

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/282254852>

High Integrity Navigation

Thesis · December 1997

DOI: 10.13140/RG.2.1.1695.2806

CITATIONS
13

READS
364

1 author:



Steven John Scheding
The University of Sydney

93 PUBLICATIONS 1,572 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Connected and Automated Vehicle [View project](#)



Multi-Sensor Data Fusion for Map Building & SLAM [View project](#)

High Integrity Navigation

Steven Scheding

A thesis submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy



Australian Centre for Field Robotics
Department of Mechanical and Mechatronic Engineering
The University of Sydney

December 1997

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the University or other institute of higher learning, except where due acknowledgement has been made in the text.

Steven Scheding

December 16, 1997

Abstract

Steven Scheding
The University of Sydney

Doctor of Philosophy
December 1997

High Integrity Navigation

This thesis examines the issue of the design of high integrity navigation systems for Autonomous Guided Vehicles (AGVs). The term high integrity is used to describe a system that is robust to failure, and at the very least can be guaranteed upon failure to attain a ‘safe’ state. This thesis seeks to provide this guarantee by determining the necessary conditions that are required for a system to be considered of high integrity.

By designing for high integrity, the designer is acknowledging the fact that at some point in time, the system *will* fail. The time between failures may be maximised by choosing high reliability components, but nevertheless the components (whether they be sensors, actuators, mechanics, computers or even mathematical models) will at some time fail in potentially unpredictable ways.

This thesis has three principle contributions. The first is the determination of the conditions necessary for a navigation system to have high integrity. These conditions may be summarised as: The navigation system must be designed so as to meet the design specifications, it must have a fault detection scheme, and *all* faults must be guaranteed to be detectable. The second contribution is the development of metrics which describe the detectability of faults within the navigation system, and from these metrics, the development of methods which effectively guarantee that all faults will be detectable. The final contribution of this thesis is the development of a consistent methodology for the overall design of high integrity navigation systems. A numerical method is developed which allows the minimum quality sensors to be chosen for which the navigation system meets the design specification. It is shown that frequency domain redundancy is necessary to ensure fault detectability, and that the navigation system should be implemented in a distributed form to minimise the effects of software or computer failure.

The concepts described in this thesis are applied to two very different experiments. The first experiment, using a scanning laser and a gyroscope, is designed to validate the theory developed in this thesis. This is achieved by comparing experimental results to those predicted by the theory. The second experiment details an ongoing project into the design of a high integrity navigation system for a Load Haul and Dump truck (LHD). The proposed navigation architecture is presented, together with initial results from the navigation sub-systems, using real data obtained in an underground mine in Queensland, Australia.

Acknowledgements

I would like to thank all of the people who encouraged and supported me during the undertaking of this research.

Firstly I would like to thank my supervisor, Dr. Eduardo Nebot, for giving me the opportunity to pursue postgraduate study at the University of Sydney. Eduardo's support and unwavering enthusiasm for *all things* has proved invaluable. I would also like to thank my co-supervisor Professor Hugh Durrant-Whyte, without whom I would still be planning paths for a now defunct indoor mobile robot. Hugh's support and expertise have also proved invaluable.

Thanks goes to the entire Sydney University mechatronics crowd: Daniel (Hangover #2056) Pagac, Hong Quang, Monica Lauda, Xiaoying (I'm thinking) Kong, Gaius Chan, Mohammad (recently departed) Bozorg, Raj Madhavan, Steve Clark, Salah Sukkarieh, Paul (not a joke) Newman, Anh Tuan Le, Adrian Bonchis, and Tim (I just want to be called Dr.) Bailey. Thanks also to the visiting scholars: Simon (scooper) Cooper, whose research was the inspiration for this thesis, Simon (baby) Julier, for being clever and for taking the time to prove it to me, Mihaly (I am not a Communist) Csorba, for providing a challenging opponent in some often intense games of Quake. Special thanks also go to K. Y. Lee for being a friend and for occasionally fixing my computer, to Keith Willis for providing cannon fodder during late night Quake-ing, and to Mike (computer guru) Stevens, for knowing the answer to everything.

Perhaps my greatest debt is to those people who have supported me during this entire process: My family and friends. A special thanks is due to Jody for feeding and supporting me above and beyond the call of duty.

To the people who helped

Contents

Declaration	i
Abstract	ii
Acknowledgements	iii
Contents	v
List of Figures	vii
List of Tables	viii
1 Introduction	1
1.1 The Objectives of this Thesis	1
1.2 Navigation	2
1.3 The Kalman Filter	4
1.4 Navigation Failure	5
1.4.1 Reliability	6
1.4.2 Integrity	6
1.5 Contributions of this Thesis	7
1.6 Thesis Structure	7
2 Navigation Systems and Their Failure Modes	10
2.1 Introduction	10
2.2 Navigation	11
2.2.1 The Continuous Time Kalman Filter	11

2.2.2	Navigation in the Frequency Domain	14
2.2.3	A Simple Example	15
2.2.4	Alternate Filter Structures	19
2.3	The Nature of Failures in Navigation Systems	22
2.4	Modelling Sensor Faults	23
2.4.1	Error Models Via Power Spectral Density Estimation	24
2.4.2	Shaping Filters	25
2.5	The Extended Kalman Filter	29
2.6	The Discrete Kalman Filter	31
2.7	Modelling Process Faults	33
2.7.1	The Tricycle Drive Robot	33
2.7.2	Error Models	35
2.7.3	Sensitivity Analysis	38
2.8	Characteristic Sensor Errors	39
2.8.1	Absolute Sensors	39
2.8.2	Rate Sensors	41
2.8.3	Sensor Error Models	41
2.9	Sensor Error and Filter Transfer	42
2.9.1	Type I Sensors	42
2.9.2	Type II Sensors	43
2.9.3	Type III Sensors	45
2.9.4	Dual Type I Sensors	46
2.9.5	Type I and Type II Sensors	47
2.9.6	Type I and Type III Sensors	52
2.9.7	Summary	57
2.10	Frequency Interpretation of the EKF	58
2.10.1	A Tricycle Drive Vehicle Example	58
2.11	Conclusion	61

3 Fault Detection	63
3.1 Introduction	63
3.2 The Need for Fault Detection	64
3.3 Fault Detection	65
3.4 Detecting Unmodelled Faults	69
3.4.1 Process faults	69
3.4.2 Process Fault Detectability	69
3.4.3 Sensor Faults	70
3.4.4 Sensor Fault Detectability	72
3.4.5 Quantifying Process Fault Detectability	73
3.4.6 Quantifying Sensor Fault Detectability	76
3.4.7 Summary	77
3.5 Guaranteeing Process Fault Detection	79
3.6 Guaranteeing Sensor Fault Detection	81
3.6.1 Redundancy	81
3.6.2 Frequency Domain Redundancy	83
3.6.3 An Example of Frequency Domain Redundancy	84
3.7 Nonlinear and Time Varying Systems	90
3.8 Fault Detection	93
3.9 Conclusion	94
4 High Integrity Design	95
4.1 Introduction	95
4.2 The Specification of Navigation Systems	96
4.3 Achieving the Specification	97
4.3.1 Sensor Suites	97
4.4 Generalising Sensor Suite Design	100
4.4.1 A Single Sensor Example	102
4.4.2 The Riccati Equation	103
4.4.3 Numerical Techniques	104
4.5 Designing for High Integrity	105

4.5.1	Redundancy	105
4.5.2	High Integrity Filter Structures	106
4.5.3	Other Architectures	110
4.6	Conclusion	111
5	Experimental Results	112
5.1	Introduction	112
5.2	The Gyro-Laser Experiment	112
5.2.1	Experimental Setup	113
5.2.2	Filter Design and Analysis	116
5.2.3	Results: The Nominal System	119
5.2.4	Results: in the Presence of Gyro Bias	128
5.2.5	Results: in the Presence of Laser Bias	129
5.2.6	Results: Laser Bias with Redundancy	132
5.2.7	Summary	135
5.3	A High Integrity Navigation System for an LHD	136
5.3.1	Experimental Setup	138
5.3.2	Navigation Architecture	140
5.3.3	The Feature-Based Navigation System	141
5.3.4	The Geometric-Based Navigation System	141
5.3.5	The Vehicle Model	141
5.3.6	The No-Slip Model	143
5.3.7	Accounting for Slip	143
5.3.8	The Importance of Slip	145
5.3.9	The Error Model	146
5.3.10	Observation Model	147
5.3.11	Implementing the Navigation System	148
5.3.12	Results	149
5.3.13	Summary	150
5.4	Conclusion	154

