 for the onshore baseline. Again, DYNAMO results are better than those from the static sessions processed with PoPS, although the differential positioning solution for the drill stem seems to be systematically biased.

Table 7.6: Errors in the Positioning of the Drill Stem (in m)

	East	North
after 1 min	-0.96 ± 0.84	+5.62 ± 0.64
after 3 min	+2.18 ± 0.48	+6.49 ± 0.37
after 5 min	+2.90 ± 0.37	+7.70 ± 0.29
after 10 min	+2.57 ± 0.27	+7.33 ± 0.20
after 20 min	+4.06 ± 0.19	+8.22 ± 0.14

Although the differential pseudo-range positioning capability of the WM101 was demonstrated, the announcement of dramatic improvements in the design of the new instrument version -- in particular more channels, each dedicated to one satellite -- strongly limit the significance of additional investigations with the present WM101 model.

7.3. OFFSHORE AZIMUTH EXPERIMENT

7.3.1. Campaign Design and Data Collection

For the measurement of subsurface ocean currents, the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Division of Oceanography uses an Acoustic Doppler Current Profiler (ADCP) system. The direction of the currents are measured with respect to the ship. To compute these currents in an absolute system of reference, for example to relate them to grid north on a map, the Azimuth of the ship (with respect to the adopted system of reference) must be known to high precision. A requirement of 0.1 degree accuracy was quoted. A gyrocompass is presently used for this task, but the drift affecting this type of instrument makes the results unreliable, unless they are frequently calibrated. Therefore, in order to test the Azimuth determination capability of GPS, an experiment was organised on a 3 day test cruise of the CSIRO's oceanography research vessel R/V FRANKLIN along the southeastern coast of Tasmania, see Figure 7.8.

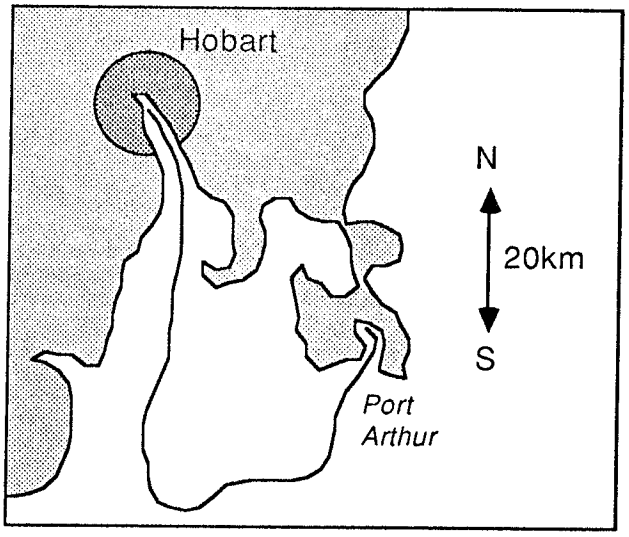


Figure 7.8: FRANKLIN Cruise

In essence, this is a problem of attitude determination and therefore at least two GPS antennae are required onboard the ship. To obtain additional redundancy in the adjustment, any geometrical relation between the position parameters associated with each of the onboard antennae should be used. Such constraints on the parameters can be processed according to the

principle of "decomposition in least squares", see Section 2.4. The optimality of the solution -- in the least squares sense -- is respected. In this experiment, the invariance of the distance between the onboard GPS antennae is enforced as a constraint on the estimated parameters. Although the Azimuth between the antennae need not be equal to the Heading of the ship, see Figure 4.7, the difference is almost constant over a short time span, and can therefore be assigned a kinematic model. Finally, the pitch between the antennae is bounded, limited by the ocean swell. The mathematical expression for these 3 types of constraints are presented in Section 4.5. A strengthening of the solution through processing of phase-rate data was also attempted. Finally smoothing would allow for a further improvement of the results, though at the post processing stage.

Data collection took place from 31 October to 1st November 1988. Two WM101 GPS receivers (and antennae) were installed on the FRANKLIN. At the time of the experiment, some of the peculiarities of the WM101 "raw" data were not known (see discussion in Section 7.2). Ideally, the antennae should be situated as far apart as possible and at as low a height above the sea surface as sky obstructions from the ship's superstructure would permit. This ideal disposition would have required a special mast to be erected at the stern of the vessel. This was not possible, and instead one antenna was placed on the ship's funnel and the other on an existing mast at the bow, which required an extension. This was necessary as both antennae must be at the same height in order to reduce the effect of the roll of the ship on the Azimuth between the antennae. The relative position of the antennae was measured by terrestrial means whilst the FRANKLIN was docked in Hobart: the horizontal distance between antennae was 29.45 metres and the height difference (Funnel - Bow) was 1.10 metres. The height above sea level of the antenna on the funnel was approximately 21 metres.

A third antenna was set up on the roof of a building at the University of Tasmania in Hobart in order to improve the accuracy of the vessel's position using the principles of differential GPS operation (see Section 4.4).

On the first day, due to a break in the connection of an antenna cable, data could only be collected by the antenna on the funnel. On the second day, malfunctions affected the cassette drive of one GPS receiver. This problem may have been caused by the high rate (2 seconds) at which data was

recorded, and for which the receiver was not originally designed. The data finally available consisted of:

- 5 hours of differential data Hobart - Funnel
- 1 hour of attitude data Funnel - Bow (+ Hobart)

The differential data provides a precise track, against which the results obtained using only the data collected on board the ship could be compared. This check was necessary, because the CSIRO does not intend to use a receiver onshore in future for such Azimuth determination operations. Pitch, roll, log and compass data were recorded throughout the GPS session on the 1st November.

7.3.2. Data Preparation

Different types of data were recorded:

- 60 second "compressed" data for the static onshore site in Hobart. The measurements result from a polynomial fit of the 2 second data, and are interpolated for each minute exactly. An investigation of the variations in the measured pseudo-ranges and phase-rates indicated that precisions of 1m and 5mm/sec, respectively, can be assumed for this type of data.

- 2 second "raw" data for both receivers onboard the ship. The precision of the pseudo-range measurements is at the 3m (rms) level, partly because of the motion of the ship during the 2 second measurement interval. Although the phase-rate is measured with a precision of say 5cm/sec in the static case, pitch and roll effects render the measurements noisy at the 50cm/sec (rms) level.

As already noted in the Rig Experiment (Section 7.2), the raw data is output at irregular epochs. In this experiment, the state vector is rather large (up to 30 parameters) and the computation of an adjustment step (Section 3.1) for each incoming measurement becomes very burdensome. Therefore, the data preparation module of DYNAMO was enhanced in order to concentrate each group of raw data measurements at a predefined epoch. This is achieved by using the phase-rate data to correct the pseudo-ranges. Although it is straightforward to compress the data at 2 second intervals (see Figure 7.9), the

processing of such data quickly indicated that this was not sufficient, because the two alternating groups of "raw" measurements are not necessarily identical for all tracking receivers. In other words, the two antennae on board may not observe the same satellites simultaneously, and this causes systematic fluctuations of their relative position. In addition, when 4 or 5 satellites are tracked, some satellites are continuously tracked whilst others share a hardware channel, and the satellites which are tracked continuously are not necessarily the same for all receivers. Hence, the observation time of each satellite is strictly identical for all receivers only when 1, 2, 3 or 6 satellites are tracked. Therefore data compression at a 4 second rate was adopted (see Figure 7.9), and only the observation window when 6 satellites were visible was considered.

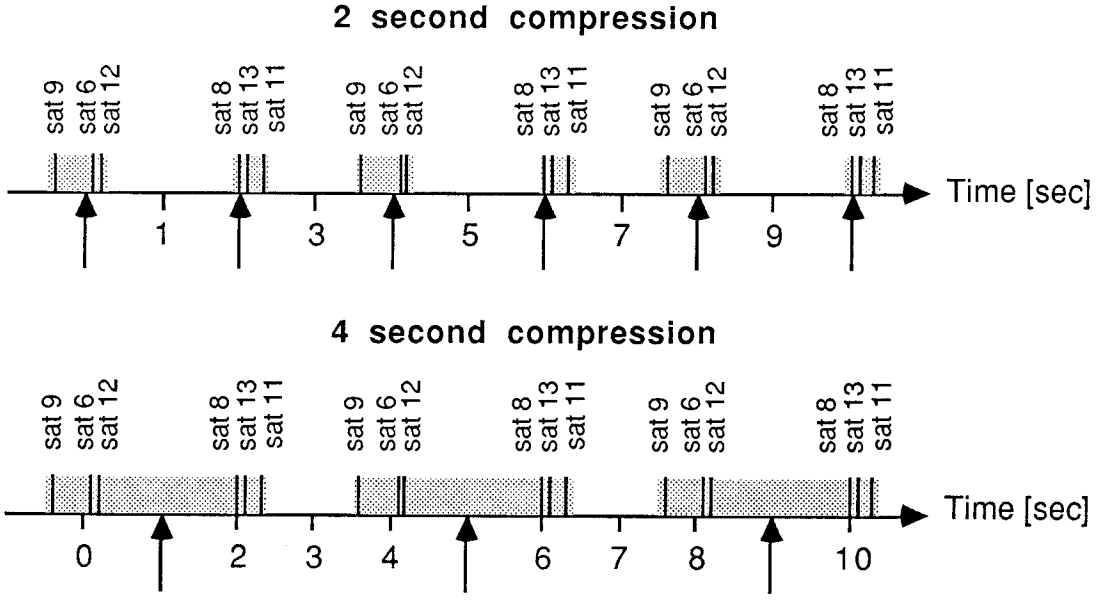


Figure 7.9: Compression of the WM101 "Raw" Data

Data compression at a 4 second rate implies that some measurement time tags must be shifted by more than one second. As the maximum range acceleration is 0.14m/sec^2 (GRANT, 1988), the phase-rate should be corrected by up to 20cm/sec and the pseudo-range by up to 15cm. Such corrections could be computed using orbital information. Alternatively, measurements could be modelled using a polynomial of degree 2 or more, however such an interpolation scheme would correlate the measurements at the successive predefined epochs. In fact, corrections due to the range acceleration term can be neglected in this case because the magnitude of the noise in the measurements clearly exceeds the effect of acceleration terms. Indeed, the

major problem is the correction of the pseudo-range with a velocity term that is relatively imprecise and may cause errors in excess of one metre. Nevertheless, this method is preferred here because of its simplicity, and because a sophisticated correction scheme would be unnecessary for other GPS receivers.

In order to facilitate the comparison of many processing options, only a 5 minute interval of data is investigated. For the filter / smoother to operate virtually in the steady state during that period, 30 seconds of raw data before and after the interval of interest were included in the analyses. For the same reason, 5 minutes of data from the reference station before and after the period to be investigated were added. The data actually processed is indicated in Figure 7.10. The FRANKLIN had left Port Arthur a half hour earlier and was cruising at a speed of 6m/sec (~12knots).

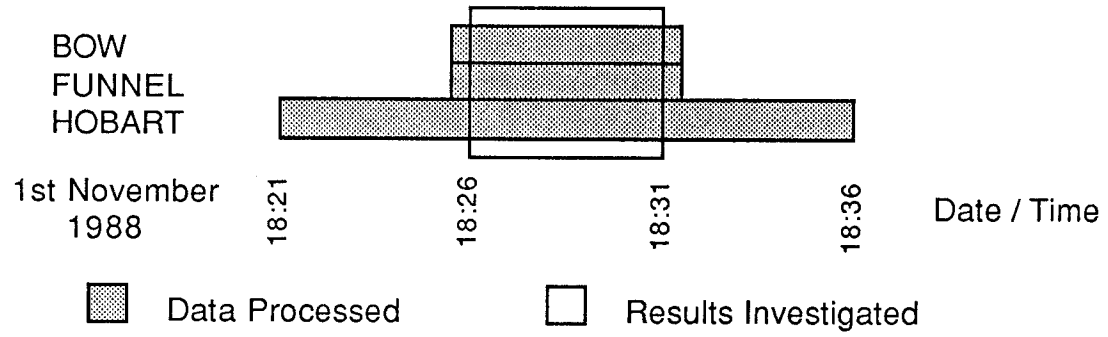


Figure 7.10: Data Processed and Results Investigated

In summary, the characteristics of the data input into the DYNAMO program are:

- the 4 second rate is rather slow, and the short term motion of the vessel cannot be easily modelled.
- a 6 satellite constellation with a PDOP of 3.1 is strong, however, because of the switching channels in the WM101 receiver, each satellite is only tracked during half the total tracking time.
- the corrections of the measurements to predefined epochs introduces some errors in the data. However, these errors are either of random nature or are largely masked by the other random errors, such as the measurement noise and short term movements of the antennae.

These facts should be borne in mind when the results obtained from this experiment are used to predict the accuracy that may be achieved with other, better suited GPS instrumentation.

7.3.3. Differential Corrections

a) Computation of Differential Corrections.

The inclusion of a shore-based reference station permits differential corrections to be computed. In DYNAMO, the preferred solution is to correlate the observations of all the receivers from the same satellite by including a common bias -- the satellite range bias -- in the functional model of the measurements (Section 4.3). Because of the slow data recording rate of 60 seconds at the reference station, it is very important to assign realistic kinematic models to the satellite range biases. Otherwise during the filtering process, predicted values of these biases and/or their predicted precision will either render them ineffective (if the assumed system noise is too large) or cause systematic errors in the position parameters (if the predicted values are inaccurate and the assumed system noise is too small).

The satellite range biases are <u>not</u> designed to account for the modelling of the range between the satellite and the receiver, <u>but only</u> the imperfection of that modelling. In other words, the orbit model provided by the Broadcast Ephemeris -- and used to compute the position of the satellites in their orbit at the measurement time -- already accounts for range-rate and acceleration terms. In a similar fashion, the satellite clocks are corrected using the clock error coefficients contained in the Navigation Message. Whilst the value of a satellite range bias may be quite large, say of the order of a few tens of metres, its variation is limited by the fluctuations of the satellite clock, atmospheric delay and orbital errors. Therefore, one may expect the satellite range biases to be "well behaved".

Although the position of the reference station is held fixed, its clock is not because a satellite clock is preferred for the definition of the time datum, as explained in Section 4.4. However, the modelling of the receiver clock at the reference station is not critical, because it is not used to model the measurements involving other receivers. Hence, as long as enough satellites can be observed, it is not worth attempting a precise modelling. As the receiver clock rate is large (over 50m/sec) it must be estimated and

propagated, hence a model of degree 1 and order 2 is chosen. Over 60 seconds, the driving noise of 1m/sec^2 assumed in previous experiments yields system noise variance terms of 60m^2/sec^2 for the clock drift and 72000m^2 for the offset (see Appendix F). In other words, the clock is not assumed to be stable. Because no reasonable model can yield accurate predictions over such a time span anyway, this rather large process noise is assumed again, see Table 7.7.

In order to select appropriate kinematic models for the satellite range biases, a filtering process involving only the reference station was designed. When 6 satellites are observed, the parameters in the state vector include 5 satellite range biases (one is held fixed) and 1 receiver clock, together with some time derivatives (HOB and SRB parameters only, see Table 7.7). For each satellite range bias, the questions are:

- which derivative has a non negligible value ?
- which derivative can be assumed to be white noise with mean zero ?