6 Summary and Conclusions	155
6.1 Introduction	155
6.2 Summary	155
6.3 Contributions	158
6.3.1 Sensor Error Models	158
6.3.2 Navigation Failure	158
6.3.3 Fault Detectability Metrics	159
6.3.4 Sensor Suite Selection	159
6.3.5 High Integrity Navigation System Design	159
6.3.6 Experimental Investigations	160
6.4 Future Work	160
6.4.1 Fault Detection	160
6.4.2 Identification	160
6.4.3 High Integrity LHD Navigation	161
A Power Spectral Density Estimation	162
B Continuous to Discrete Time Conversion	164
C Innovation Whiteness Testing	166
D A Numerical Solution of the Riccati Equation	169
E The Gyro-Laser Experiment Discretisation Equations	171
E.1 Discretisation	171
F LHD Discretisation Equations	174
F.1 Discrete Vehicle Model	174
F.2 State Prediction	175
F.3 Observation Prediction	178
F.4 Update Equations	179
Bibliography	180

List of Figures

2.1	Block Diagram of Continuous Time Kalman Filter	14
2.2	Bode Magnitude Plot - Transfer from Position Sensor to Position State	18
2.3	Bode Magnitude Plot - Transfer from Position Sensor to Velocity State	19
2.4	Direct Kalman Filter	21
2.5	Indirect Feedforward Kalman Filter	21
2.6	Indirect Feedback Kalman Filter	21
2.7	Sensor Model - Observation Corrupted by White Noise	26
2.8	Sensor Model - Observation Corrupted by Coloured Noise	26
2.9	Sensor Model - Observation Corrupted by Coloured Noise Generated from a White Noise Source	26
2.10	Augmented System Model with Observation Noise Shaping Filter	28
2.11	Some Shaping Filters and their Corresponding Autocorrelation and PSD Functions	28
2.12	Block Diagram of Continuous Time Kalman Filter with Shaping State(s)	29
2.13	The Discrete Estimation Process	32
2.14	Tricycle Drive Mobile Robot Geometry	34
2.15	Tricycle Drive Mobile Robot Error Source Geometry	36
2.16	Bode Diagram of the Transfer Function of a Single Type I Sensor	43
2.17	Bode Diagram of the Transfer Function of a Single Type II Sensor	44
2.18	Bode Diagram of the Transfer Function of Dual Type I Sensors	46
2.19	Bode Diagram of the Combined Transfer Function of Dual Type I Sensors	47
2.20	Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type II Sensors	49
2.21	Bode Diagram of the Transfer Functions from a Type II Sensor for a System Consisting of Type I and Type II Sensors	49

2.22	Bode Diagram of the Combined Transfer Function of Type I and Type II Sensors to the Position State	50
2.23	Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type II Sensors	51
2.24	Bode Diagram of the Transfer Functions from a Type II Sensor for a System Consisting of Type I and Type II Sensors	51
2.25	Bode Diagram of the Combined Transfer Function of Type I and Type II Sensors to the Position State	52
2.26	Bode Diagram of the Combined Transfer Function of Type I and Type II Sensors to the Position State, Showing Frequency ‘Gap’	53
2.27	Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type III Sensors	54
2.28	Bode Diagram of the Transfer Functions from a Type III Sensor for a System Consisting of Type I and Type III Sensors	55
2.29	Bode Diagram of the Combined Transfer Function of Type I and Type III Sensors to the Position State	55
2.30	Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type III Sensors	56
2.31	Bode Diagram of the Transfer Functions from a Type III Sensor for a System Consisting of Type I and Type III Sensors	56
2.32	Bode Diagram of the Combined Transfer Function of Type I and Type III Sensors to the Position State	57
2.33	Bode Magnitude Plot of the Transfer Function From the Range Measurement to the Position States x (top) and y (bottom)	60
2.34	Bode Magnitude Plot of the Transfer Function From the Bearing Measurement to the Position States x (top) and y (bottom)	60
2.35	Bode Magnitude Plot of the Transfer Function From the Range Measurement (top) and the Bearing Measurement (bottom) to the Orientation State ϕ	61
3.1	The Factors that may Lead to Navigation Failure	64
3.2	The Nature of Fault Processes	65
3.3	Kalman Filter Block Diagram with Injected Process Fault	69
3.4	Some Common Sensor Hardware Failures	71
3.5	Fault Plot - Showing Percentage of Process Fault Transmitted to the Innovations and the Corresponding Phase for a Process Fault Injected into the Position State	74

3.6	Fault Plot - Showing Percentage of Process Fault Transmitted to the Innovations and the Corresponding Phase for a Process Fault Injected into the Velocity State	75
3.7	Bode Plot of the Filter Transfer Function - Top Two Subplots are the Magnitude and Phase Plots from the Sensor to the Estimate \hat{x} - The Bottom Two Subplots are the Magnitude and Phase Plots from the Sensor to the Estimate $\hat{\dot{x}}$	76
3.8	Fault Plot - Showing Percentage of Sensor Fault Transmitted to the Innovations and the Corresponding Phase	77
3.9	Fault Plot - Showing Percentage of Process Fault Injected into the Position State Transmitted to the (a) Position Innovation and (b) the Velocity Innovation	80
3.10	Fault Plot - Showing Percentage of Process Fault Injected into the Velocity State Transmitted to the (a) Position Innovation and (b) the Velocity Innovation	80
3.11	Bode Plot of the Filter Transfer Function - Identical for Both Sensors	82
3.12	Fault Plot - Fault in One Sensor Only	82
3.13	Fault Plot - Identical Faults in Both Sensors	83
3.14	Fault Plot - Percentage of Sonar Fault Transmitted to (a) Sonar Innovations and (b) the Laser Innovations	85
3.15	Fault Plot - Percentage of Laser Fault Transmitted to (a) Laser Innovations and (b) the Sonar Innovations	85
3.16	Map of the Sensor Trial Area	86
3.17	Raw Sonar Data	87
3.18	Raw Laser Data	87
3.19	Estimated Range	89
3.20	(a) Laser Innovation Sequence with 1σ and 2σ bounds and (b) Sonar Innovation Sequence with 1σ and 2σ bounds	89
3.21	Fault Plot - Percentage of Range Measurement Fault Transmitted to Range Measurement Innovations	91
3.22	Fault Plot - Percentage of Range Measurement Fault Transmitted to Bearing Measurement Innovations	91
3.23	Fault Plot - Percentage of Bearing Measurement Fault Transmitted to Bearing Measurement Innovations	92
3.24	Fault Plot - Percentage of Bearing Measurement Fault Transmitted to Range Measurement Innovations	92
4.1	Bound Plot Showing Desired Frequency Response	97