The following alternatives were tested, and the comments that follow will further illustrate the modelling "dilemna" mentioned in Section 4.2. For each investigated model, the ratio mean error *a posteriori* / mean error *a priori* was monitored and the intensity of the driving noise selected so that the ratio would be close to unity.

- degree 2 and order 3: In this case, the acceleration is a random walk. Unlike a white noise the estimated value of the acceleration is representative of the "motion", because it is correlated with the value of the acceleration at past or subsequent epochs. Hence, it makes sense to propagate the acceleration into velocity and position. Such a model is useful to ascertain the magnitude of possible acceleration terms. For all satellites, the acceleration term was found to be negligible.

- degree 2 and order 2: the acceleration is estimated as a nuisance parameter. However, the estimated value is not representative of the motion, because the acceleration at any past or subsequent epochs is more likely to be equal to the expectation of the white noise, which is zero by assumption. In this case, the estimated acceleration is not used to predict the next position. However, the computation of the system noise is based on the zero mean

white noise assumption for the acceleration, that is, the assumed uncertainty on the acceleration is propagated into the position. This model is useful to test the acceleration term, because deficiencies in the modelling will show up as systematic errors in the estimated acceleration. For example, if an entire time series of estimated accelerations have positive sign, the zero mean white noise assumption is not appropriate. None of the satellite range biases were found to exhibit a systematic trend.

- degree 1 and order 2: the acceleration is not estimated, and of course not propagated. The reduction in the size of the state vector is justified when the acceleration is effectively a zero mean white noise with a limited variance. However if this assumption is not met, some or all of the estimated parameters may be biased. This is why time series of accelerations were first estimated and their behaviour investigated. The magnitude of the velocity terms is also rather limited, as it never exceeds 0.01m/sec, however systematic trends are obvious, and hence it is preferable to estimate the velocity (~ satellite clock drift) and propagate it into position (~ satellite clock offset). This model was finally selected, and a standard deviation of $0.003 m/sec^2$ was found appropriate for all the satellite range biases.

- A further reduction to degree 0, that is, estimating only the position, would introduce systematic errors in the predicted pseudo-ranges. Note that in cases where computer time and storage should be saved or if the ultimate accuracy is not required, the maximum error of 0.6m induced in the modelled pseudo-range through ignoring velocity and acceleration terms may be tolerated.

b) Effect of Differential Corrections.

If the positions of the onboard antennae are estimated without applying differential corrections shore-to-ship to the measurements, that is, the onshore reference station and the satellite range biases are ignored, it is implicitly assumed that measurements onshore and onboard the ship are uncorrelated. This of course is not valid and the adjustment produces large residuals whenever more than 4 satellites are observed. When 6 satellites are observed, the unfiltered solution yields an average ratio "mean error *a posteriori* / mean error *a priori*" exceeding 5. That is, the average residual of a pseudo-range exceeds 15m. Time series of pseudo-range residuals exhibit obvious correlations between sites and between epochs. In other words, the expectancy of the residuals is far from zero ! Whilst a filter is efficient to reduce

the magnitude of random variations, it does not markedly reduce the size of the residuals and the ratio *a posteriori / a priori* remains over 4.

However, the relative position of the two antennae onboard is not affected by differential corrections, as long as the same satellites are observed. If this is not the case, the relative position may be affected by up to several tens of metres. In an unfiltered solution, the relative position is biased only for these epochs for which different satellites are tracked at the two sites. With a filter, the bias in the relative position when different satellites are observed is reduced, but the results for subsequent epochs are also affected and it takes some time until an accurate solution is produced again. However, the filter may include a rejection criteria to prevent such problems. Nevertheless, the simplest method is to reject measurements to any satellite that is not simultaneously recorded by both onboard receivers. For a shipboard environment where the risk of losing lock at either site is relatively high, this rejection strategy may lead to a significant loss of information. During the 5 minute period under consideration, the tracking to one or more satellites was interrupted twice for a few seconds.

The inclusion of the reference station permits a much better modelling of the measurements, and virtually eliminates the jumps in the solution caused by a change in the set of satellites tracked. Losses of lock are less likely to occur at the reference site, and indeed the 60 second data compression with the WM101 receiver onshore implies that only long duration of losses of lock would result in the lack of measurements. Indeed, because of this internal processing capability, the WM101 receiver is a reliable and convenient reference station. Whilst a reference station makes any real time processing much more difficult, because of the required data link between receivers on the ship and the shore, this is no problem if only post processing is required. As the additional amount of data is relatively small, the only drawback is an increase in computer time due to a larger state vector. The improved modelling of the measurements means that the expectancy of the residuals is actually zero, and that their variations are random. This makes the monitoring of the filter / smoother much easier.

7.3.4. State Vector and Kinematic Models

Table 7.7 shows the kinematic models assigned to all parameters for the most sophisticated filter / smoother formulation. All the processing options tested are merely subsets of this complete configuration. For example, all SRB parameters as well as Time HOB are removed (or their degree set to -1) when the differential corrections are ignored. The parameters for the stochastic constraints Azimuth - Heading and Pitch are treated similarly when the constraints are not applied in some solutions.

Table 7.7: State Vector Content and Kinematic Models

Parameter		Degree	Order	Process Noise	Correlation Time
East	HOB	-1	0	0	
North	HOB	-1	0	0	
Vert.	HOB	-1	0	0	
Time	HOB	1	2	1m/sec^2	∞
East	FUN	1	2	0.1m/sec^2	∞
North	FUN	1	2	0.1m/sec^2	∞
Vert.	FUN	1	1	1m/sec	8sec
Time	FUN	1	2	1m/sec^2	∞
East	BOW	1	2	0.1m/sec^2	∞
North	BOW	1	2	0.1m/sec^2	∞
Vert.	BOW	1	1	1m/sec	8sec
Time	BOW	1	2	1m/sec^2	∞
SRB	6	-1	0	0	
SRB	8	1	2	0.003m/sec^2	∞
SRB	9	1	2	0.003m/sec^2	∞
SRB	11	1	2	0.003m/sec^2	∞
SRB	12	1	2	0.003m/sec^2	∞
SRB	13	1	2	0.003m/sec^2	∞
Dist.	BOW-FUN	-1	0	0	
Az-H	BOW-FUN	0	1	0.01deg/sec	∞
Pitch	BOW-FUN	0	1	1deg/sec	8sec

The content of the state vector can be deduced from the degree of each component listed in Table 7.7, see Section 4.2. The total number of states is obtained as:

$$\text{Dimension} = \sum (\text{degree}+1) \qquad (7\text{-}1)$$

and hence never exceeds 30.

7.3.5. Position

Table 7.8 presents the results obtained using a number of processing options. The value of PDOP for the period investigated is 3.1. It is defined as:

$$PDOP = \sqrt{\sigma_E^2 + \sigma_N^2 + \sigma_V^2} \; / \; \sigma_{PR} \qquad (7\text{-}2)$$

for an unfiltered single point position where:

σ_E, σ_N and σ_V	are the mean errors in the East, North and Vertical directions,
σ_{PR}	is the mean error in a pseudo-range measurement.

The same value for PDOP is obtained when several sites are processed together, as long as the correlation between measurements at different sites is not modelled, because this is like computing several independent single point positions. Using (7-2), it is possible to define a "PDOP" value for any processing method. By doing so, the improvement in the results due to a refinement of the estimation procedure can be expressed by a single number. This is similar to substituting PDOP for the BDOP3 indicator in the case of a static tracking session, see Figure 1.4. All the options shown in Table 7.8 include the onshore reference station. For the Differential Navigation Solution (1), the slight increase in "PDOP" with respect to the standard value (that is, 3.2 instead of 3.1) is due to the addition of satellite range biases in the modelled pseudo-ranges. Although these parameters are stochastically modelled with a very small system noise, they are not completely deterministic and therefore diminish some of the strength in the solution. This is however a low price for the increase in reliability.

The average ratio "mean error *a posteriori* / mean error *a priori* " given in Table 7.8 indicates that the functional and stochastic models used are in good agreement. The ratio gives an overall indication for the adjustment steps, not only for position parameters. Indeed, all precisions shown may be multiplied by 0.85 to obtain the rms value, but more importantly the invariance of the ratio

with respect to the processing options indicates that all the refinements progressively introduced in the estimation procedure -- phase-rate, filtering, smoothing and constraints -- are adequately modelled.

Table 7.8: Precision of Absolute Position Parameters FUNNEL

Solution		E [m]	N [m]	V [m]	"PDOP"	Ratio
(1) NAV	PR only	5.43	2.37	7.46	3.2	0.80
(2) FIL	PR only	2.51	1.77	1.90	1.2	0.84
(3) SMO	PR only	1.56	1.10	1.81	0.9	0.85
(4) FIL	PR +PH	2.30	1.52	1.86	1.1	0.85
(5) SMO	PR +PH	1.50	1.02	1.77	0.8	0.85
(6) F+C	PR +PH	1.58	1.11	1.40	0.8	0.86
(7) S+C	PR +PH	1.11	0.78	1.33	0.6	0.86

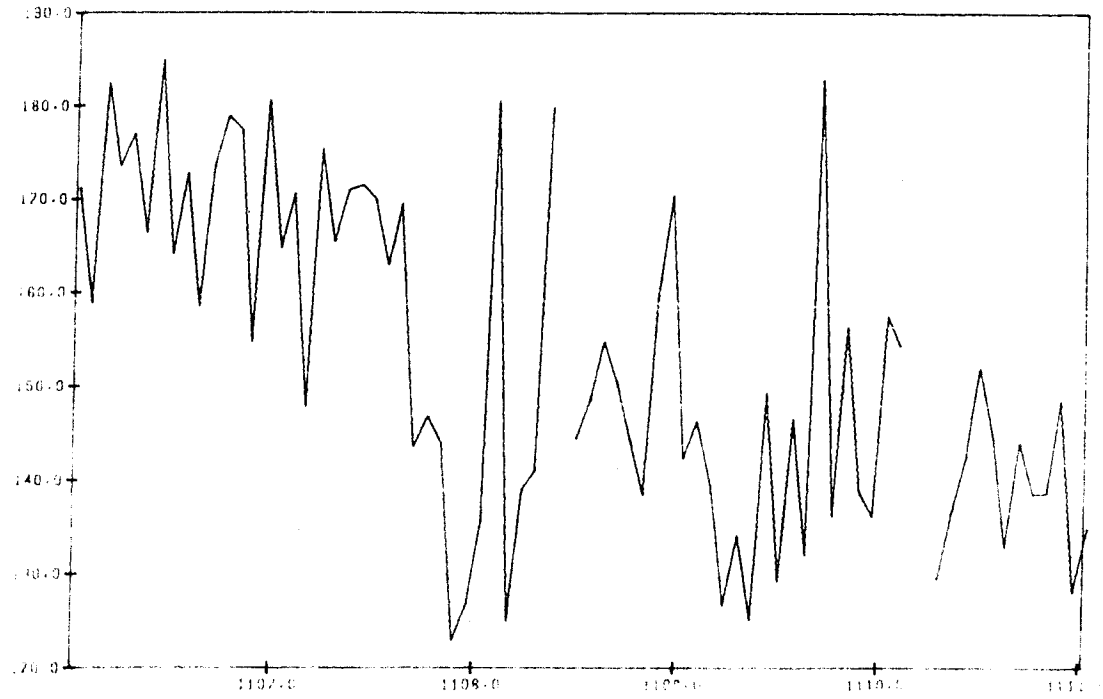
For the processing option combining smoothing and constraints (S+C), the R-T-S formulation was used for the smoother (Section 3.5), rather than the new algorithm (Section 3.6), because the latter formulation implicitly assumes that the improvement in the subsequent estimates is caused exclusively by measurements.

7.3.6. Azimuth

Table 7.9 gives the estimated precision for the baseline joining the onboard antennae (Distance & Azimuth) and the absolute velocity of the vessel (Speed & Heading) for various processing options. The precision of the estimated Azimuth may change significantly between epochs, especially in the unconstrained solutions, because it depends on the estimated magnitude of the distance between the antennae. The same holds for the Heading and speed of the ship.

Table 7.9: Precision of Relative Position BOW-FUNNEL and Absolute Velocity FUNNEL.

Solution		Distance 1σ in m	Azimuth 1σ in deg	Speed 1σ in m/sec	Heading 1σ in deg
(1) NAV	PR only	5.11	13.25	-.-	-.-
(2) FIL	PR only	2.95	6.54	0.30	2.86
(3) SMO	PR only	1.77	3.97	0.16	1.57
(4) FIL	PR +PH	2.59	5.84	0.24	2.39
(5) SMO	PR +PH	1.67	3.77	0.15	1.46
(6) F+C	PR +PH	0.05	1.94	0.19	1.51
(7) S+C	PR +PH	0.05	1.12	0.12	1.08



Navigation Solution

Figure 7.11: Estimated Azimuth [deg] versus Time [min in day], Pseudo-Range Data Only

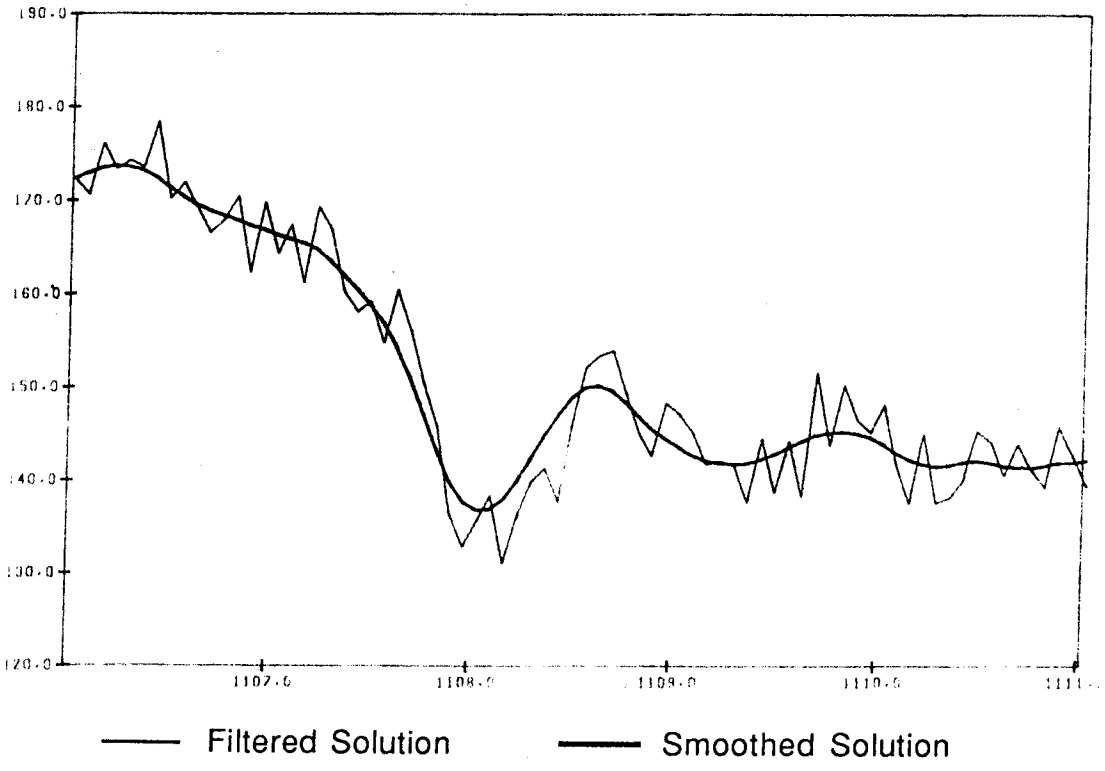


Figure 7.12: Estimated Azimuth [deg] versus Time [min in day], Pseudo-Range Data Only