4.2	Frequency Response of Sensor with Time Correlated Measurement Noise	98
4.3	(a) Frequency Response of Sensor with Time Correlated Measurement Noise and (b) Frequency Response of Poor Quality Aiding Sensor	99
4.4	Overall Frequency Response of Sensor Suite	100
4.5	A Comparison of Sensor Types with their Transfer Functions and Detectability Plots	101
4.6	Psuedocode to Find the set of Sensor Noise Covariance Matrices	105
4.7	Direct Kalman Filter with Two Sensor Suites	107
4.8	High Integrity Kalman Filter with Two Sensor Suites	107
4.9	High Integrity Kalman Filter State Innovation Construction	108
5.1	The GCS Laser	114
5.2	The Andrews Gyro	114
5.3	The Experimental Setup	115
5.4	Bode Diagram of the Transfer Functions from (a) the Laser and (b) the Gyro to the states ϕ , $\dot{\phi}$ and x_{sf}	118
5.5	Bode Diagram of the Combined Transfer Functions to the State ϕ	119
5.6	Fault Plots - Percentage of Fault in Laser Sensor Transmitted to (a) the Laser Innovations (b) the Gyro Innovations	120
5.7	Fault Plots - Percentage of Fault in Gyro Transmitted to (a) the Laser Innovations (b) the Gyro Innovations	120
5.8	Estimated Orientation ϕ	121
5.9	Integrated Gyro Output	121
5.10	Estimated Angular Rate $\dot{\phi}$	122
5.11	Estimated Position (a) x and (b) y	123
5.12	Estimated Shaping State x_{sf}	124
5.13	Gyro Innovation Sequence	124
5.14	Gyro Innovation Sequence with 1σ and 2σ Bounds	125
5.15	Laser Innovation Sequence with 2σ Bound	125
5.16	Position Standard Deviation (a) x and (b) y	126
5.17	(a) Orientation Standard Deviation and (b) Angular Rate Standard Deviation	127
5.18	Shaping State Standard Deviation	127
5.19	(a) Estimated Orientation ϕ and (b) Estimated Angular Rate $\dot{\phi}$	128

5.20 Gyro Observation of Angular Rate $\dot{\phi}$	129
5.21 Estimated Shaping State x_{sf}	130
5.22 Gyro Innovations	130
5.23 Estimated Orientation ϕ	131
5.24 (a) Laser Innovations and (b) Gyro Innovations	132
5.25 Fault Plot - Percentage of Simulated Laser Fault Transmitted to Simulated Laser Innovations	133
5.26 Fault Plot - Percentage of Simulated Laser Fault Transmitted to Real Laser Innovations	134
5.27 Fault Plot - Percentage of Simulated Laser Fault Transmitted to Gyro Innovations	134
5.28 (a) Estimated Orientation ϕ and (b) Gyro Innovations	135
5.29 (a) Simulated Laser Innovations and (b) Real Laser Innovations	136
5.30 The LHD Showing Sensor Layout	137
5.31 The Data Logging System	139
5.32 The Trials Area	139
5.33 The Proposed Navigation System Architecture	140
5.34 Estimated LHD Path - Using ICP Approach	142
5.35 LHD Kinematic Geometry	142
5.36 LHD Kinematic Geometry - Including Slip Angles	144
5.37 The IMU and Laser used in the field trials	148
5.38 Estimated LHD Path - stars represent retroreflective markers	150
5.39 Position standard deviation during run (x and y)	151
5.40 Vehicle Velocity and Steer Angle	151
5.41 Estimated Wheel Radius	152
5.42 Estimated Slip Angle Alpha	152
5.43 Estimated Slip Angle Beta	153
5.44 Gyro Innovation Sequence with 1σ and 2σ bounds	153

List of Tables

2.1 Parameter Minimum and Maximum Values	59
5.1 The Noise Parameters	117

Chapter 1

Introduction

1.1 The Objectives of this Thesis

This thesis seeks to provide an autonomous navigation systems engineer with the tools and insights needed to design a high integrity navigation system. The term high integrity is used to describe a system that is robust to failure, and at the very least can be guaranteed upon failure to attain a ‘safe’ state.

To achieve this goal, the following three questions are addressed:

- What constitutes a navigation failure?
- How can navigation failures be detected?
- How does one design a high integrity navigation system?

By designing for high integrity, the designer is acknowledging the fact that at some point in time, the system *will* fail. The time between failures may be maximised by choosing high reliability components, but nevertheless the components (whether they be sensors, actuators, mechanics, computers or even mathematical models) will at some time fail in potentially unpredictable ways.

A navigation system is said to have high integrity if a guarantee is provided that while individual components may fail, the system as a whole will never fail in a catastrophic manner, and therefore will not require any human intervention. Without the guarantee of

integrity, autonomous navigation systems will be difficult, if not impossible, to deploy in real-world applications.

This thesis aims to provide this guarantee by determining the necessary conditions that are required for a navigation system to be considered of high integrity.

1.2 Navigation

In this thesis, the term navigation refers to the position determination (localisation) of Autonomous Guided Vehicles (AGVs). This is arguably the most important competency that an AGV must possess to obtain truly autonomous, useful operation.

AGVs are typically used in commercial applications to eliminate human operators from the tedious task of moving payload from one point to another[29, 52, 96]. In order to achieve this task, the AGV must be able to repeatedly and accurately position itself within its environment. This is not a trivial task.

It is the job of the navigation system to determine, in either absolute or relative terms, the position of the vehicle. This position determination must also be accurate enough to achieve the desired task. For example, a factory AGV designed to move palettes between different workspaces must be able to position itself such that the palette can be retrieved at the workspace. This type of task typically requires centimetre or sub-centimetre accuracy.

To achieve this, the AGV must continuously monitor its motion to ensure that any errors inherent in the system do not grow without bound. This is analogous to using feedback in a control system.

In order to monitor AGV motion, sensors must be used. These sensors must be able to provide enough information such that vehicle pose (position and orientation) can be estimated.

The choice of sensors is usually application specific, for example, a high speed dynamic system will typically use inertial sensors to measure accelerations acting on the vehicle. In landmark based navigation, lasers, radars and sonars can be used to locate objects in the environment (either natural or artificial) and outdoors, sensors such as GPS¹ can be used.

¹Global Positioning System

A simple example of a localisation methodology is triangulation. Commercially available triangulation systems use a laser scanner with high rotation rate. The environment is populated with precisely surveyed beacons (retroreflective tapes), which the laser scans for. When the angle to three beacons is known, triangulation is used to determine the vehicle's position and orientation. The disadvantage of this methodology is that the vehicle must be assumed to be stationary during the time it takes for three beacons to be scanned. Clearly, the higher the speed of the vehicle, the less this assumption holds.

The manufacturer NDC[71] sells a system which uses triangulation in the way described above. MacLeod Technology's COANIC system[59] differs only in that it uses *active* beacons. When a beacon is scanned by the onboard laser, it responds by transmitting an identification code and time stamp.

There are other simple localisation systems that rely only on a single sensor, but do not require the user to modify the local environment. The most widely used system of this type is GPS, which calculates global position based on information received from satellites (active beacons). The accuracy of these systems is typically in the order of $100m$, for conventional civilian systems, down to sub-meter accuracy for differential carrier phase GPS.

GPS systems, however, do not tend to be robust[14]. The GPS receiver must be receiving signals from at least four satellites to produce a position estimate. This means that the receiver must be in full view of a large portion of the sky at all times. Another problem seen in GPS systems is multi-path reflection, where the signal from a satellite is received after having bounced off an object, rather than being received directly.

Single sensor systems all suffer from the same problem. In the absence of other information it is impossible to tell if the sensor is performing correctly. There is no choice but to trust the information supplied by that sensor. In [75], it is argued that navigation sensors fall into two very broad categories, those that measure internal vehicle parameters, and those that measure parameters external to the vehicle, and that for robustness, multi-sensor navigation systems must comprise sensors of both types.

If multi-sensor navigation systems are to be used for reliability, a method for *fusing* the sensor information must be chosen. This is generally achieved by treating localisation as an estimation problem.

1.3 The Kalman Filter

A consistent methodology for estimating position from navigation sensors is through the use of Kalman filtering and, for nonlinear systems, through the use of the extended Kalman filter.

The Kalman filter is a linear statistical algorithm used to recursively estimate the states of interest. The states of interest will usually consist of the vehicle pose and other relevant vehicle parameters. In map building, the state vector can be augmented with feature positions, so that they too may be estimated.

To aid in the estimation of the states, the Kalman filter requires that there be two mathematical models: the process model and the observation model. The process model describes how the states evolve over time, together with an estimate of the errors committed by the system. The observation model explicitly describes the information supplied by a sensor as a function of the state, together with a model of measurement noise.