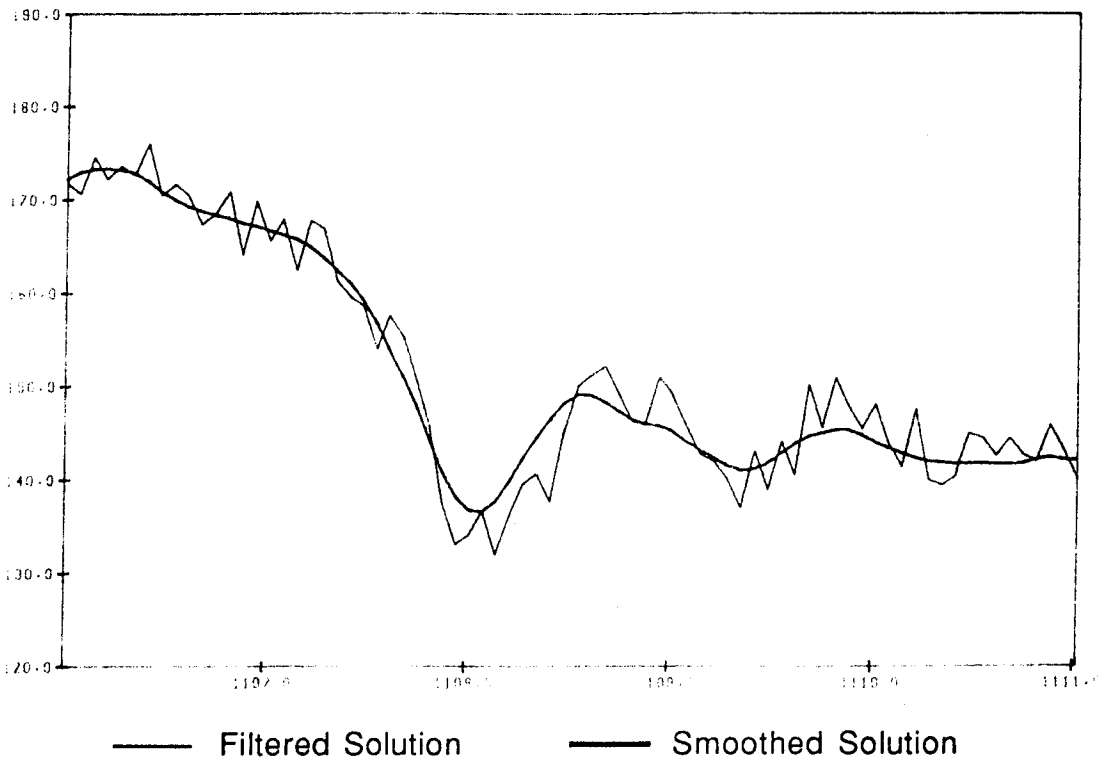


Figure 7.13: Estimated Azimuth [deg] versus Time [min in day], Pseudo-Range and Phase-Rate Data

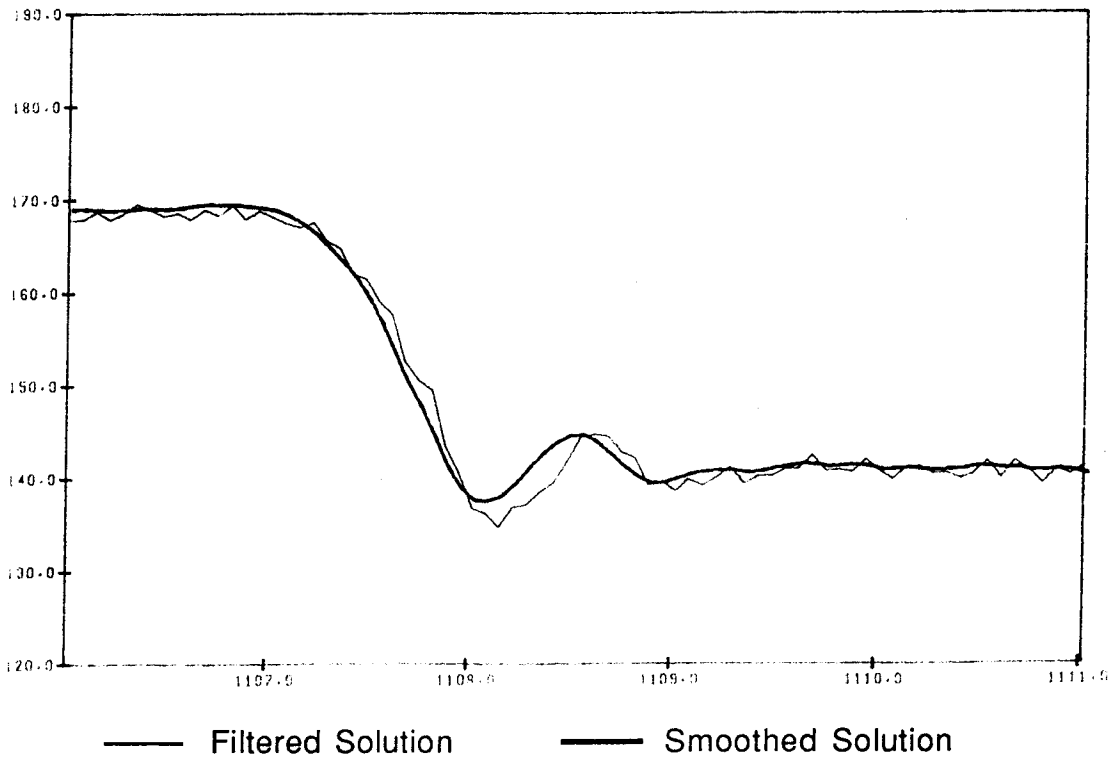
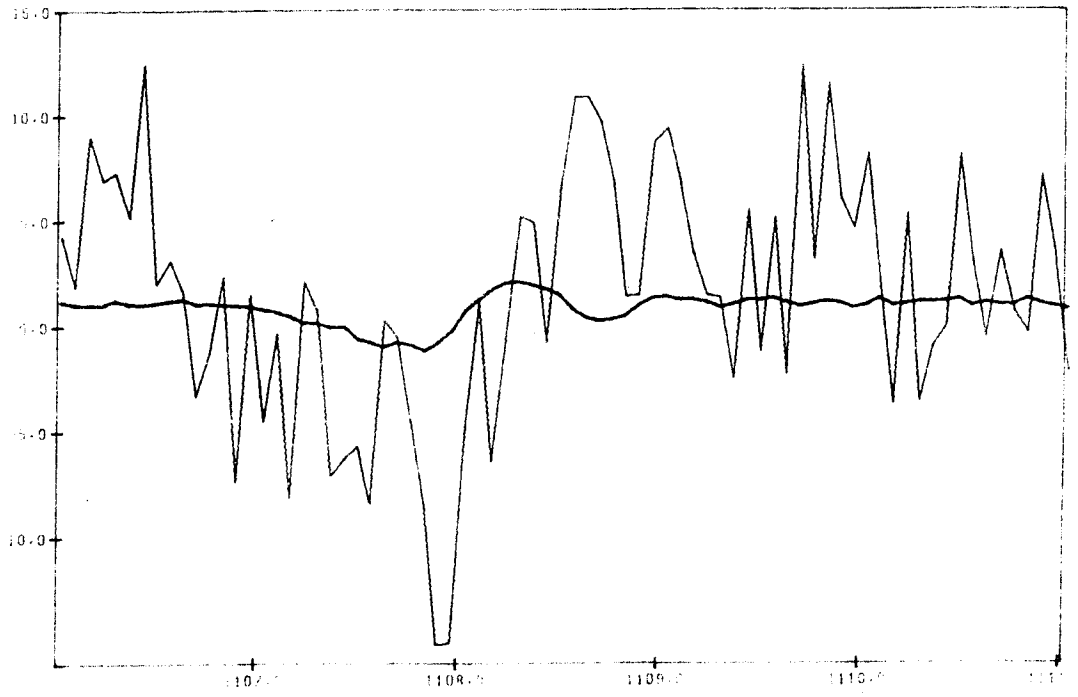


Figure 7.14: Estimated Azimuth [deg] versus Time [min in day], Pseudo-Range and Phase-Rate Data + Constraints



Filtered Solution, pseudo-range data only
Smoothed Solution, pseudo-range and phase-rate data + constraints

Figure 7.15: Estimated Difference Azimuth minus Heading [deg] versus Time [min in day]

The application of a constraint to the difference between the Azimuth and the Heading makes sense because the determination of the Heading is more precise than that of the Azimuth, see solutions (2) to (5). Although the phase-rate data provides direct velocity information, its contribution to the determination of the Heading is very small, compare (2) and (4), or (3) and (5). By filtering and smoothing, the kinematic model assigned to the motion of the ship (that is, to both antennae) provides the main contribution, compare (1), (2) and (3).

Figures 7.11 to 7.14 show the evolution of the estimated Azimuth over the 5 minute data interval selected. There is an improvement in the behaviour of the solutions, as they become more sophisticated. This effect is strong when a kinematic model is introduced (from Figure 7.11 to Figure 7.12) or when constraints are applied (from Figure 7.13 to Figure 7.14). The input of phase-rate data has a limited effect (from Figure 7.12 to Figure 7.13). Figure 7.15 illustrates the variations in the difference Azimuth - Heading during the same time interval for the solutions (1) and (7), that is filtered using pseudo-range data only and smoothed using also phase-rate data and constraints. The most sophisticated solution yields results that appear much more satisfactory. In

particular, the oscillation of the difference Azimuth - Heading around zero during the large change in Azimuth (from about 170 to about 140 degrees) is likely to be real. However, it is difficult to ascertain whether such results are "over-smoothed".

7.3.7. Design of a Specific Estimation Procedure

Although the DYNAMO program can process the data collected, it is far from optimal with respect to computational efficiency. DYNAMO is convenient for feasibility studies because various processing schemes can be investigated easily. However, for an implementation of the filter / smoother in an operational environment, a specific design is more efficient. For this Azimuth determination application, significant simplifications can be made:

- Shore-to-ship corrections are not necessary as they do not affect the relative position of the onboard antennae. However, only data from satellites observed simultaneously at both sites should be processed. Simplification 1: **No onshore reference station is required.**

- Because of the small separation of the onboard antennae, common measurement biases caused by satellite clocks, orbit modelling or atmospheric delays can be completely eliminated by differencing the data observed at both antennae. When computing the differences, measurements involving a satellite that is not simultaneously observed by the other antenna can be eliminated. Simplification 2: **Differencing the data between receivers reduces the volume of input and ensures consistency.**

- The position of the ship need not be estimated by the filter. The coordinates computed internally by the onboard GPS receivers are good enough. Simplification 3: **Only the difference in position between the onboard antennae need be estimated.**

- The difference in position between the onboard antennae need not be expressed in a global 3-Dimensional system of coordinates. Rather, it can be expressed by the slope distance, Azimuth and pitch (or elevation angle) between the antennae. In addition, the slope distance can be measured accurately prior to the processing of GPS data, hence it need

not be estimated, nor input as a constraint. Simplification 4: **In the functional model for the differenced pseudo-ranges, only the Azimuth, the pitch and the offset between receiver clocks need be estimated.**

- In DYNAMO, the measurement model for a gyrocompass would be expressed with respect to 3-Dimensional position parameters. For a specific estimation procedure, this is not necessary. Simplification 5: **Because the Azimuth is a state parameter, the measurement model for gyrocompass data is straightforward.**

- In DYNAMO the stochastic constraint on the pitch requires extra parameter(s) in the state vector. This is not necessary for a specific estimation procedure. Simplification 6: **Because the pitch is a state parameter, bounds for its variations can be expressed by a kinematic model.**

These simplifications result in a simple estimation procedure that requires only limited computer power. However, the relation between Azimuth and Heading has been completely neglected. If the Heading can be determined precisely and independently, the state vector can be augmented and a stochastic constraint imposed, as in DYNAMO. For example, the average of the phase-rates measured at both antennae can provide such a determination of the Heading. Note that the derivatives of Easting and Northing position parameters in the model of the phase-rate measurements (Section 4.3) can be replaced by Heading and Speed. In this case, the expression for the constraint is also simple, because both Azimuth and Heading are in the state vector.

However such a convenient estimation procedure is not applicable if the relation between the velocity and the change in position has to be considered, because absolute positions are not estimated. In other words, the simplified procedure cannot include a kinematic model of the ship's motion. Such a model has proven to be valuable in this experiment. Indeed, it is only possible to neglect such kinematic modelling if the absolute velocity vector can be measured precisely, because an improvement in the determination of the Heading through assumptions regarding the (predictability of the) ship's motion is no longer required in this case.

7.3.8. Future Investigations

From the results of this first trial, the goal of a 0.1 degree precision in the Azimuth of the FRANKLIN cannot be achieved with GPS pseudo-ranges alone. Even with an optimal satellite constellation, an increased measurement rate and using receivers with dedicated channels, this is not likely to be possible.

A marked improvement using phase-rate data could only be achieved with a high measurement rate, say 0.1 second, so that variations would not be masked by an averaging process. The precise phase-rate measurements would be integrated into position and provide a good relation between positions estimated at close epochs, say a few seconds apart. In such a set-up, the phase-rate data resembles the accelerations measured with an inertial sensor. In particular, the error in position grows with time, and in the long term the accuracy of the position mainly depends on the pseudo-range data.

Results of attitude determination using GPS carrier phase data have been published (for example NESBØ, 1988 or KRUCZYNSKI et al, 1988). The advantage and drawbacks are basically the same as described in Section 1.5 for the intermittent positioning mode, that is, an increase in precision but a decrease in reliability. At sea, there is an additional problem because it is not possible to regress to the static mode after a loss of lock to allow for a new determination of the integer biases. Such an approach is therefore not thought to be appropriate for the FRANKLIN, at least for the immediate future.

Indeed, the discrete sampling of GPS pseudo-range data makes a kinematic model necessary, and the real problem is the modelling of the short term behaviour of the ship. However, the emphasis on the modelling of the motion can be reduced if inertial sensors are used. For example, a strong improvement could be achieved by recording pitch and roll information for each antenna separately, because the pseudo-range and phase-rate measurements could then be modelled more accurately. Note that the phase-rate data should be corrected using the time derivatives of pitch and roll. However, such effects may not have to be modelled if the elevation of the antennae above sea level is small.