For a linear system subject to Gaussian, uncorrelated, zero mean measurement and process noises, the Kalman filter is the optimal minimum mean squared error estimator. The effect of these assumptions on navigation system behaviour will be examined in detail in the chapters that follow.

An additional benefit of using the Kalman filter estimator is that it keeps track of the uncertainties present in the system via the computation of a covariance matrix. This is important in many applications where it is desired to know how well (or how poorly) a system is performing.

The Kalman filter has been used successfully in many multi-sensor AGV applications, including transport vehicles[28], road vehicles[48, 51] and indoor mobile robotics[17].

Due to its recursive nature, the Kalman filter is ideally suited to representation in the time domain. The time domain, however, is not well suited to describing the physical operation of the Kalman filter.

The Kalman filter (as will be shown in Chapter 2) is in fact a low pass filter. As such its operation is best understood in the frequency domain with the aid of tools such as Bode diagrams. Bode diagrams have been used extensively in fields of control and analog filter design for many years. They provide a designer with a method for analysing the transfer of

information from a systems inputs (in this case sensors) to the outputs (the estimated state) in terms of frequency response.

In essence, frequency domain techniques are useful to Kalman filter analysis as they allow a more compact description of filter operation. Rather than sensors and processes being characterised by their time propagation, they are instead characterised by their information contribution at any particular frequency.

One advantage of frequency domain analysis is that it allows the designer to determine for a given set of sensors and process models which combination will perform the best, without the expense of field testing[21].

In this thesis, frequency domain techniques are used extensively in the analysis and design of navigation systems.

1.4 Navigation Failure

Navigation failure may be thought of as the inability of the navigation system to accurately determine the pose of the AGV. Fault detection, in its most general form, will be defined in this thesis as *the ability to detect navigation failure*.

There is an increasing demand for AGVs which can operate at relatively high speed, in unstructured environments such as ports[28], mines[85], and even highways[80]. These applications typically require the automation of large high-payload vehicles. As the size, speed and mass of the AGV increases, so does the risk associated with its operation. A larger, faster moving vehicle can potentially do more damage than a smaller, slower moving one.

The potential of navigation systems to cause harm cannot be understated. As well as the potential for loss of life and damage to equipment, a navigation system failure may also have severe *political* implications. This is due to the perception that an autonomous navigation system will be safer than its manned counterpart. Any catastrophic failure of the navigation system will inevitably tarnish this perception, making widespread acceptance of AGVs increasingly difficult.

To the best of the authors knowledge, in Australia at least, there are no guidelines which define the specifications an AGV system must conform to in order to be considered safe. It

is therefore the responsibility of the navigation system designer to anticipate legislation of this sort. The designer *must* be able to provide a guarantee that any fault will not cause a catastrophic failure of the navigation system.

1.4.1 Reliability

Most commercial, military and aerospace navigation systems attempt to minimise the risk associated with navigation systems failing by providing *reliable* navigation systems. The reliability of a navigation system is measured by its Mean Time Between Failures (MTBF). In practise, a high reliability navigation system is constructed by using high reliability parts. This means that if there is no redundancy, the navigation system will only be as reliable as its most unreliable component.

An interesting survey of reliability, specifically in the area of robotics, is given in [15]. This paper discusses NASA and military standards relevant to the design of reliable systems. Methods are presented which allow the degree of reliabilty to be calculated via the use of fault trees[4, 93].

Regardless of the *degree* of reliability, there is still potential for navigation failure. Reliable systems only postpone the inevitable, they do not prevent it. In itself, reliability is highly desirable, but reliability does not provide a guarantee of system performance when a fault does occur.

1.4.2 Integrity

To overcome the reliability problem, a navigation system should be designed to have high integrity, that is, to be robust to system failure. Inevitably, this requires that the navigation system have some form of fault detection. Furthermore, there must be a guarantee that all faults will be detectable.

It is an interesting point that a system designed using high reliabilty components may not itself have high integrity. However, a system constructed of relatively low reliability components may indeed be designed to have high integrity. Ideally, a navigation system should have both high integrity and high reliability.

This thesis seeks to quantify the essential elements that must be present in a navigation system for it to be considered to have high integrity.

1.5 Contributions of this Thesis

The main contributions of this thesis, in order of importance, are:

- The determination of the causes of navigation failure.
- The unification of sensor suite design with fault detection such that the resulting navigation system may be thought of as having high integrity with respect to navigation failure.
- The derivation of a metric which describes the detectability of faults within the Kalman filter framework, and the development of methods to guarantee fault detectability.
- The development of automated design techniques for selecting sensor suites, such that certain design requirements are satisfied.
- The demonstration of navigation failure in a real system, and the design and implementation of a navigation system for a large autonomous mining vehicle.

Together, these contributions provide the necessary framework to allow a navigation system to be designed such that it meets the design specifications, and that all navigation failures will be detectable, thus providing the assurances needed for high integrity navigation.

1.6 Thesis Structure

This thesis does not have one central literature review. Rather, the literature relevant to a particular topic is cited where that topic is presented.

In **Chapter 2**, the Kalman filter is proposed as the navigation algorithm most useful to the automation of large autonomous outdoor vehicles. The Kalman filter is then examined in the frequency domain. Frequency domain analysis provides a simple understanding of

how sensor information is used by the filter to estimate the vehicle states. Alternate filter structures are discussed which explicitly take into account the frequency response of the sensors used.

The Kalman filter analysis is extended to include the concept of navigation failure. The possible sources of failure are identified, and modelling techniques are introduced which compensate for many ‘common’ navigation failures.

Lastly, the nonlinear and discrete time Kalman filters are discussed. The nonlinear or extended Kalman filter is shown to have a bounded frequency response, such that linear approximations afford an insight into nonlinear behaviour.

Chapter 3 begins by examining methods by which unmodelled faults may be detected. Innovations techniques are chosen as the only fault detection techniques that satisfy the criteria of being non-model based and retaining filter optimality.

Process faults are examined, and a measure of process fault detectability is introduced. This detectability metric is used to propose a method by which all process faults are rendered detectable. Similarly, sensor faults are analysed, and the concept of frequency domain redundancy is introduced as a method of guaranteeing sensor fault detectability.

It is shown that these concepts, developed for linear systems, apply equally well to bounded nonlinear and time varying linear systems.

Finally, a recursive time domain whiteness test is introduced which can be used to test for the presence of process and sensor faults.

Chapter 4 discusses the issue of designing high integrity navigation systems. From an initial design specification, it is shown how sensor suites should be chosen, and a method is given to aid in the automation of this design process. Once the navigation system has been designed, the arguments presented in Chapter 3 indicate that frequency domain redundancy should be employed to ensure fault detectability.

Architectures other than the direct Kalman filter are discussed, motivated by the need to distribute the navigation system over multiple computers, to take into account hardware and software failures. Theory is presented which shows that the detectability of faults in a navigation system implemented in distributed form is identical to that of the conventional navigation system.

In **Chapter 5**, experimental results are presented. The first set of results, from a simple system consisting of a laser and a gyroscope, is used to demonstrate and verify the concepts presented in Chapter 3. The second experiment describes the initial results of an ongoing research project into the design and implementation of a high integrity navigation system for an underground mining vehicle. This experiment emphasises the modelling techniques described in Chapter 2, and shows how the key concepts developed in this thesis may be applied in a real application.

Finally, **Chapter 6** provides conclusions and suggestions for further work.

Chapter 2

Navigation Systems and Their Failure Modes

2.1 Introduction

Perhaps the most important and difficult task in the guidance of an autonomous vehicle is localisation. This chapter introduces the Kalman filter as the data fusion algorithm used to estimate an AGV's position from uncertain sensor measurements. The continuous time Kalman filter is presented in Section 2.2 together with its frequency domain counterpart. The frequency domain is shown to be a compact, intuitive representation of filter behaviour. Alternate filter structures are examined which explicitly take into account the frequency domain characteristics of different sensors.

In Section 2.3, the Kalman filter is employed to determine the sources and nature of errors which can potentially lead to navigation failure. The possible sources of error are found to be the process model and the observation model. In Section 2.4, techniques for modelling sensor error are discussed. Section 2.5 introduces the extended Kalman filter as a solution to the nonlinear navigation problem. It is shown that the extended Kalman filter uses the same theoretical framework as the time varying linear Kalman filter. Section 2.6 introduces the discrete time Kalman filter.

Section 2.7 describes a simple method for determining the sources of error in process models and how these error sources may be included in the model as states to be estimated. In

Section 2.8, a brief survey of navigation sensors is given, and in Section 2.9, a detailed analysis of the effects of sensor error modelling on Kalman filter performance is presented.

Finally, in Section 2.10 frequency domain analysis techniques are extended to nonlinear systems, and it is shown that for the most common types of processes encountered in AGV navigation, the frequency response of these systems is bounded.

2.2 Navigation

In AGV systems, Navigation is the process whereby vehicle pose is computed based upon sensor measurements. In this thesis, the Kalman filter algorithm is adopted in order to solve the localisation problem.

The Kalman filter is a powerful tool for multi-sensor estimation problems. By posing localisation as a probabilistic estimation problem, the Kalman filter allows an estimate of state and covariance to be obtained from imprecise sensors measurements and imperfect vehicle models.

2.2.1 The Continuous Time Kalman Filter

The physical world is best described in continuous time. Therefore most physical processes which occur within the world may be modelled and analysed in continuous time form. In the context of navigation systems for AGVs there are two main models of interest, the vehicle model and the observation model(s). These are almost always developed as continuous time equations.

The Kalman filter has a continuous time form, and if continuous time process (or vehicle) and observation models are known, is the optimum minimum mean squared error estimator for linear systems. This section summarises the continuous time Kalman filter and its operation.

Consider a linear process represented in state space form by the equation

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t) \quad (2.1)$$

where $\mathbf{x}(t)$ is the state vector containing the states of interest, $\mathbf{F}(t)$ is a matrix mapping the state $\mathbf{x}(t)$ to the state velocity $\dot{\mathbf{x}}(t)$, $\mathbf{B}(t)$ is a matrix mapping the control inputs $\mathbf{u}(t)$ to the state velocity, and $\mathbf{G}(t)$ is a matrix mapping the process noise vector $\mathbf{w}(t)$ to the state velocity. The process noise $\mathbf{w}(t)$ is assumed Gaussian with zero mean and strength $\mathbf{Q}(t)$:

$$\begin{aligned} \mathbb{E}[\mathbf{w}(t)] &= 0, \\ \mathbb{E}[\mathbf{w}(t)\mathbf{w}^T(t+\tau)] &= \mathbf{Q}(t)\delta(\tau) \end{aligned} \quad (2.2)$$

where $\delta(\tau)$ is the Dirac delta function.

The linear observation model is assumed to be represented by

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{v}(t) \quad (2.3)$$

where $\mathbf{H}(t)$ is a matrix mapping the state vector to observation space and $\mathbf{D}(t)$ is a matrix mapping the observation noise $\mathbf{v}(t)$ to observation space. The process noise $\mathbf{v}(t)$ is assumed Gaussian with zero mean and strength $\mathbf{R}(t)$:

$$\begin{aligned} \mathbb{E}[\mathbf{v}(t)] &= 0, \\ \mathbb{E}[\mathbf{v}(t)\mathbf{v}^T(t+\tau)] &= \mathbf{R}(t)\delta(\tau) \end{aligned} \quad (2.4)$$

Further, it is also assumed that the process and observation noises are uncorrelated

$$\mathbb{E}[\mathbf{v}(t)\mathbf{w}^T(\tau)] = 0, \quad \forall t, \tau \quad (2.5)$$

Following from these definitions, the continuous time Kalman filter may be derived as[63]

$$\hat{\mathbf{x}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{K}(t)[\mathbf{z}(t) - \mathbf{H}(t)\hat{\mathbf{x}}(t)] \quad (2.6)$$

where $\hat{\mathbf{x}}(t)$ is the estimate of the state vector $\mathbf{x}(t)$, and $\mathbf{K}(t)$ is the Kalman gain matrix.

The term $[\mathbf{z}(t) - \mathbf{H}(t)\hat{\mathbf{x}}(t)]$ is known as the filter innovations, defined as the difference between the observation $\mathbf{z}(t)$ and the predicted observation $\mathbf{H}(t)\hat{\mathbf{x}}(t)$. The innovations are commonly represented by the symbol $\nu(t)$.

The covariance update equation, which is a measure of the uncertainty present in the system,

is given by the solution of the continuous time matrix Riccati equation:

$$\dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t) - \mathbf{K}(t)\mathbf{H}(t)\mathbf{P}(t) \quad (2.7)$$

where $\mathbf{P}(t)$ is the covariance matrix and $\mathbf{K}(t)$ is given by

$$\mathbf{K}(t) = \mathbf{P}(t)\mathbf{H}^T(t)\mathbf{R}^{-1}(t) \quad (2.8)$$

Equations 2.6 and 2.7 are differential equations and in general are solved by integrating forward from some initial conditions

$$\hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0, \quad \hat{\mathbf{P}}(t_0) = \hat{\mathbf{P}}_0 \quad (2.9)$$

where $\hat{\mathbf{x}}_0$ is the initial value of the state vector and $\hat{\mathbf{P}}_0$ is the initial value for the covariance matrix which should be large enough to account for the uncertainty in the initial state estimate.

In terms of vehicle navigation, the Kalman filter may be interpreted as follows. The state vector \mathbf{x} will always contain the position and orientation of the vehicle together with other relevant parameters such as slip angles or beacon positions. These are the quantities that will be estimated by the Kalman filter.

The process model \mathbf{F} will usually contain a model of the vehicle, together with models of the other parameters in the state vector. The vehicle model can be derived by any means, i.e. kinematic or dynamic. This model describes how the vehicle's motion changes with time, and is used by the filter to predict the state when no observations are available.

The observation model \mathbf{H} is used to obtain predicted observations with respect to the vehicle state.

The covariance matrix \mathbf{P} provides a representation of the uncertainty in the state estimate due to the fact that both the sensor observations, and process model predictions are corrupted by noise. The Kalman filter minimises this uncertainty in a minimum mean squared error sense.

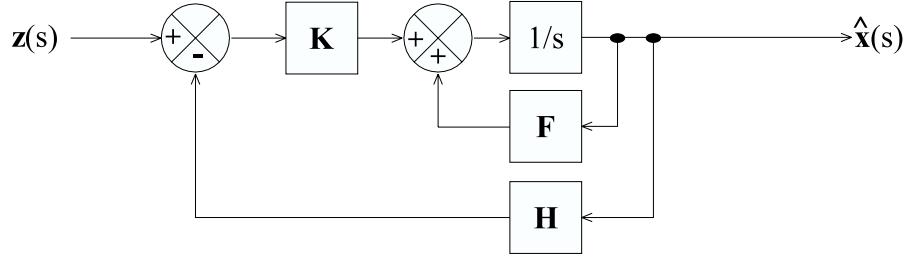


Figure 2.1: Block Diagram of Continuous Time Kalman Filter

2.2.2 Navigation in the Frequency Domain

The Kalman filter is a data fusion algorithm that is traditionally derived and implemented in the time domain, as shown in Section 2.2.1. However much insight into filter operation may be gleaned by examining the filter behaviour in the frequency domain.

The frequency response of the Kalman filter may be readily analysed using classical control techniques. However, in order to compute a closed form solution for the filter transfer function, the system matrices must be assumed to be time invariant, with control inputs assumed perfectly known. This assumption gives rise to system equations of the form

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{w}(t) \quad (2.10)$$

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{D}\mathbf{v}(t) \quad (2.11)$$

The only further assumption that must be made in order to compute the Kalman filter transfer function is that the filter has reached steady state. In fact, it can be shown that the Kalman filter reaches steady state conditions rapidly after initialisation[63], and the gain \mathbf{K} can therefore be considered a constant. Commonly, in practical systems, the steady state filter gains are used to reduce computational and memory requirements[6, 83].