Azimuth determination using a high quality gyrocompass is certainly more precise than using a dual GPS set-up, because this type of sensor reacts much better to short term variations in the motion of the vessel. However the drift affecting such inertial sensors is well known. The advantage of GPS is mainly in the "long term" positioning. That is, unlike using a compass, the attitude determination is "noisy" but not subject to a drift. Therefore, GPS could be used to permanently calibrate the compass, and hence special calibration procedures presently employed would become unnecessary. This approach seems most appropriate for future investigations. In summary, the awkward modelling of the short term motion of the vessel could be replaced by the reliable measurements from the gyrocompass, whilst the long term accuracy of the positioning process would be ensured by GPS. The Kalman filter / smoother remains a suitable technique for parameter estimation and the application of constraints on the parameters is still a valid means of improving the consistency of the results. However, a larger system noise could be assumed for the parameters of the motion, thus ensuring that the inertia effect of the filter is limited.

For the conduct of a similar experiment in the future, some points deserve particular attention:

1) A special mast should be mounted on the ship's stern, thus allowing for a 50 metre baseline and a height of the antennae above sea level of about 15 metres. Special scaffoldings at both the bow and the stern may permit the antennae to be set up outside the hull, thus allowing for a lower elevation and a longer baseline.

2) The recording of data on cassettes must be avoided, for example by logging the data directly into a personal computer using the input/output RS-232 port available on GPS receivers.

3) All cable connections must be checked before leaving the dock, as antennae may not be accessible at sea.

4) It would be preferable to cruise at as high a speed as possible (say 12 knots), as this results in a more precise determination of the ship's Heading.

8. CONCLUSIONS

The design of an appropriate kinematic GPS positioning procedure must take a number of considerations into account, including accuracy, flexibility and real-time operation.

Accuracy is an important consideration for positioning results. Accurate results should possess two essential features:

- They should be reliable, that is, all major blases are absent from the solution. The differential use of GPS receivers provides effective techniques for the removal of most systematic errors.

- They should be precise, that is, short term variations of the positioning solutions are limited or, in other words, the solution is "well behaved". Such a reduction in the "noise" affecting the solution may be effectively achieved by Kalman filtering and smoothing techniques.

Real-time differential operation may place the reliability of the positioning system, and hence the accuracy of the results, at risk because of possible malfunctions of the necessary data link between the receivers. The more data that must be transmitted, the less reliable the system becomes. Therefore the transmission of pseudo-range correction factors is the minimum that is acceptable for real-time applications. Furthermore, the transmission rate can be kept low, less than the measurement rate of the GPS receiver. On the other hand, considering the different algorithms presented, the most precise results are consistently obtained after smoothing, which is impossible in real-time. In any case, with respect to both precision and reliability, the highest accuracy is not possible in real-time.

Flexibility in the choice of the estimation procedure may be essential when new applications are considered. Most GPS instrument manufacturers are incorporating simple differential positioning capabilities into their receivers which require a minimum of operator input. Although such largely automatic operation may be appropriate in most cases, especially for real-time operation, a complete reliance on internal data processing presents several drawbacks:

- Limited capability upgrade -- the simple "black-box" differential positioning systems provided by GPS manufacturers can only be upgraded by them.

- Specific operation -- the problem of "turn-key" systems is that they are often incompatible with other manufacturer's systems. Hopefully with the adoption of a common standard for the differential correction message it will be possible to mix and match systems.

If the user designs his own software, it may be possible to build in enhancements. This permits a better modelling of the changing / prevailing conditions. However, although flexibility relies mainly on external processing, some internal processing may be integrated in a flexible scheme. For example, this was realised in the Offshore Azimuth Experiment (Section 7.3) by using the 60 second compressed GPS data at the static onshore reference station. Such data can be thought of as "prefiltered".

The DYNAMO software package was designed as a research tool. Therefore, accuracy and flexibility have been the main points of concern throughout the software development. Real-time operation was not envisaged, and the data must all be transported to a central facility. This approach was adopted for a number of reasons:

- The system can handle any number of tracking receivers and observed satellites, hence always giving a unique and optimum result.
- Performance "tuning" can be based on the dynamics encountered. In other words DYNAMO could be used for land, marine or air differential positioning with only a small change in the filter "settings".
- Various schemes of differential (pseudo-range) navigation can be tested, as they are merely "degenerate" examples of a filter (that is, no filter !).
- With an appropriate extension, any sensor data can be incorporated, whether GPS (such as carrier phase data) or other (INS, etc.).

Many of the routines are common to the UNSW static (surveying) software package, and permit data from many GPS receiver types to be processed.

The filtered and smoothed results of the TI-4100 onshore experiment at Sydney Airport are encouraging (Section 7.1). There appears to be a two-fold

improvement in the quality of the Differential Kalman Filter approach over the results obtained from the Differential Navigation approach, and even slightly better after smoothing. The analysis of the WM101 Rig Experiment did not show such optimistic results (Section 7.2). However, this is mainly due to an adverse relation between the precision of the measurements and the (small) amplitude of the actual motion as well as to some characteristics of the instruments, which were not designed for kinematic operation.

The application of constraints on some of the position parameters has proven to be very useful in the Offshore Azimuth Experiment. The deterministic constraint on the distance between the onboard antennae and the stochastic constraint on the pitch (elevation angle between onboard antennae) consistently improve the position precisions. The stochastic constraint on the difference between Azimuth and Heading plays a major role in the determination of the Azimuth, as the Heading can generally be measured or modelled more precisely than the Azimuth. The size of the measurement residuals is almost unaffected by the application of these constraints. In other words, the measurements and the constraints do not contradict each other in the adjustment, but contribute to the strength of the solution in different ways.

The effectiveness of kinematic modelling and constraints has been demonstrated for GPS navigation, and such information should be used whenever available and together with measurements in the estimation procedure.

Future developments include:

- The development of data logging software for GPS receivers.
- The testing of the software on further datasets in order to ascertain the limits of capability. In particular, the Offshore Azimuth Experiment should be repeated and planned according to the indications gained during the first trial.
- Incorporating carrier phase processing in order to improve the precision of the results.
- The transfer of the software from the VAX computer to a Personal Computer.

REFERENCES

ABIDIN, H.Z. (1988) "Real-Time Marine Kinematic Differential GPS Positioning and Velocity Determination". The Institute of Navigation, Satellite Division, Student Paper Competition, Colorado Springs.

AXELRAD, P. & PARKINSON, B.W. (1988) "Closed Loop Navigation and Guidance for Gravity Probe B Orbit Insertion". Technical Meeting of the Institute of Navigation, Satellite Division, Colorado Springs.

BIERMAN, G.J. (1973) "Fixed Interval Smoothing with Discrete Measurement". **Int. Journal of Control**, Vol.118, No.1.

BIERMAN, G.J. (1977) "**Factorization Methods for Discrete Sequential Estimation**". Academic Press, New York, 241pp.

BJERHAMMAR, A. (1973) "**Theory of Errors and Generalised Matrix Inverses**". Elsevier Scientific Publishing Company.

BOSSLER, J.D. (1972) "Bayesian Inference in Geodesy". PhD Dissertation, Dept of Geod. Sciences, Ohio State University, 79pp.

BRYSON, A.E. & FRAZIER, M. (1962) "Smoothing for Linear and Nonlinear Dynamic Systems". Technical Report, ASD-TDR-63-119, Wright Patterson Air Force Base, Ohio, 353-364.

BUCY, R.S. & JOSEPH, P.D. (1968) "**Filtering for Stochastic Processes with Application to Guidance**". Wiley, New York, 195pp.

CARLSON, N.A. (1987) "Federated Square Root Filter for Decentralised Parallel Processing". Proceedings of the National Aerospace Electronics Conference, Dayton, Ohio, 1448-1456.

CARLSON, N.A. (1988) "Distributed Filtering". NAVTECH Seminars, Technical Meeting of the Institute of Navigation, Satellite Division, Colorado Springs.

CHEZELLES, N. de (1988) "Le système de radionavigation GPS-NAVSTAR". Association Française de Topographie, **revue XYZ**, No.36, 32-34.

CHISHOLM, G.S. (1987) "Integration of GPS into Hydrographic Survey Operations". Unisurv S-29, School of Surveying, University of New South Wales, 194pp.

CLYNCH, J.R. & HARPER, W. (1987) "Evaluation of a Large Ship GPS System: Direct Computation and Dynamic Differential". **Marine Positioning**, 87-96.

CROSS, P.A. (1983) "Advanced Least Squares Applied to Position-Fixing". Working paper No.6, Dept. of Land Surveying, NE London Polytechnic, 205pp.

DOVE, M.J. (1977) "Kalman Filter Techniques in Marine Integrated Navigation Systems". **Navigation**, Vol.30.

ECKELS, R. (1987) "Surveying with GPS in Australia". Unisurv S-28, University of New South Wales, 216pp.

EVANS, A.G. (1986) "Comparison of GPS Pseudorange and Biased Doppler Range Measurements to Demonstrate Signal Multipath Effects". Fourth International Geodetic Symposium on Satellite Positioning, Austin TX.

GELB, A. ed. (1974) "**Applied Optimal Estimation**". MIT Press, Cambridge, Mass., 374pp.

GRANT, D. B. (1988) "Combination of Terrestrial and GPS Data in Earth Deformation Studies in New Zealand". PhD thesis, University of New South Wales, 357pp.

GRANT, D. B., RIZOS, C. & STOLZ, A. (1989) "Time in GPS". Report of the School of Surveying, University of New South Wales, in prep.

HARVEY, B.R. (1985) "The Combination of VLBI and Ground Data for Geodesy and Geophysics". Unisurv S-27, University of New South Wales, 239pp.

HATCH, R.R. (1982) "The Synergism of GPS Code and Carrier Measurements". Proceedings of the Third International Geodetic Symposium on Satellite Doppler Positioning, Vol.2, 1213-1231.

HWANG, P.Y.C. & BROWN, R.G. (1985) "GPS Geodesy: Experimental Results Using the Kalman Filter Approach". Institute of Electrical and Electronics Engineers EASCON Conference, Washington D.C.

JORGENSEN, P.S. (1980) "Combined Pseudo-Range and Doppler Positioning for the Stationary Navstar User". Position Location and Navigation Symposium, IEEE Publication 80CH1597-4, 450-458.

KALMAN, R. E. (1960) "A New Approach to Linear Filtering and Prediction Problems". **J. Basic Eng**. 82D, 35-45.

KERR, T.H. (1985) "Decentralised Filtering and Redundancy Management / Failure Detection for Multisensor Integrated Navigation Systems". National Meeting of the Institute of Navigation, San Diego CA.

KING, R.W., MASTERS, E.G., RIZOS, C., STOLZ, A. & COLLINS, J. (1987) "**Surveying with GPS**". Dümmler, Bonn, 128pp.

KLEUSBERG, A. (1986) "Kinematic Relative Positioning Using GPSCode and Carrier Beat Phase Observations". **Marine Geodesy**, Vol.10, No.3/4, 257-274.

KRAKIWSKY, E.J. (1975) "A Synthesis of Recent Advances in the Method of Least Squares". Lecture notes No.42, University of New Brunswick, 125pp.

KRUCZYNSKI, L.R., LI, P.C., EVANS, A.G. & HERMANN, B.R. (1988) "Using GPS to Determine Vehicle Attitude". Technical Meeting of the Institute of Navigation, Satellite Division, Colorado Springs, 7pp.

LIEBELT, P.B. (1967) "**An introduction to optimal estimation**". Addison-Wesley, Reading, Mass., 273pp.

LINDLOHR, W. & WELLS, D. (1985) "GPS Design Using Undifferenced Carrier Beat Phase Observations". **Manuscripta Geodaetica**, Vol.10, 255-295.

LOOMIS, P., KREMER, G. & REYNOLDS, J. (1988) "Correction Algorithms for Differential GPS Reference Stations". Technical Meeting of the Institute of Navigation, Satellite Division, Colorado Springs, 7pp.

McBURNEY, P.W. & BROWN, R.G. (1988) "Self-Contained GPS Integrity Monitoring Using a Censored Kalman Filter". Technical Meeting of the Institute of Navigation, Satellite Division, Colorado Springs, 10pp.

MADER, G.L. (1986) "Decimeter Level Aircraft Positioning Using GPS Carrier Phase Measurements". Int. Symp. Geod. Pos., Austin TX, 15pp.

MEIJER, S.A. de (1983) "Navigation Filters". Department of Marine Technology, Delft University of Technology, 112pp.

MERMINOD, B. (1988) "The Resolution of the Cycle Ambiguities". Report to the Swiss Science Foundation, University of New South Wales, 57pp.

MERMINOD, B. & RIZOS, C. (1988) "A Surveyor's Perspective of Kalman Filtering". **Aust. J. Geod. Phot. Surv.**, No.49, 1-38.

NESBØ, I. (1988) "Applications of GPS Determined Attitude for Navigation". Technical Meeting of the Institute of Navigation, Satellite Division, Colorado Springs, 6pp.

PACHELSKI, W. (1980) "On the Decomposition in Least Squares (with Examples of its Application in Satellite Geodesy)". Report 91, German Geodetic Commission (theoretical geodesy), 58pp.

PAPOULIS, A. (1962) "**The Fourier Integral and its Applications**". McGraw-Hill.

RAUCH, H.E., TUNG, F. & STRIEBEL, C.T. (1965) "Maximum Likelihood Estimates of Linear Dynamic Systems". **AAIA Journal**, 3(8), 1445-1450.

REMONDI, B.W. (1985) "Performing Centimeter-Level Surveys in Seconds with GPS Carrier Phase: Initial Results". Technical Memorandum NGS-43, National Oceanic and Atmospheric Administration, Rockville MD, 18pp.

RUSSELL, S.S. & SCHAIBLY, J.H. (1980) "Control Segment and User Performance". Special issue of **Navigation** on GPS, 74-80.

SCHWARZ, K.P. (1983) "Kalman Filtering and Optimal Smoothing" in KRAKIWSKY E.J. (ed.) Papers for the CIS Adjustment and Analysis Seminars, Calgary, 31pp.

SCHWARZ, K.P., CANNON, M.E. & WONG, R.V.C. (1987) "The Use of GPS in Exploration Geophysics - A Comparison of Kinematic Models". Department of Surveying Engineering, University of Calgary, 20pp.

SEEBER, G., HEIMBERG, F., SCHUCHARDT, A. & WÜBBENA, G. (1986) "Developments for the Operational Use of GPS in Kinematic Modes". Activity Report, Institut für Erdmessung, Universität Hannover.

STOLZ, A. & HEIN, G.W. (1988) "Rapid High-Precision Relative GPS Surveying Using the TI-4100". Submitted to **Australian Surveyor**.

SWIFT, E.R. (1985) "NSWC's GPS Orbit/Clock Determination System". First International Symposium on Precise Positioning with GPS, Rockville MD, Volume 1, 51-62.

TALBOT, N.C. (1988) "GPS Surveying Under Selective Availability". Submitted to **Canadian Surveyor**.

TEUNISSEN, P.J.G., SALZMANN, M.A. (1988) "Performance Analysis of Kalman Filters". Faculty of Geodesy, Delft University of Technology, 18pp.

WANLESS, B. & LACHAPELLE, G. (1988) "NOVAS - an Automated Program for the Precise Reduction of GPS Static Carrier Phase Observations". Submitted to **Manuscripta Geodaetica**.

WONG, R.V.C., SCHWARZ, K.P. & CANNON, M.E. (1988) "High-Accuracy Kinematic Positioning by GPS-INS". **Navigation**, Vol.35, No.2, 275-287.

YUNCK, T.P., WU, S.C. & WU, J. (1987) "Precise Near-Earth Navigation with GPS: A Survey of Techniques". The Telecommunications and Data Acquisition Progress Report 42-91, July-September, 29-45.

APPENDIX A: DECOMPOSITION IN LEAST SQUARES

A.1. The Combined Case

Equation (2-9) is written in a form that resembles a condition equation:

$$\mathbf{Bv} = \mathbf{w} + \mathbf{A}\delta\mathbf{x} \qquad \text{with weight matrix: } \mathbf{P} = \mathbf{Q}_{ll}^{-1} \qquad (2\text{-}12)$$

In a first step, no attempt is made to estimate $\mathbf{x}$. At this stage, $\delta\mathbf{x}$ is set to zero and a classical condition adjustment $\mathbf{Bv} = \mathbf{w}$ is performed. Adjusted residuals are obtained using (2-2).

$$\hat{\mathbf{v}} = \mathbf{Q}_{ll}\mathbf{B}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}(\mathbf{w} + \mathbf{A}\delta\mathbf{x}) \qquad (A\text{-}1)$$

In a second step, the expression of the residuals is regarded as an observation equation and written in the usual form (2-5):

$$\mathbf{Q}_{ll}\mathbf{B}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}\mathbf{w} - \hat{\mathbf{v}} = -\mathbf{Q}_{ll}\mathbf{B}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}\mathbf{A}\,\delta\mathbf{x} \qquad (A\text{-}2)$$

$$\mathbf{w}^* - \hat{\mathbf{v}} = \mathbf{A}^*\,\delta\mathbf{x}$$

with weight matrix: $\mathbf{P} = \mathbf{Q}_{ll}^{-1}$. This system of observation equations is solved in the usual manner (2-6) to give:

$$\delta\hat{\mathbf{x}} = (\mathbf{A}^{*T}\mathbf{P}\mathbf{A}^*)^{-1}\mathbf{A}^{*T}\mathbf{P}\,\mathbf{w}^*$$

$\mathbf{A}^*$ and $\mathbf{w}^*$ can be replaced by their original expression in (A-2), and after some simplifications:

$$\delta\hat{\mathbf{x}} = -(\mathbf{A}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}\mathbf{A})^{-1}\mathbf{A}^T(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^T)^{-1}\mathbf{w} \qquad (2\text{-}10)$$

This is the standard formula for the solution of the combined case. Thus the combined case can be solved by applying the condition and the parametric equation methods successively.

A.2. Adjustment with Constrained Parameters

The observation equations and the parameter constraints yield the following system of equations:

$$\overset{\circ}{\mathbf{v}} - \mathbf{v} = \mathbf{A}\delta\mathbf{x} \qquad \text{with weight matrix: } \mathbf{P} = \mathbf{Q}_{ll}^{-1} \qquad (2\text{-}5)$$

$$\mathbf{U}\delta\mathbf{x} = \mathbf{t} \qquad (2\text{-}13)$$

Of course, it is preferable to linearise both relations around the same approximate values $\overset{\circ}{\mathbf{x}}$, so that the increment $\delta\mathbf{x}$ is the same in both equations. This is a problem of minimisation under constraints and the Lagrange method of undetermined multipliers can be applied:

$$\mathbf{v}^T\mathbf{P}\,\mathbf{v} - 2\,\mathbf{k}_1^T(\overset{\circ}{\mathbf{v}} - \mathbf{v} - \mathbf{A}\delta\mathbf{x}) - 2\,\mathbf{k}_2^T(\mathbf{U}\delta\mathbf{x} - \mathbf{t}) \rightarrow \text{minimum} \qquad (A\text{-}3)$$

The derivation with respect to $\mathbf{v}$, $\mathbf{x}$, $\mathbf{k}_1$ and $\mathbf{k}_2$ and the minimum condition lead to:

$$\begin{bmatrix} \mathbf{P} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}^T & -\mathbf{U}^T \\ \mathbf{I} & \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}} \\ \delta\hat{\mathbf{x}} \\ \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \overset{\circ}{\mathbf{v}} \\ \mathbf{t} \end{bmatrix} \qquad (A\text{-}4)$$

premultiplying the third line by $\mathbf{P}$ and subtracting the first line eliminates $\hat{\mathbf{v}}$:

$$\begin{bmatrix} \mathbf{PA} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^T & -\mathbf{U}^T \\ \mathbf{U} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \delta\hat{\mathbf{x}} \\ \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{P}\overset{\circ}{\mathbf{v}} \\ \mathbf{0} \\ \mathbf{t} \end{bmatrix} \qquad (A\text{-}5)$$

premultiplying the first line by $\mathbf{A}^T$ and adding the second line eliminates $\mathbf{k}_1$:

$$\begin{bmatrix} \mathbf{A}^T\mathbf{PA} & -\mathbf{U}^T \\ \mathbf{U} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \delta\hat{\mathbf{x}} \\ \mathbf{k}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{P}\overset{\circ}{\mathbf{v}} \\ \mathbf{t} \end{bmatrix} \qquad (A\text{-}6)$$

premultiplying the first line by $-\mathbf{U}(\mathbf{A}^T\mathbf{PA})^{-1}$ and adding the second line eliminates $\delta\hat{\mathbf{x}}$:

$$\mathbf{U}(\mathbf{A}^T\mathbf{PA})^{-1}\mathbf{U}^T\mathbf{k}_2 = \mathbf{t} - \mathbf{U}(\mathbf{A}^T\mathbf{PA})^{-1}\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}} \qquad (A\text{-}7)$$

Let $\delta\hat{\mathbf{y}}$ be the solution of the parametric system (2-5) alone, neglecting the constraints on on the parameters (2-13). The solution of such observation equations, according to (2-6) and (2-7), yields for $\hat{\mathbf{y}}$ and its covariance:

$$\delta\hat{\mathbf{y}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}} \qquad \text{and} \qquad \mathbf{Q}_{\hat{y}\hat{y}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \tag{A-8}$$

Substituting in (A-7): $\mathbf{U}\mathbf{Q}_{\hat{y}\hat{y}}\mathbf{U}^T\ \mathbf{k}_2 = \mathbf{t} - \mathbf{U}\delta\hat{\mathbf{y}}$

$$\mathbf{k}_2 = (\mathbf{U}\mathbf{Q}_{\hat{y}\hat{y}}\mathbf{U}^T)^{-1}(\mathbf{t} - \mathbf{U}\delta\hat{\mathbf{y}}) \tag{A-9}$$

Replacing in the first line of (A-6):

$$\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} - \mathbf{U}^T\mathbf{k}_2 = \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}$$

$$\mathbf{Q}_{\hat{y}\hat{y}}^{-1}\,\delta\hat{\mathbf{x}} - \mathbf{U}^T(\mathbf{U}\mathbf{Q}_{\hat{y}\hat{y}}\mathbf{U}^T)^{-1}(\mathbf{t} - \mathbf{U}\delta\hat{\mathbf{y}}) = \mathbf{Q}_{\hat{y}\hat{y}}^{-1}\delta\hat{\mathbf{y}}$$

$$(\hat{\mathbf{x}} - \hat{\mathbf{y}}) = \mathbf{Q}_{\hat{y}\hat{y}}\ \mathbf{U}^T(\mathbf{U}\mathbf{Q}_{\hat{y}\hat{y}}\mathbf{U}^T)^{-1}(\mathbf{t} - \mathbf{U}\delta\hat{\mathbf{y}}) \tag{A-10}$$

Comparing with the residuals from a condition adjustment (2-2):

$$\hat{\mathbf{v}} = \mathbf{Q}_{ll}\,\mathbf{B}^T\,(\mathbf{B}\,\mathbf{Q}_{ll}\,\mathbf{B}^T)^{-1}\,\mathbf{w}$$

the expression (A-10) can be considered as the solution of a condition equation with:

	$(\hat{\mathbf{x}} - \hat{\mathbf{y}})$	for the residuals $\hat{\mathbf{v}}$
	$\mathbf{Q}_{\hat{y}\hat{y}}$	for the covariance of the observations $\mathbf{Q}_{ll}$
and	$(\mathbf{t} - \mathbf{U}\delta\hat{\mathbf{y}})$	for the vector of misclosures $\mathbf{w}$.

A complete formal equivalence with the classical condition case is obtained by setting the following correspondance between the terms:

$\hat{\mathbf{x}} = \hat{\mathbf{y}} - \hat{\mathbf{z}}$	and	$\mathbf{U}\mathbf{z} = \mathbf{w}$	with weight matrix: $\mathbf{Q}_{\hat{y}\hat{y}}$	(A-11)
$\hat{l} = l - \hat{\mathbf{v}}$	and	$\mathbf{B}\mathbf{v} = \mathbf{w}$	with weight matrix: $\mathbf{Q}_{ll}$	(classical)

Thus the system can be solved by applying the parametric and the condition equation methods successively.

APPENDIX B: BAYESIAN LEAST SQUARES

Consider the parametric system (2-5):

$$\overset{\circ}{\mathbf{v}} - \mathbf{v} = \mathbf{A}\delta\mathbf{x} \qquad \text{with weight matrices: } \mathbf{P} = \mathbf{Q}_{ll}^{-1} \text{ and } \mathbf{P}_{\mathbf{x}}^{\circ}$$

The minimisation of the quadratic form $\mathbf{v}^T\mathbf{P}\,\mathbf{v} + \delta\mathbf{x}^T\mathbf{P}_{\mathbf{x}}^{\circ}\,\delta\mathbf{x}$ is achieved using the Lagrange method of undetermined multipliers:

$$\mathbf{v}^T\mathbf{P}\,\mathbf{v} + \delta\mathbf{x}^T\mathbf{P}_{\mathbf{x}}^{\circ}\,\delta\mathbf{x} - 2\,\mathbf{k}^T(\overset{\circ}{\mathbf{v}} - \mathbf{v} - \mathbf{A}\delta\mathbf{x}) \rightarrow \text{minimum}$$

The derivation with respect to $\mathbf{v}$, $\mathbf{x}$, $\mathbf{k}$ and the minimum condition lead to:

$$\begin{bmatrix} \mathbf{P} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{P}_{\mathbf{x}}^{\circ} & \mathbf{A}^T \\ \mathbf{I} & \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}} \\ \delta\hat{\mathbf{x}} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \overset{\circ}{\mathbf{v}} \end{bmatrix}$$

premultiplying the third line by $\mathbf{P}$ and subtracting the first line eliminates $\hat{\mathbf{v}}$:

$$\begin{bmatrix} \mathbf{PA} & -\mathbf{I} \\ \mathbf{P}_{\mathbf{x}}^{\circ} & \mathbf{A}^T \end{bmatrix} \begin{bmatrix} \delta\hat{\mathbf{x}} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{P}\overset{\circ}{\mathbf{v}} \\ \mathbf{0} \end{bmatrix}$$

premultiplying the first line by $\mathbf{A}^T$ and adding the second line eliminates $\mathbf{k}$:

$$(\mathbf{A}^T\mathbf{PA} + \mathbf{P}_{\mathbf{x}}^{\circ})\,\delta\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

The expression for the normal equations can be written as:

$$(\mathbf{A}^T\mathbf{PA} + \mathbf{I}\mathbf{P}_{\mathbf{x}}^{\circ}\mathbf{I})\,\delta\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

which is the solution of the system of parametric equations, see (2-6):

$$\begin{bmatrix} \overset{\circ}{\mathbf{v}} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{v} \\ \mathbf{v}_{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{I} \end{bmatrix} \delta\mathbf{x} \qquad \text{with weight matrix } \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\mathbf{x}}^{\circ} \end{bmatrix}$$

Thus, the a priori information on the parameters is equivalent to adding an observation equation in the system:

$$\overset{\circ}{\mathbf{v}} - \mathbf{v} = \mathbf{A}\delta\mathbf{x} \qquad \text{with weight matrix } \mathbf{P}$$

$$\mathbf{0} - \mathbf{v}_{\mathbf{x}} = \delta\mathbf{x} \qquad \text{with weight matrix } \mathbf{P}_{\overset{\circ}{\mathbf{x}}}$$

The effect of the added observation is to maintain the estimated value of $\mathbf{x}$ close to the approximate value $\overset{\circ}{\mathbf{x}}$ used for linearisation, providing a resistance to a change in the parameters induced by the measurements. This can be thought of as an "inertia" effect.

APPENDIX C: LEAST SQUARES FILTER

C.1. The Bayes Form

The covariance matrix of the filtered estimate can be computed by applying the law of propagation of variances, assuming the stochastic independence of $\tilde{\mathbf{x}}$ and l:

$$\delta\hat{\mathbf{x}} = (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}(\mathbf{P}_{\tilde{x}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}) \tag{3-7a}$$

Therefore:

$$\begin{aligned}\mathbf{Q}_{\hat{x}\hat{x}} &= (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{P}_{\tilde{x}}\,\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{P}_{\tilde{x}}\,(\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \\ &\quad + (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\,\mathbf{Q}_{ll}\,\mathbf{P}\mathbf{A}\,(\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \\ &= (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,(\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})\,(\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\end{aligned}$$

Thus: $$\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \tag{3-8}$$

The covariance matrix of the residuals is obtained in a similar fashion. The adjusted value of $\mathbf{x}$ can be substituted in (2-5) and the adjusted residual $\hat{\mathbf{v}}$ becomes:

$$\hat{\mathbf{v}} = \mathring{\mathbf{v}} - \mathbf{A}\delta\hat{\mathbf{x}} \tag{C-1}$$

The law of propagation of variances is applied:

$$\mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}l} - \mathbf{Q}_{l\hat{x}}\mathbf{A}^T + \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T \tag{C-2}$$

From (3-7a), the covariance of $\hat{\mathbf{x}}$ and l is easily derived:

$$\mathbf{Q}_{\hat{x}l} = (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{P}\mathbf{Q}_{ll} = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{P}\mathbf{Q}_{ll} = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T$$

Substituting into (C-2):

$$\begin{aligned}\mathbf{Q}_{\hat{v}\hat{v}} &= \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T + \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T \\ \mathbf{Q}_{\hat{v}\hat{v}} &= \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\end{aligned} \tag{3-9}$$

This expression is exactly the same as for the classical parametric case (2-8), though the definition of $\mathbf{Q}_{\hat{x}\hat{x}}$ is different.