With constant system and gain matrices, the continuous time Kalman filter update equations given in Equation 2.6 may be rewritten by taking a Laplace transform with zero initial conditions as

$$s\hat{\mathbf{x}}(s) = \mathbf{F}\hat{\mathbf{x}}(s) + \mathbf{K}[\mathbf{z}(s) - \mathbf{H}\hat{\mathbf{x}}(s)] \quad (2.12)$$

which may be represented by the block diagram shown in Figure 2.1.

Rearranging Equation 2.12 the transfer function between the state vector and the measure-

ment vector may be expressed as

$$\begin{aligned}\frac{\hat{\mathbf{x}}(s)}{\mathbf{z}(s)} &= [s\mathbf{I} - \mathbf{F} + \mathbf{K}\mathbf{H}]^{-1}\mathbf{K} \\ &= \mathbf{T}(s)\end{aligned}\quad (2.13)$$

As mentioned previously, a closed form solution for the transfer function of the Kalman filter for time varying linear systems and nonlinear systems cannot in general be computed. These systems can only be studied on a case by case basis. This idea is expanded and investigated in Section 2.10.

Taking the limit of the Kalman filter transfer function as s approaches infinity

$$\lim_{s \rightarrow \infty} \mathbf{T}(s) = \lim_{s \rightarrow \infty} [s\mathbf{I} - \mathbf{F} + \mathbf{K}\mathbf{H}]^{-1}\mathbf{K} = 0 \quad (2.14)$$

shows that the frequency response of the filter is always low pass since the transfer approaches zero. Also, as the poles may be found by taking the determinant of the matrix $[s\mathbf{I} - \mathbf{F} + \mathbf{K}\mathbf{H}]$ and are the same for all of the states, the states will therefore have an identical cutoff frequency ω_c above which all the sensor inputs are attenuated.

No such generalisations can be made about the low frequency behaviour of the Kalman filter. The Kalman filter transfer function can therefore be written in the form

$$\mathbf{T}(s) = \mathbf{T}_{kf}(s) \times \frac{1}{(s + \omega_c)^n} \quad (2.15)$$

which is a low pass filter with cutoff frequency ω_c where ω_c is at a higher frequency than any zeros. The transfer function $\mathbf{T}_{kf}(s)$ is determined by the filter equations.

2.2.3 A Simple Example

The low pass nature of the Kalman filter can best be demonstrated by a simple example.

Example 2.2.1 Consider a system that can be adequately described by a constant velocity

model where measurements are made of the position as

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \quad (2.16)$$

$$z(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + v(t) \quad (2.17)$$

where $[x(t), \dot{x}(t)]^T$ is the state vector (position and velocity), $z(t)$ the continuous time position observation, and $w(t)$ and $v(t)$ are the process and observation noises respectively. Equations 2.16 and 2.17 are of the form given in Equations 2.1 and 2.3.

The noise sources $w(t)$ and $v(t)$ are uncorrelated, Gaussian and zero mean with covariance

$$\mathbb{E}[w(t)w^T(t+\tau)] = Q\delta(\tau) \quad (2.18)$$

$$\mathbb{E}[v(t)v^T(t+\tau)] = R\delta(\tau) \quad (2.19)$$

The continuous time Kalman filter for this system is given by Equation 2.6 as

$$\begin{aligned} \begin{bmatrix} \hat{x}(t) \\ \hat{\dot{x}}(t) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}(t) \\ \hat{\dot{x}}(t) \end{bmatrix} + \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{12}(t) & P_{22}(t) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{R} [z(t) - \hat{x}(t)] \\ &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}(t) \\ \hat{\dot{x}}(t) \end{bmatrix} + \begin{bmatrix} \frac{P_{11}(t)}{R} \\ \frac{P_{12}(t)}{R} \end{bmatrix} [z(t) - \hat{x}(t)] \end{aligned} \quad (2.20)$$

where the covariance $\mathbf{P}(t)$ satisfies the algebraic Riccati equation given in Equation 2.7 as (after performing all matrix multiplications)

$$\begin{aligned} \begin{bmatrix} \dot{P}_{11}(t) & \dot{P}_{12}(t) \\ \dot{P}_{12}(t) & \dot{P}_{22}(t) \end{bmatrix} &= \begin{bmatrix} P_{12}(t) & P_{22}(t) \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} P_{12}(t) & 0 \\ P_{22}(t) & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & Q \end{bmatrix} \\ &- \begin{bmatrix} \frac{P_{11}^2(t)}{R} & \frac{P_{11}(t)P_{12}(t)}{R} \\ \frac{P_{11}(t)P_{12}(t)}{R} & \frac{P_{22}^2(t)}{R} \end{bmatrix} \end{aligned} \quad (2.21)$$

which may be solved for the steady state case by substituting $\dot{\mathbf{P}}(t) = 0$, yielding

$$\mathbf{P} = \begin{bmatrix} \sqrt{2}Q^{\frac{1}{4}}R^{\frac{3}{4}} & \sqrt{QR} \\ \sqrt{QR} & \sqrt{2}Q^{\frac{3}{4}}R^{\frac{1}{4}} \end{bmatrix} \quad (2.22)$$

allowing the steady state Kalman gain to be calculated from Equation 2.8 as

$$\mathbf{K} = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} \frac{P_{11}(t)}{R} \\ \frac{P_{12}(t)}{R} \end{bmatrix} = \begin{bmatrix} \sqrt{2}Q^{\frac{1}{4}}R^{-\frac{1}{4}} \\ Q^{\frac{1}{2}}R^{-\frac{1}{2}} \end{bmatrix} \quad (2.23)$$

The filter transfer function may be calculated from Equation 2.13 as

$$\frac{\hat{\mathbf{x}}(s)}{\mathbf{z}(s)} = \begin{bmatrix} \frac{sK_1+K_2}{s^2+K_1s+K_2} \\ \frac{K_2s}{s^2+K_1s+K_2} \end{bmatrix} \quad (2.24)$$

showing that the filter behaves as a second order system, which may be represented in the form

$$\frac{\hat{\mathbf{x}}(s)}{\mathbf{z}(s)} = \begin{bmatrix} \frac{\sqrt{2}\omega_n s + \omega_n^2}{s^2 + \sqrt{2}\omega_n s + \omega_n^2} \\ \frac{\omega_n^2 s}{s^2 + \sqrt{2}\omega_n s + \omega_n^2} \end{bmatrix} \quad (2.25)$$

where

$$\begin{aligned} K_1 &= \sqrt{2}\omega_n = \sqrt{2}Q^{\frac{1}{4}}R^{-\frac{1}{4}} \\ K_2 &= \omega_n^2 = Q^{\frac{1}{2}}R^{-\frac{1}{2}} \end{aligned} \quad (2.26)$$

The Kalman filter for this system is therefore a second order system with natural frequency $\omega_n = Q^{\frac{1}{4}}R^{-\frac{1}{4}}$ and damping ratio $\frac{1}{\sqrt{2}}$.

The cutoff frequency is completely determined by the ratio of process to observation noise strengths $\frac{Q}{R}$.

The transfer function from the position observation to the position state may therefore be represented by the Bode magnitude plot shown in Figure 2.2. This system is characterised as having a zero at $\frac{\omega_n}{\sqrt{2}}$ and cutoff frequency at ω_n .