C.2. The Kalman Form

C.2.1. State Vector Estimate

The matrix identity:

$$(\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} = (\mathbf{P}_{\tilde{x}}^{-1} - \mathbf{P}_{\tilde{x}}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{P}_{\tilde{x}}^{-1}\mathbf{A}^T + \mathbf{P}^{-1})^{-1}\mathbf{A}\mathbf{P}_{\tilde{x}}^{-1}) \quad (3\text{-}10)$$

can be applied to the Bayesian expression of the filtered estimate (3-7):

$$\delta\hat{\mathbf{x}} = (\mathbf{P}_{\tilde{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}(\mathbf{P}_{\tilde{x}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}})$$

$$\delta\hat{\mathbf{x}} = (\mathbf{P}_{\tilde{x}}^{-1} - \mathbf{P}_{\tilde{x}}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{P}_{\tilde{x}}^{-1}\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\mathbf{A}\mathbf{P}_{\tilde{x}}^{-1})(\mathbf{P}_{\tilde{x}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}})$$

The inverses of the weight matrices can be replaced by the corresponding covariance matrices:

$$\delta\hat{\mathbf{x}} = (\mathbf{Q}_{\tilde{x}\tilde{x}} - \mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}})(\mathbf{P}_{\tilde{x}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}})$$

$$\delta\hat{\mathbf{x}} = (\mathbf{I} - \mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\mathbf{A})(\delta\tilde{\mathbf{x}} + \mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}) \quad (C\text{-}3)$$

Some simplifications can be made. Let $\mathbf{K}$ be defined as:

$$\mathbf{K} = \mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T + \mathbf{Q}_{ll})^{-1} \quad (3\text{-}12)$$

The estimate of $\delta\mathbf{x}$ becomes:

$$\delta\hat{\mathbf{x}} = (\mathbf{I} - \mathbf{K}\mathbf{A})(\delta\tilde{\mathbf{x}} + \mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\mathbf{P}\,\mathring{\mathbf{v}}) \quad (C\text{-}4)$$

The coefficient of the observation residuals $\mathring{\mathbf{v}}$ is:

$$(\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\mathbf{P} =$$

$$= (\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T - \mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T)\mathbf{P}$$

$$= \mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\,(\,\mathbf{I} - (\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T)\mathbf{P}$$

$$= \mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\,((\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll}) - (\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T)\mathbf{P}$$

$$= \mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\,((\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T)\mathbf{P}$$

$$= \mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\,(\mathbf{A}\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T + \mathbf{Q}_{ll})^{-1}\,\mathbf{Q}_{ll}\,\mathbf{P}$$

$$= \mathbf{K}\,\mathbf{Q}_{ll}\,\mathbf{P} - \mathbf{K}$$

Finally: $$(\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\mathbf{P} = \mathbf{K} \qquad \text{(C-5)}$$

Substituting in (C-4), the filtered estimate of $\delta\mathbf{x}$ can be expressed as:

$$\delta\hat{\mathbf{x}} = (\mathbf{I} - \mathbf{K}\mathbf{A})\,\delta\tilde{\mathbf{x}} + \mathbf{K}\mathring{\mathbf{v}} \qquad \text{(C-6)}$$

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} + \mathbf{K}\,(\mathring{\mathbf{v}} - \mathbf{A}\delta\tilde{\mathbf{x}}\,) \qquad \text{(3-11a)}$$

C.2.2. Covariance Matrix of the Filtered Estimate

The covariance matrix of the filtered estimate can be computed from (C-6). According to the law of propagation of variances and the assumed stochastic independence of $\tilde{\mathbf{x}}$ and l:

$$\begin{aligned}
\mathbf{Q}_{\hat{x}\hat{x}} &= (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}}\,(\mathbf{I} - \mathbf{K}\mathbf{A})^T + \mathbf{K}\,\mathbf{Q}_{ll}\,\mathbf{K}^T \qquad \text{(3-13)}\\
&= (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}} - (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T\mathbf{K}^T + \mathbf{K}\,\mathbf{Q}_{ll}\,\mathbf{K}^T\\
&= (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}} - ((\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}}\,\mathbf{A}^T - \mathbf{K}\,\mathbf{Q}_{ll})\,\mathbf{K}^T\\
\text{(C-5)} \Rightarrow \quad &= (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}} - (\mathbf{K}\,\mathbf{Q}_{ll} - \mathbf{K}\,\mathbf{Q}_{ll})\,\mathbf{K}^T
\end{aligned}$$

Thus: $$\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}} \qquad \text{(3-14)}$$

On the other hand, another relation between the filtered and predicted covariances had been derived for the Bayes filter:

$$\mathbf{Q}_{\hat{x}\hat{x}}^{-1} = \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} + \mathbf{A}^T\mathbf{P}\mathbf{A} \qquad \text{(3-8)}$$

This relation still holds for the Kalman filter and the equivalence of (3-8) and (3-14) can be proven. A convenient method is to make use of the matrix identity (3-10).

C.2.3. Residuals on the Observations

Starting with the parametric expression of the residuals (C-1), and following the usual procedure, the covariance of the vector of residuals is computed:

$$\mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}l} - \mathbf{Q}_{l\hat{x}}\mathbf{A}^T + \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T \qquad \text{(C-2)}$$

From (C-6), the covariance of $\hat{\mathbf{x}}$ and l is easily derived:

$$\mathbf{Q}_{\hat{x}l} = \mathbf{K}\mathbf{Q}_{ll} \qquad \text{(C-7)}$$

Combining (C-5) and (3-14), a useful identity can be derived:

$$\mathbf{K}\,\mathbf{Q}_{ll} = (\mathbf{I} - \mathbf{K}\mathbf{A})\,\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A}^T = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T \qquad \text{(C-8)}$$

Thus: $\mathbf{Q}_{\hat{x}l} = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T$

After substitution in (C-7):

$$\mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T$$

As expected, this result is identical to the expression (3-9) obtained for the Bayes filter.

C.2.4. Another Expression for the Gain Matrix

Considering (C-8), the gain matrix can be computed as:

$$\mathbf{K} = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{P} \qquad \text{(C-9)}$$

Although this expression is simple, it is not applicable when filtering data because $\mathbf{Q}_{\hat{x}\hat{x}}$ is not available. However, this formula can be used to recompute the gain matrix once the data has been filtered, for example during the smoothing phase.

APPENDIX D: LEAST SQUARES FILTER / SMOOTHER

D.1. State Vector Estimates

The minimisation of the quadratic form:

$$\Sigma v^T P v = v_{x_0}{}^T P_{\hat{x}_0} v_{x_0} + v^T P v + w^T P_w w \rightarrow \text{minimum} \qquad (3\text{-}18)$$

is sought and applied to the system of parametric equations (3-19):

$$\begin{aligned} \hat{x}_0 - v_{x_0} &= x_0 && P_{\hat{x}_0} && (D\text{-}1)\\ - w &= \Phi x_0 - x && P_w && (D\text{-}2)\\ \mathring{v} - v &= A\,\delta x && P && (D\text{-}3) \end{aligned}$$

The quadratic form is first expressed as a function of x_0 and x:

$$\begin{aligned} \text{From (D-1):}\quad v_{x_0}{}^T P_{\hat{x}_0} v_{x_0} &= x_0{}^T P_{\hat{x}_0} x_0 - 2\, x_0{}^T P_{\hat{x}_0} \hat{x}_0 + \hat{x}_0{}^T P_{\hat{x}_0} \hat{x}_0\\ \text{(D-2):}\quad w^T P_w w &= x_0{}^T \Phi^T P_w \Phi\, x_0 - 2\, x^T P_w \Phi\, x_0 + x^T P_w x\\ \text{(D-3):}\quad v^T P v &= \delta x^T A^T P A\, \delta x - 2\, \delta x^T A^T P\, \mathring{v} + \mathring{v}^T P\, \mathring{v} \end{aligned}$$

Differentiation with respect to the unknowns and the minimum condition gives:

$$\begin{aligned} \frac{\partial \Sigma v^T P v}{\partial x_0} &= 2\,(P_{\hat{x}_0} \hat{\hat{x}}_0 - P_{\hat{x}_0} \hat{x}_0 + \Phi^T P_w \Phi\, \hat{\hat{x}}_0 - \Phi^T P_w \hat{x}) = 0\\ \frac{\partial \Sigma v^T P v}{\partial x} &= 2\,(A^T P A\, \delta\hat{x} - A^T P\, \mathring{v} - P_w \Phi\, \hat{\hat{x}}_0 + P_w \hat{x}) = 0 \end{aligned}$$

This leads to the system of normal equations, equivalent to (3-20):

$$\begin{aligned} (P_{\hat{x}_0} + \Phi^T P_w \Phi)\, \hat{\hat{x}}_0 - \Phi^T P_w\, \hat{x} &= P_{\hat{x}_0} \hat{x}_0 && (D\text{-}4)\\ - P_w \Phi\, \hat{\hat{x}}_0 + P_w \hat{x} + A^T P A\, \delta\hat{x} &= A^T P\, \mathring{v} && (D\text{-}5) \end{aligned}$$

Premultiplying (D-5) by Φ^T and adding (D-4) gives:

$$P_{\hat{x}_0} \hat{\hat{x}}_0 + \Phi^T A^T P A\, \delta\hat{x} = P_{\hat{x}_0} \hat{x}_0 + \Phi^T A^T P\, \mathring{v} \qquad (D\text{-}6)$$

The smoothed estimate of x_0 can be extracted:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\,(\overset{\circ}{\mathbf{v}} - \mathbf{A}\,\delta\hat{\mathbf{x}}) \qquad (3\text{-}21)$$

Substituting in (D-5):

$$-\mathbf{P}_w\Phi\,\hat{\mathbf{x}}_0 - \mathbf{P}_w\Phi\,\mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\,(\overset{\circ}{\mathbf{v}} - \mathbf{A}\,\delta\hat{\mathbf{x}}) + \mathbf{P}_w\hat{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

$$(\mathbf{P}_w\Phi\,\mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\,\Phi^T\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P}\mathbf{A})\,\delta\hat{\mathbf{x}} + \mathbf{P}_w\,\hat{\mathbf{x}} = \\ = \mathbf{P}_w\Phi\,\hat{\mathbf{x}}_0 + (\mathbf{P}_w\Phi\,\mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P} + \mathbf{A}^T\mathbf{P})\,\overset{\circ}{\mathbf{v}}$$

All terms are premultiplied by $\mathbf{P}_w^{-1}$, yielding:

$$\hat{\mathbf{x}} + (\Phi\,\mathbf{P}_{\hat{\mathbf{x}}_0}\Phi^T + \mathbf{P}_w^{-1})\,\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} = \Phi\,\hat{\mathbf{x}}_0 + (\Phi\,\mathbf{P}_{\hat{\mathbf{x}}_0}\Phi^T + \mathbf{P}_w^{-1})\,\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

According to the prediction equations (3-2) and (3-3), this simplifies to:

$$\hat{\mathbf{x}} + \mathbf{P}_{\tilde{\mathbf{x}}}^{-1}\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} = \tilde{\mathbf{x}} + \mathbf{P}_{\tilde{\mathbf{x}}}^{-1}\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

The state vector $\overset{\circ}{\mathbf{x}}$ used for linearisation is subtracted from both sides:

$$\delta\hat{\mathbf{x}} + \mathbf{P}_{\tilde{\mathbf{x}}}^{-1}\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} = \delta\tilde{\mathbf{x}} + \mathbf{P}_{\tilde{\mathbf{x}}}^{-1}\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

All terms are premultiplied by $\mathbf{P}_{\tilde{\mathbf{x}}}$:

$$(\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})\,\delta\hat{\mathbf{x}} = \mathbf{P}_{\tilde{\mathbf{x}}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$$

$$\delta\hat{\mathbf{x}} = (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,(\mathbf{P}_{\tilde{\mathbf{x}}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}) \qquad (3\text{-}7a)$$

This is the form already obtained with the Basic Least Squares Filter, see Section 3.3.2.

D.2. Covariance Matrix of the Smoothed Estimate

By defining: $\mathbf{H} = \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\Phi^T\mathbf{A}^T\mathbf{Q}_{ll}^{-1}$ (3-24)

the expression (3-21) for the new estimate of $\mathbf{x}_0$ becomes simply:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{H}\,\hat{\mathbf{v}} \tag{3-22}$$

The covariance can be computed in the usual way:

$$\mathbf{Q}_{\hat{\hat{\mathbf{x}}}_0\hat{\hat{\mathbf{x}}}_0} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} + \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{v}}}\mathbf{H}^T + \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{x}}_0} + \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T \tag{D-7}$$

From (C-6) and (3-2), $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}_0$ are related through:

$$\hat{\mathbf{x}} = (\mathbf{I} - \mathbf{K}\mathbf{A})\,\Phi\,\hat{\mathbf{x}}_0 + \mathbf{K}\,\mathring{\mathbf{v}}$$

and the covariance of $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}_0$ can be deduced:

$$\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}_0} = (\mathbf{I} - \mathbf{K}\mathbf{A})\,\Phi\,\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} \tag{D-8}$$

Using $\hat{\mathbf{v}} = \mathring{\mathbf{v}} - \mathbf{A}\delta\hat{\mathbf{x}}$ (C-1)

the covariance of $\hat{\mathbf{v}}$ and $\hat{\mathbf{x}}_0$ is established, assuming the independence of $\hat{\mathbf{x}}_0$ and $\boldsymbol{l}$:

$$\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{x}}_0} = -\,\mathbf{A}\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}_0} \tag{D-9}$$

Inserting (D-8) into (D-9):

$$\begin{aligned}
\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{x}}_0} &= -\,\mathbf{A}(\mathbf{I} - \mathbf{K}\mathbf{A})\,\Phi\,\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} \\
&= -\,\mathbf{A}\Phi\,\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} + \mathbf{A}\mathbf{K}\mathbf{A}\,\Phi\,\mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} \\
&= -\,\mathbf{Q}_{\boldsymbol{l}\boldsymbol{l}}\mathbf{H}^T + \mathbf{A}\mathbf{K}\,\mathbf{Q}_{\boldsymbol{l}\boldsymbol{l}}\,\mathbf{H}^T \\
(\text{C-8}) \Rightarrow \quad &= -\,\mathbf{Q}_{\boldsymbol{l}\boldsymbol{l}}\mathbf{H}^T + \mathbf{A}\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}^T\mathbf{H}^T \\
&= -\,(\mathbf{Q}_{\boldsymbol{l}\boldsymbol{l}} - \mathbf{A}\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}^T)\,\mathbf{H}^T \\
(\text{3-9}) \Rightarrow \quad \mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{x}}_0} &= -\,\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T
\end{aligned} \tag{D-10}$$

Substituting now in (D-7):

$$\begin{aligned}
\mathbf{Q}_{\hat{\hat{\mathbf{x}}}_0\hat{\hat{\mathbf{x}}}_0} &= \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} - \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T - \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T + \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T \\
&= \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} - \mathbf{H}\mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}}\mathbf{H}^T
\end{aligned} \tag{3-23}$$

APPENDIX E: SMOOTHING ALGORITHMS

E.1. Observation Free Expression for the Smoothed Estimate

Starting with the expression for the smoothed estimate:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\,(\overset{\circ}{\mathbf{v}} - \mathbf{A}\,\delta\hat{\mathbf{x}}) \qquad (3\text{-}21)$$