The Bode magnitude plot for the transfer function from the position sensor to the estimated velocity is shown in Figure 2.3. This transfer is characterised by a zero at 0 and, again a cutoff frequency of ω_n . It is important not to misinterpret the bandpass nature of the transfer from the position sensor to the velocity state. A small change in position does indeed produce only a small change in velocity, as indicated by the Bode plot attenuating low frequency position measurements. The main reason for the bandpass nature of this

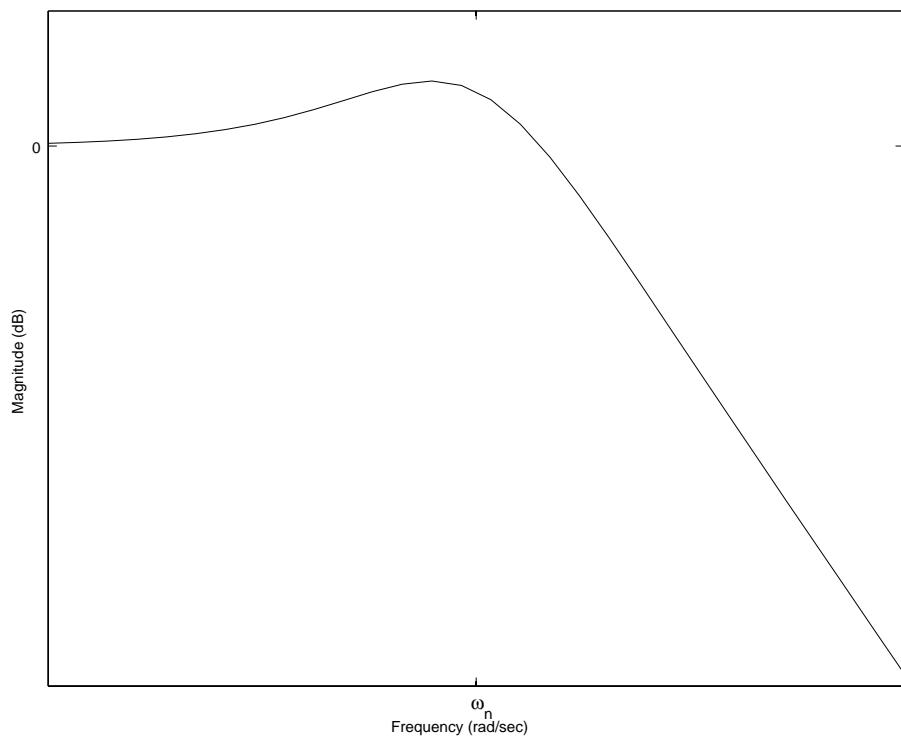


Figure 2.2: Bode Magnitude Plot - Transfer from Position Sensor to Position State

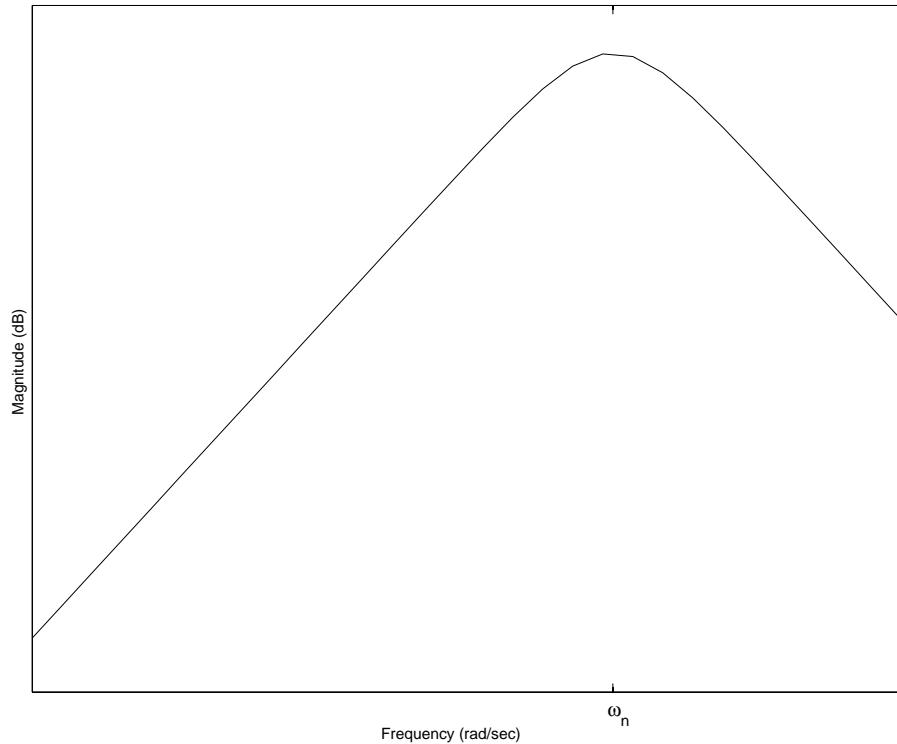


Figure 2.3: Bode Magnitude Plot - Transfer from Position Sensor to Velocity State

transfer function, however, is that the Kalman filter contains an internal model of the system and thus differentiates the estimated position to obtain velocity information.

This example has shown how the Kalman filter may be interpreted in the frequency domain using classical control techniques. In general, though, an algebraic solution to the Riccati equation is *not* possible, and numerical techniques must be used. This will be further expanded in Chapter 4.

2.2.4 Alternate Filter Structures

As shown in the previous section, the Kalman filter is inherently low pass. This low pass nature is due to the sensor signal being band limited in broad band noise. The filter essentially ‘chooses’ a cutoff frequency above which the sensor input is considered to be entirely noise. This frequency domain behaviour may be exploited with appropriately chosen filter structures.

Figure 2.4 shows the structure of the direct Kalman filter, by far the most widely used form.

As can be seen in the figure, the observations are essentially low pass filtered to provide the state estimate. This structure, however, does not take explicit advantage of the frequency characteristics of the sensors employed. High frequency information may be lost. It is also possible that the sampling rates required to track high frequency manouevres may place an unreasonable computational burden on the system.

To overcome these problems, a filter structure known as the indirect feedforward Kalman filter was developed[42, 63]. In this structure (shown in Figure 2.5), the sensor supplying the high frequency information (labelled observation 1) is passed directly to the output. Periodically, though, the high frequency sensor is sampled and differenced with a sensor supplying comparable information although at low frequency (observation 2). This difference (or error) is then optimally estimated by the Kalman filter to account for observation noise, and then fed forward to correct the signal from the high frequency sensor.

In this structure, the transfer function from observation 2 to the state estimate is low pass, however the transfer function from observation 1 to the output is the complement of this (high pass). Therefore, this filter structure exploits the frequency characteristics of different sensors. Typically, rate sensors such as accelerometers and gyros may be characterised as being good high frequency sensors, and indeed this filter structure is commonly used as the basis for Inertial Navigation Systems (INS).

The main assumption that is made when using this filter structure is that the dynamics of the high frequency sensor noise (which may not be white) vary slowly with respect to the sample time of the filter. It should also be noted that an advantage of this form of the Kalman filter is that if the Kalman filter should fail (perhaps due to a computer power down), there will still be an estimate of the state provided by the high frequency sensor, though never corrected.

One problem with the feedforward version of the indirect Kalman filter is that the error between the high frequency sensor and the low frequency sensor may potentially grow without bound. This violates a basic assumption of the Kalman filter: that the errors in the system must be small. To alleviate this problem, the indirect feedback Kalman filter (shown in Figure 2.6) was developed[63]. In this structure, the estimated error is fed back directly to the sensors. In this way, the sensor error is always kept small.