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 - \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}} \qquad (E\text{-}1)$$

$\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}$ can be extracted from the expression for the filtered estimate.

$$\delta\hat{\mathbf{x}} = (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\,(\mathbf{P}_{\tilde{\mathbf{x}}}\,\delta\tilde{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}}) \qquad (3\text{-}7a)$$

$$\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}} = (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})\,\delta\hat{\mathbf{x}} - \mathbf{P}_{\tilde{\mathbf{x}}}\,\delta\tilde{\mathbf{x}}$$

$$\mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}} = \mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{P}_{\tilde{\mathbf{x}}}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \qquad (E\text{-}2)$$

Thus, the term containing the measurements is expressed as a function of both the predicted and filtered state vector. Substituting now in (E-1):

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 - \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{P}_{\tilde{\mathbf{x}}}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}})$$

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{P}_{\tilde{\mathbf{x}}}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}})$$

Let $$\mathbf{J} = \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{P}_{\tilde{\mathbf{x}}} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\Phi^T\,\mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1} \qquad (3\text{-}27)$$

The smoothed estimate can now be written as:

$$\hat{\hat{\mathbf{x}}}_0 = \hat{\mathbf{x}}_0 + \mathbf{J}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \qquad (3\text{-}25)$$

In a similar manner, the covariance matrix can be transformed.
From (3-23) and (3-24):

$$\mathbf{Q}_{\hat{\hat{\mathbf{x}}}_0\hat{\hat{\mathbf{x}}}_0} = \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0} - \mathbf{P}_{\hat{\mathbf{x}}_0}^{-1}\Phi^T\mathbf{A}^T\mathbf{P}\,(\mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}^T)\,\mathbf{P}\mathbf{A}\Phi\,\mathbf{P}_{\hat{\mathbf{x}}_0}^{-1} \qquad (E\text{-}3)$$

A useful relation can be extracted from (3-8):

$$\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = (\mathbf{P}_{\tilde{\mathbf{x}}} + \mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}$$

Hence $\mathbf{A}^T\mathbf{P}\mathbf{A} = \mathbf{Q}_{\hat{x}\hat{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}$ (E-4)

The central part of the last term in (E-3) is considered separately:

$$
\begin{aligned}
& \mathbf{A}^T\mathbf{P}\,(\mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T)\,\mathbf{P}\mathbf{A} = \\
& = \mathbf{A}^T\mathbf{P}\mathbf{A} - \mathbf{A}^T\mathbf{P}\mathbf{A}\,\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{A}^T\mathbf{P}\mathbf{A} = \\
\text{(E-4)} \Rightarrow \quad & = (\mathbf{Q}_{\hat{x}\hat{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}) - (\mathbf{Q}_{\hat{x}\hat{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1})\,\mathbf{Q}_{\hat{x}\hat{x}}\,(\mathbf{Q}_{\hat{x}\hat{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}) = \\
& = \mathbf{Q}_{\hat{x}\hat{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} - \mathbf{Q}_{\hat{x}\hat{x}}^{-1} + \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} + \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} = \\
& = \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} - \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{Q}_{\tilde{x}\tilde{x}}^{-1} = \\
& = -\,\mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}\,(\mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\tilde{x}\tilde{x}})\,\mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}
\end{aligned}
$$

Substituting now in (E-3):

$$
\begin{aligned}
\mathbf{Q}_{\hat{\hat{x}}_0\hat{\hat{x}}_0} &= \mathbf{Q}_{\hat{x}_0\hat{x}_0} + \mathbf{P}_{\hat{x}_0}^{-1}\Phi^T\,\mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}\,(\mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\tilde{x}\tilde{x}})\,\mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}\Phi\,\mathbf{P}_{\hat{x}_0}^{-1} \\
\mathbf{Q}_{\hat{\hat{x}}_0\hat{\hat{x}}_0} &= \mathbf{Q}_{\hat{x}_0\hat{x}_0} + \mathbf{J}\,(\mathbf{Q}_{\hat{x}\hat{x}} - \mathbf{Q}_{\tilde{x}\tilde{x}})\,\mathbf{J}^T
\end{aligned}
\qquad \text{(3-26)}
$$

The terms explicitly related to the observations l, $\mathbf{Q}_{ll}$, $\mathbf{P}$ and $\mathbf{A}$ have been completely removed.

E.2. From Predicted to Filtered Residuals

Considering the update step of the rearranged MB-F algorithm:

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{K}\mathbf{A})^T\tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{Q}_{\tilde{v}\tilde{v}}^{-1}\,\tilde{\mathbf{v}} \qquad \text{(3-34)}$$

the replacement of the predicted residual through the filtered one is sought. The inverse of the covariance matrix of the predicted residual is obtained by applying the matrix identity (3-10) to (3-33):

$$
\begin{aligned}
& \mathbf{Q}_{\tilde{v}\tilde{v}} = \mathbf{Q}_{ll} + \mathbf{A}^T\mathbf{Q}_{\tilde{x}\tilde{x}}\mathbf{A} && \text{(3-33)} \\
\text{(3-10)} \Rightarrow \quad & \mathbf{Q}_{\tilde{v}\tilde{v}}^{-1} = \mathbf{P} - \mathbf{P}\mathbf{A}\,(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1})^{-1}\,\mathbf{A}^T\mathbf{P} \\
\text{(3-8)} \Rightarrow \quad & \mathbf{Q}_{\tilde{v}\tilde{v}}^{-1} = \mathbf{P} - \mathbf{P}\mathbf{A}\,\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{A}^T\mathbf{P}
\end{aligned}
$$

Thus
$$
\begin{aligned}
\mathbf{Q}_{\tilde{v}\tilde{v}}^{-1}\,\tilde{\mathbf{v}} &= (\mathbf{P} - \mathbf{P}\mathbf{A}\,\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{A}^T\mathbf{P})\,\tilde{\mathbf{v}} \\
&= \mathbf{P}\,\tilde{\mathbf{v}} - \mathbf{P}\mathbf{A}\,\mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{A}^T\mathbf{P}\,\tilde{\mathbf{v}}
\end{aligned}
\qquad \text{(E-5)}
$$

Recalling the alternative expression for the Kalman gain matrix:

$$\mathbf{K} = \mathbf{Q}_{\hat{x}\hat{x}}\,\mathbf{A}^T\mathbf{P} \qquad \text{(C-9)}$$

and inserting in (E-5):

$$\mathbf{Q}_{\tilde{v}\tilde{v}}^{-1}\,\tilde{\mathbf{v}} = \mathbf{P}\,\tilde{\mathbf{v}} - \mathbf{PA}\,\mathbf{K}\,\tilde{\mathbf{v}}$$

The Kalman expression for the adjusted parameters can be used:

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} + \mathbf{K}\,\tilde{\mathbf{v}} \qquad \text{(3-11b)}$$

Inserting in (E-5):

$$\begin{aligned} \mathbf{Q}_{\tilde{v}\tilde{v}}^{-1}\,\tilde{\mathbf{v}} &= \mathbf{P}\,\tilde{\mathbf{v}} - \mathbf{PA}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \\ &= \mathbf{P}\,(\tilde{\mathbf{v}} - \mathbf{A}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}})) \\ \text{(3-37)} \Rightarrow \mathbf{Q}_{\tilde{v}\tilde{v}}^{-1}\,\tilde{\mathbf{v}} &= \mathbf{P}\,\hat{\mathbf{v}} \end{aligned} \qquad \text{(E-6)}$$

The assumption inherent in (3-11b) that the model of the observations has been linearised around the predicted state is not necessary for this proof, and (E-6) can also be obtained using (3-11a), (3-32) and (3-36). However, the manipulations are more tedious. (E-6) can be inserted in (3-34), and the update step becomes:

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{KA})^T\,\tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} \qquad \text{(3-40)}$$

The update of the adjoint covariance matrix is obtained in a similar fashion.

E.3. From Filtered to Smoothed Residuals

The replacement of the filtered residuals through the smoothed ones starts with:

$$\hat{\mathbf{v}} = \hat{\hat{\mathbf{v}}} + \mathbf{A}(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) \qquad \text{(3-42)}$$

Considering the smoothing step in (3-40), the filtered and smoothed state vectors are related through:

$$\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}} = - \mathbf{Q}_{\hat{x}\hat{x}}\, \tilde{\mathbf{z}} \qquad \text{(E-7)}$$

Inserting (E-7) in (3-42), the expression for the vector of filtered residuals becomes

$$\hat{\mathbf{v}} = \hat{\hat{\mathbf{v}}} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\, \tilde{\mathbf{z}}$$

Substituting in the update step (3-40)

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{K}\mathbf{A})^T \tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,(\hat{\hat{\mathbf{v}}} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\, \tilde{\mathbf{z}})$$

Recalling the alternative expression of the Kalman gain matrix (C-9)

$$\hat{\mathbf{z}} = (\mathbf{I} - \mathbf{Q}_{\hat{x}\hat{x}}\, \mathbf{A}^T\mathbf{P}\mathbf{A})^T \tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,(\hat{\hat{\mathbf{v}}} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\, \tilde{\mathbf{z}})$$
$$\hat{\mathbf{z}} = \tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\, \tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}} + \mathbf{A}^T\mathbf{P}\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\, \tilde{\mathbf{z}}$$

Hence $$\hat{\mathbf{z}} = \tilde{\mathbf{z}} - \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}} \qquad \text{(3-43)}$$

The update of the adjoint covariance matrix:

$$\mathbf{Q}_{\hat{z}\hat{z}} = \mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{A}^T\mathbf{K}^T\, \mathbf{Q}_{\tilde{z}\tilde{z}} - \mathbf{Q}_{\tilde{z}\tilde{z}}\, \mathbf{K}\mathbf{A} + \mathbf{A}^T\mathbf{P}\, \mathbf{Q}_{\hat{\hat{v}}\hat{\hat{v}}}\, \mathbf{P}\mathbf{A}$$

is obtained in a similar fashion.

E.4. Equivalence of the New Algorithm with R-T-S

The proof is based on the classical principle of recursive algorithms:

1) the first step of the algorithm is correct
2) the next step of the algorithm is correct

For the state vector:

1) For the first epoch to smooth, no subsequent measurement is available and thus $\hat{\hat{\mathbf{x}}} = \hat{\mathbf{x}}$ and $\hat{\hat{\mathbf{v}}} = \hat{\mathbf{v}}$.

To prove: $$\mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} = \mathbf{Q}_{\tilde{x}\tilde{x}}^{-1}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \qquad \text{(E-8)}$$

$$
\begin{aligned}
(3\text{-}36) \Rightarrow \qquad \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} &= \mathbf{A}^T\mathbf{P}\,(\overset{\circ}{\mathbf{v}} - \mathbf{A}\delta\hat{\mathbf{x}}) \\
&= -\,\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\,\overset{\circ}{\mathbf{v}} \\
(E\text{-}2) \Rightarrow \qquad &= -\,\mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{A}^T\mathbf{P}\mathbf{A}\,\delta\hat{\mathbf{x}} + \mathbf{P}_{\tilde{\mathbf{x}}}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \\
&= \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}})
\end{aligned}
$$

2) The relation between $\hat{\hat{\mathbf{x}}}_0$ and $\hat{\hat{\mathbf{x}}}$ is established, by considering a full cycle of the new algorithm (3-48):

$$
\begin{aligned}
\tilde{\mathbf{z}} &= -\,\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && \text{smoothing} \\
\hat{\mathbf{z}} &= -\,\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) + \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}} && \text{update} \\
\tilde{\mathbf{z}}_0 &= -\,\Phi^T\left(\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) + \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}}\right) && \text{prediction} \\
\hat{\hat{\mathbf{x}}}_0 &= \hat{\mathbf{x}}_0 + \mathbf{Q}_{\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0}\,\Phi^T\left(\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) + \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}}\right) && \text{smoothing}
\end{aligned}
$$

Comparing with the R-T-S expression, it appears that the equivalence requires:

$$
\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) + \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}} = \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \tilde{\mathbf{x}}) \qquad (E\text{-}9)
$$

The proof follows:

$$
\begin{aligned}
& \mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && + \mathbf{A}^T\mathbf{P}\,\hat{\hat{\mathbf{v}}} = \\
(3\text{-}42) \Rightarrow \quad & \mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && + \mathbf{A}^T\mathbf{P}\,(\hat{\mathbf{v}} - \mathbf{A}(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}})) = \\
& \mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && + \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} - \mathbf{A}^T\mathbf{P}\mathbf{A}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) = \\
& (\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1} - \mathbf{A}^T\mathbf{P}\mathbf{A})\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && + \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} = \\
& \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && + \mathbf{A}^T\mathbf{P}\,\hat{\mathbf{v}} = \\
(E\text{-}8) \Rightarrow \quad & \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \hat{\mathbf{x}}) && + \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}\,(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) = \mathbf{Q}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^{-1}\,(\hat{\hat{\mathbf{x}}} - \tilde{\mathbf{x}})
\end{aligned}
$$

The proof of the equivalence for the covariance matrix is similar.

APPENDIX F: KINEMATIC MODELS

F.1. Continuous Form

Considering a one-dimensional, third order system where only the acceleration drift is affected by a white random noise u, (4-8) becomes:

$$\dot{\mathbf{x}} = \mathbf{F}\,\mathbf{x} + \mathbf{G}u$$

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \\ \dddot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\alpha \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \qquad \text{(F-1)}$$

In addition, the driving noise u is assumed to have a Gaussian probability distribution and expectation zero. Note that the third line of (F-1) describes a first order Gauss-Markov process:

$$\dddot{x} = -\alpha\,\ddot{x} + u$$

However, the second line of (F-1) does <u>not</u> describe a second order Markov process, see definition (4-2), although $\ddot{x}$ is a second order random process.

The time behaviour of the system in each dimension may be described by differential equations of various orders. As long as each dimension has its own driving noise, the extension from a scalar function to a vector of independent scalar functions is straightforward: it will result in a block diagonal structure of both Φ and $\mathbf{Q}_{ww}$. This situation is assumed in DYNAMO. For example, such a structure is not obtained if the driving noise is assumed to affect the heading of the ship (or one of its time derivatives) whilst Easting and Northing are estimated. In such a case, there will be correlations between the noises affecting Easting and Northing (and their time derivatives).

F.2. Transition Matrix

In order to use the Taylor series expansion (4-10)

$$\Phi(t,t_0) = \mathbf{I} + \mathbf{F}(t - t_0) + \frac{1}{2!}\mathbf{FF}(t - t_0)^2 + \frac{1}{3!}\mathbf{FFF}(t - t_0)^3 + \ldots$$

powers of **F** must first be computed:

$$\mathbf{F}^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -\alpha \\ 0 & 0 & \alpha^2 \end{bmatrix}, \quad \mathbf{F}^3 = \begin{bmatrix} 0 & 0 & -\alpha \\ 0 & 0 & \alpha^2 \\ 0 & 0 & -\alpha^3 \end{bmatrix}, \quad \mathbf{F}^4 = \begin{bmatrix} 0 & 0 & \alpha^2 \\ 0 & 0 & -\alpha^3 \\ 0 & 0 & \alpha^4 \end{bmatrix} \quad \text{and so on ...}$$

Thus the transition matrix is of the form:

$$\Phi = \begin{bmatrix} 1 & 0 & \phi_3 \\ 0 & 1 & \phi_2 \\ 0 & 0 & \phi_1 \end{bmatrix} \qquad \text{where:}$$

$$\phi_1 = 1 + \frac{(-\alpha\Delta t)}{1!} + \frac{(-\alpha\Delta t)^2}{2!} + \frac{(-\alpha\Delta t)^3}{3!} + \frac{(-\alpha\Delta t)^4}{4!} + \ldots = e^{-\alpha\Delta t}$$

$$\phi_2 = 0 + 1 + \frac{(-\alpha\Delta t)}{2!} + \frac{(-\alpha\Delta t)^2}{3!} + \frac{(-\alpha\Delta t)^3}{4!} + \ldots = (1 - e^{-\alpha\Delta t}) / \alpha$$

$$\phi_3 = 0 + 0 + 1 + \frac{(-\alpha\Delta t)}{3!} + \frac{(-\alpha\Delta t)^2}{4!} + \ldots = (e^{-\alpha\Delta t} + \alpha\Delta t - 1) / \alpha^2$$

F.3. Covariance Matrix of the System Noise

The formula (4-18) is applied: $\mathbf{Q}_{ww} = \int_{t_0}^{t} \Phi(t,\tau)\, \mathbf{G}\, \mathbf{Q'}_{uu}\, \mathbf{G}^T\, \Phi^T(t,\tau)\, d\tau$

In this case: $\Phi(t,\tau)\, \mathbf{G}\, \mathbf{Q'}_{uu}\, \mathbf{G}^T\, \Phi^T(t,\tau) = \begin{bmatrix} 1 & 0 & \phi_3 \\ 0 & 1 & \phi_2 \\ 0 & 0 & \phi_1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} q'_{uu} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \phi_3 & \phi_2 & \phi_1 \end{bmatrix}$

$$= \begin{bmatrix} \phi_3^2 & \phi_3\phi_2 & \phi_3\phi_1 \\ \phi_2\phi_3 & \phi_2^2 & \phi_2\phi_1 \\ \phi_1\phi_3 & \phi_1\phi_2 & \phi_1^2 \end{bmatrix} q'_{uu}$$

The integration results in: $\mathbf{Q_{ww}} = \begin{bmatrix} q_{33} & q_{32} & q_{31} \\ q_{23} & q_{22} & q_{21} \\ q_{13} & q_{12} & q_{11} \end{bmatrix}$

where:

$$
\begin{aligned}
q_{11} &= q'_{uu} (1 - e^{-2\alpha\Delta t}) / 2\alpha \\
q_{12} = q_{21} &= q'_{uu} (1 - 2e^{-\alpha\Delta t} + e^{-2\alpha\Delta t}) / 2\alpha^2 \\
q_{22} &= q'_{uu} (- 3 + 2\alpha\Delta t + 4e^{-\alpha\Delta t} - e^{-2\alpha\Delta t}) / 2\alpha^3 \\
q_{13} = q_{31} &= q'_{uu} (1 - 2\alpha\Delta t e^{-\alpha\Delta t} - e^{-2\alpha\Delta t}) / 2\alpha^3 \\
q_{23} = q_{32} &= q'_{uu} (1 - 2\alpha\Delta t + \alpha^2\Delta t^2 + 2\alpha\Delta t e^{-\alpha\Delta t} - 2e^{-\alpha\Delta t} + e^{-2\alpha\Delta t}) / 2\alpha^4 \\
q_{33} &= q'_{uu} (1 + 2\alpha\Delta t - 2\alpha^2\Delta t^2 + 2/3\alpha^3\Delta t^3 - 4\alpha\Delta t e^{-\alpha\Delta t} - e^{-2\alpha\Delta t}) / 2\alpha^5
\end{aligned}
$$

The uncommon numbering of the rows and columns is justified, because a limitation to order 2 or 1 reduces the matrices Φ and $\mathbf{Q_{ww}}$ to the corresponding lower right part of the matrices developed for order 3.

F.4. Random Walk as a Special Case

If the correlation time tends to infinity, that is $\alpha \rightarrow 0$, the noise on the acceleration becomes a random walk and the expressions of Φ and $\mathbf{Q_{ww}}$ simplify to:

$$
\begin{aligned}
\phi_1 &= 1 \\
\phi_2 &= \Delta t \\
\phi_3 &= \Delta t^2/2
\end{aligned}
$$

$$
\begin{aligned}
q_{11} &= q'_{uu} \Delta t && \ddot{x} \text{ is a random walk} \\
q_{12} = q_{21} &= q'_{uu} \Delta t^2/2 \\
q_{22} &= q'_{uu} \Delta t^3/3 && \dot{x} \text{ is a second order random walk} \\
q_{13} = q_{31} &= q'_{uu} \Delta t^3/6 \\
q_{23} = q_{32} &= q'_{uu} \Delta t^4/8 \\
q_{33} &= q'_{uu} \Delta t^5/20 && x \text{ is a third order random walk}
\end{aligned}
$$

Publications from

THE SCHOOL OF SURVEYING, THE UNIVERSITY OF NEW SOUTH WALES.

All prices include postage by surface mail. Air mail rates on application. (Effective April 1989)

To order, write to Publications Officer, School of Surveying, The University of New South Wales, P.O. Box 1, Kensington N.S.W., 2033 AUSTRALIA

NOTE: ALL ORDERS MUST BE PREPAID

UNISURV REPORTS - G SERIES

Price (including postage): $3.50

G14. A. Stolz, "The computation of three dimensional Cartesian coordinates of terrestrial networks by the use of local astronomic vector systems", Unisurv Rep. 18, 47 pp.

G16. R.S. Mather et al, "Communications from Australia to Section V, International Association of Geodesy, XV General Assembly, International Union of Geodesy and Geophysics, Moscow 1971", Unisurv Rep. 22, 72 pp.

G17. Papers by R.S. Mather, H.L. Mitchell & A. Stolz on the following topics:- Four-dimensional geodesy, Network adjustment and Sea surface topography, Unisurv G17, 73 pp.

G18. Papers by L. Berlin, G.J.F. Holden, P.V. Angus-Leppan, H.L. Mitchell & A.H. Campbell on the following topics:- Photogrammetry co-ordinate systems for surveying integration, Geopotential networks and Linear measurement, Unisurv G18, 80 pp.

G19. R.S. Mather, P.V. Angus-Leppan, A. Stolz & I. Lloyd, "Aspects of four-dimensional geodesy", Unisurv G19, 100 pp.

G20. Papers by J.S. Allman, R.C. Lister, J.C. Trinder & R.S. Mather on the following topics:- Network adjustments, Photogrammetry, and 4-Dimensional geodesy, Unisurv G20, 133 pp.

G21. Papers by E. Grafarend, R.S. Mather & P.V. Angus-Leppan on the following topics:- Mathematical geodesy, Coastal geodesy and Refraction, Unisurv G21, 100 pp.

G22. Papers by R.S. Mather, J.R. Gilliland, F.K. Brunner, J.C. Trinder, K. Bretreger & G. Halsey on the following topics:- Gravity, Levelling, Refraction, ERTS imagery, Tidal effects on satellite orbits and Photogrammetry, Unisurv G22, 96 pp.

G23. Papers by R.S. Mather, E.G. Anderson, C. Rizos, K. Bretreger, K. Leppert, B.V. Hamon & P.V. Angus-Leppan on the following topics:- Earth tides, Sea surface topography, Atmospheric effects in physical geodesy, Mean sea level and Systematic errors in levelling, Unisurv G23, 96 pp.

G24. Papers by R.C. Patterson, R.S. Mather, R. Coleman, O.L. Colombo, J.C. Trinder, S.U. Nasca, T.L. Duyet & K. Bretreger on the following topics:- Adjustment theory, Sea surface topography determinations, Applications of LANDSAT imagery, Ocean loading of Earth tides, Physical geodesy, Photogrammetry and Oceanographic applications of satellites, Unisurv G24, 151 pp.

G25. Papers by S.M. Nakiboglu, B. Ducarme, P. Melchior, R.S. Mather, B.C. Barlow, C. Rizos, B. Hirsch, K. Bretreger, F.K. Brunner & P.V. Angus-Leppan on the following topics:- Hydrostatic equilibrium figures of the Earth, Earth tides, Gravity anomaly data banks for Australia, Recovery of tidal signals from satellite altimetry, Meteorological parameters for modelling terrestrial refraction and Crustal motion studies in Australia, Unisurv G25, 124 pp.

G26. Papers by R.S. Mather, E.G. Masters, R. Coleman, C. Rizos, B. Hirsch, C.S. Fraser, F.K. Brunner, P.V. Angus-Leppan, A.J. McCarthy & C. Wardrop on the following topics:- Four-dimensional geodesy, GEOS-3 altimetry data analysis, analysis of meteorological measurements for microwave EDM and Meteorological data logging system for geodetic refraction research, Unisurv G26, 113 pp.

G27. Papers by F.K. Brunner, C.S. Fraser, S.U. Nasca, J.C. Trinder, L. Berlin, R.S. Mather, O.L. Colombo & P.V. Angus-Leppan on the following topics:- Micrometeorology in geodetic refraction, LANDSAT imagery in topographic mapping, adjustment of large systems, GEOS-3 data analysis, Kernel functions and EDM reductions over sea, Unisurv G27, 101 pp.

G29. Papers by F.L. Clarke, R.S. Mather, D.R. Larden & J.R. Gilliland on the following topics:- Three dimensional network adjustment incorporating ξ, η and N, Geoid determinations with satellite altimetry, Geodynamic information from secular gravity changes and Height and free-air anomaly correlation, Unisurv G29, 87 pp.

From June 1979 Unisurv G's name was changed to Australian Journal of Geodesy, Photogrammetry and Surveying. These can be ordered from The Managing Editor, Australian Journal of Geodesy, Photogrammetry and Surveying, Institution of Surveyors - Australia , Box 4793 G.P.O., Sydney, N.S.W., 2001, AUSTRALIA.

UNISURV REPORTS - S SERIES

S8 - S19	Price (including postage):		$7.50
S20 onwards	Price (including postage):	Individuals	$18.00
		Institutions	$25.00

S8 A. Stolz, "Three-D Cartesian co-ordinates of part of the Australian geodetic network by the use of local astronomic vector systems", Unisurv Rep. S 8, 182 pp, 1972.

S9 H.L. Mitchell, "Relations between MSL & geodetic levelling in Australia", Unisurv Rep. S 9, 264 pp, 1973.

S10 A.J. Robinson, "Study of zero error & ground swing of the model MRA101 tellurometer", Unisurv Rep. S 10, 200 pp, 1973.

S12. G.J.F. Holden, "An evaluation of orthophotography in an integrated mapping system", Unisurv Rep. S 12, 232 pp, 1974.

S14. Edward G. Anderson, "The Effect of Topography on Solutions of Stokes` Problem", Unisurv Rep. S 14, 252 pp, 1976.

S15. A.H.W. Kearsley, "The Computation of Deflections of the Vertical from Gravity Anomalies", Unisurv Rep. S 15, 181 pp, 1976.

S16. K. Bretreger, "Earth Tide Effects on Geodetic Observations", Unisurv S 16, 173 pp, 1978.

S17. C. Rizos, "The role of the gravity field in sea surface topography studies", Unisurv S 17, 299 pp, 1980.

S18. B.C. Forster, "Some measures of urban residual quality from LANDSAT multi-spectral data", Unisurv S 18, 223 pp, 1981.

S19. Richard Coleman, "A Geodetic Basis for recovering Ocean Dynamic Information from Satellite Altimetry", Unisurv S 19,332 pp, 1981.

S20. Douglas R. Larden, "Monitoring the Earth's Rotation by Lunar Laser Ranging", Unisurv Report S 20, 280 pp, 1982.

S25. Ewan G. Masters, "Applications of Satellite Geodesy to Geodynamics", Unisurv Report S25, 208 pp, 1984.

S27. Bruce R. Harvey, "The Combination of VLBI and Ground Data for Geodesy and Geophysics", Unisurv Report S27, 239 pp, 1985.

S28. Rod Eckels, "Surveying with GPS in Australia", Unisurv S28, 220 pp, 1987.

S29. Gary S. Chisholm, "Integration of GPS into hydrographic survey operations", Unisurv S29, 190 pp, 1987.

S30. Gary Alan Jeffress, "An investigation of Doppler satellite positioning multi-station software", Unisurv S30, 118 pp, 1987.

S31. Jahja Soetandi, "A model for a cadastral land information system for Indonesia", Unisurv S31, 168 pp, 1988.

S33. R. D. Holloway, "The integration of GPS heights into the Australian Height Datum", Unisurv S33, 151 pp.,1988.

S34. Robin C. Mullin, "Data update in a Land Information Network", Unisurv S34, 168 pp. 1988.

S35. Bertrand Merminod, "The use of Kalman filters in GPS Navigation", Unisurv S35, 203 pp., 1989.

PROCEEDINGS

Prices include postage by surface mail

P1. P.V. Angus-Leppan (Editor), "Proceedings of conference on refraction effects in geodesy & electronic distance measurement", 264 pp. Price: $6.50

P2. R.S. Mather & P.V. Angus-Leppan (Eds), "Australian Academy of Science/International Association of Geodesy Symposium on Earth`s Gravitational Field & Secular Variations in Position", 740 pp. Price $12.50

MONOGRAPHS

Prices include postage by surface mail

M1. R.S. Mather, "The theory and geodetic use of some common projections", (2nd edition), 125 pp. Price $9.00

M2. R.S. Mather, "The analysis of the earth's gravity field", 172 pp. Price $5.50

M3. G.G. Bennett, "Tables for prediction of daylight stars", 24 pp. Price $2.00

M4. G.G. Bennett, J.G. Freislich & M. Maughan, "Star prediction tables for the fixing of position", 200 pp. Price $5.00

M5. M. Maughan, "Survey computations", 98 pp. Price $8.00

M6. M. Maughan, "Adjustment of Observations by Least Squares", 61 pp. Price $7.00

M7. J.M. Rueger, "Introduction to Electronic Distance Measurement", (2nd Edition), 140 pp. Price $14.00

M8. A.H.W. Kearsley, "Geodetic Surveying". 77pp. Price $8.00

M9. ** R.W. King, E.G. Masters, C. Rizos, A. Stolz and J. Collins, "Surveying with GPS", 128 pp .

M10. W. Faig, "Aerial Triangulation and Digital Mapping", 102. pp. Price $13.00

M11. W.F. Caspary, "Concepts of Network and Deformation Analysis", 183 pp. Price $22.00

M12. F.K. Brunner, "Atmospheric Effects on Geodetic Space Measurements", 110 pp. Price $13.00

** No longer for external sale from the School of Surveying. Now published by Ferd. Dümmlers Verlag, Kaiserstrasse 31 - 37, D - 5300 Bonn 1, Federal Republic of Germany