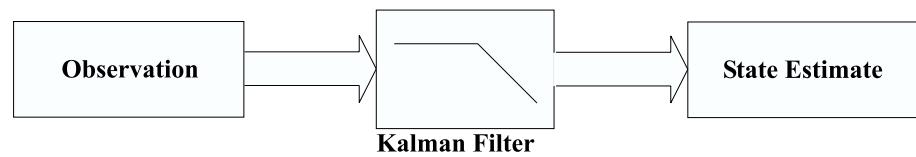


Figure 2.4: Direct Kalman Filter

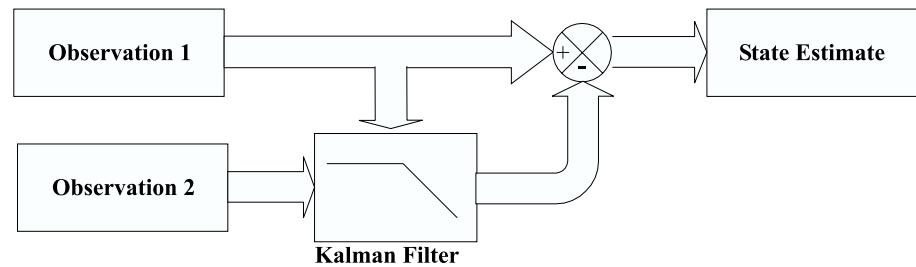


Figure 2.5: Indirect Feedforward Kalman Filter

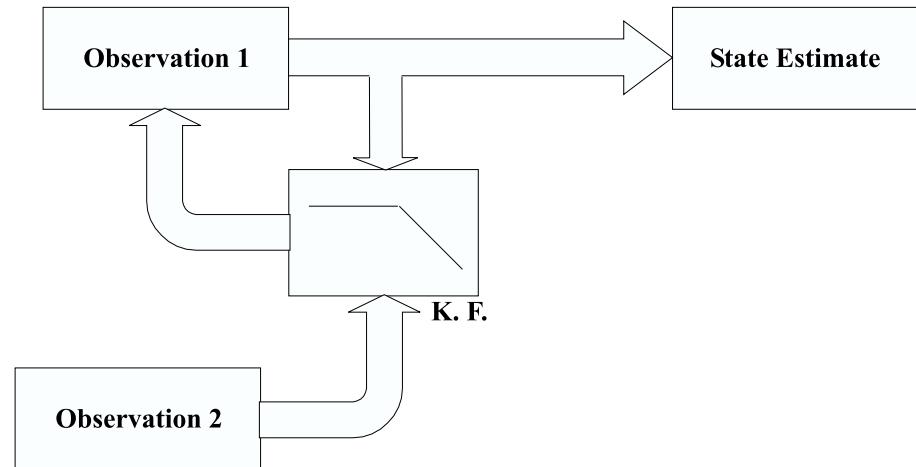


Figure 2.6: Indirect Feedback Kalman Filter

2.3 The Nature of Failures in Navigation Systems

The preceding sections have discussed the continuous time Kalman filter and its frequency domain interpretation. This gives an understanding of the correct operation of Kalman filter based navigation systems in terms of frequency domain behaviour. To be able to design a high integrity navigation system, however, it is not only important to understand the way in which the Kalman filter operates, but also to understand what can cause it to fail.

In broad terms, navigation failure may be defined as anything that causes the Kalman filter to produce inconsistent estimates of state.

In terms of vehicle navigation, this means that a navigation failure occurs when the Kalman filter algorithm indicates that the vehicle is at one position, whereas in reality it is at another, and the difference between the estimated state and the true state is larger than the estimated error.

Although this definition of navigation failure may seem obvious, the sources of navigation failure are not. To determine the causes of navigation failure, the assumptions made in the Kalman filter derivation summarised in Section 2.2 must be re-examined. The assumptions made were as follows;

- The process noise is white
- The observation noise is white
- The system is linear and of stated order

The first two assumptions may be voided, and thus a failure occur, if there is significant modelling error, or if indeed, the noise processes are non-white.

No model, however, can ever completely describe the real world. As a designer one must strive to capture in the form of mathematical models what seem to be the essential elements of real world behaviour. Almost inevitably this involves neglecting dynamics that are considered to have a negligible effect on system behaviour for the benefit of model simplicity.

Typically, mathematical models of real world systems are an attempt to capture the essentials of system behaviour. In frequency terms, this generally means creating a model which in the low frequency ranges is accurate, but which may neglect some high frequency dynamics.

In terms of the Kalman filter, the effect of using process and/or observation models which do not accurately reflect the true process and true observations is to cause filter divergence. This is for two reasons. The first, is that as the process model \mathbf{F} and the observation model \mathbf{H} are used essentially for prediction, incorrect models will cause the predictions to be biased, leading to biased estimates of state. The second reason is that the Kalman gain, which is a function of the process model, the observation model and the process and observation noises (as detailed in Section 2.2), will be computed incorrectly. Therefore incorrect weight will be placed on new measurements.

The third assumption made in the derivation of the Kalman filter states that any nonlinearities present in the system will also cause the Kalman filter to fail.

The following sections describe the mathematical tools used to model these potential faults, including methods used to deal with model error, non-white measurement and process noise, nonlinear models, and the conversion of the system to the discrete time case.

2.4 Modelling Sensor Faults

In the real world, not all observation models (or process models for that matter) fit neatly into the Kalman filter framework. Some sensors display error characteristics which are clearly non-white. As stated, non-white measurement noise violates the basic assumptions used in the derivation of the Kalman filter.

Example 2.4.1 Consider a gyro, which may be simplistically modelled as a measurement of the rate of change of some angle θ in the presence of additive white noise $v(t)$, as

$$z(t) = \dot{\theta}_t(t) + v(t) \quad (2.27)$$

where the subscript t indicates the true state, and $v(t)$ is a white noise sequence with strength

$$\mathbb{E} [v(t)v^T(t + \tau)] = R\delta(\tau) \quad (2.28)$$

In many applications, the state of interest is $\theta(t)$, which may be estimated by integrating the observation described by Equation 2.27, as

$$\hat{\theta}(t) = \int_0^t [\dot{\theta}_t(\tau) + v(\tau)] d\tau = \theta_t(t) + \int_0^t v(\tau) d\tau \quad (2.29)$$

The error, is therefore not white, in fact it is Brownian motion - the integral of white noise. It can be shown[74] that the estimate of the state $\hat{\theta}$ will drift from the true value at a rate proportional to the square root of the length of the sample and the variance of the observation noise

$$\delta\theta(\tau) = \hat{\theta}(\tau) - \theta_t(\tau) = R\sqrt{\tau} \quad (2.30)$$

The variance of the estimate will also increase with time.

This example shows that if the gyro is used as an observation of the angle θ , then the assumptions made in the derivation of the Kalman filter will be voided as the sensor noise will be non-white (in fact it will be Brownian). This will almost certainly cause the Kalman filter to diverge.

Many other sensors exhibit noise characteristics which are non-white and therefore cannot be directly incorporated into the Kalman filter framework. A short survey of navigation sensors and their error characteristics is given in Section 2.8.

2.4.1 Error Models Via Power Spectral Density Estimation

Many sensors with non-white measurement errors, have errors which are nevertheless quantifiable. In the case where the sensor measurement error has a rational power spectral density function, a shaping filter may be employed to model the non-white noise of the sensor. This section describes how a power spectral density may be computed for a given signal, and in the next section, techniques are given for creating sensor error models for these systems which can be incorporated into the Kalman filter.

A PSD $\Psi(s)$ is a measure of signal power versus frequency, and as such compactly represents the statistics of the sensor error. A white noise source for example has equal power at all frequencies.

The PSD has a time domain analogy - Autocorrelation, which is simply the inverse Fourier transform of the PSD. The detailed formulation of PSDs and autocorrelation functions from sensor data may be found in Appendix A.

Once the statistics of the sensor error have been computed in the form of a PSD, it is necessary to transform this into state space form, enabling a mathematical model of the errors to be developed. The problem is essentially one of fitting a curve that most closely approximates the PSD using the lowest order model possible. This can often be achieved by fitting a curve by hand; however, the technique of spectral estimation [2, 7, 74] provides a firm theoretical framework in order to achieve this goal.

Spectral factorisation (or estimation) allows a state space system to be formulated which, when driven by a white noise source, produces coloured noise with a PSD that is identical to that of the sensor error. Thus, a sensor's coloured noise may be modelled as the output from a linear system driven by a white noise source.

To illustrate these concepts, consider Figures 2.7 to 2.9. Figure 2.7 shows a sensor model where the observation is corrupted by white noise ($v(t)$) only. The term $\mathbf{Hx}(t)$ is used to represent the 'true' observation and $z(t)$ is the output from the sensor. This is the sensor model that is implicitly assumed in the Kalman filter derivation

In Figure 2.8 a sensor model is shown where the observation is corrupted by coloured noise $n(t)$. This model is representative of many real sensors, and if used directly in a Kalman filter will almost certainly cause the filter to diverge.

Finally, Figure 2.9 shows a system identical to that of Figure 2.8 except that the coloured noise source has been replaced by a PSD model driven by white noise. As will be seen in the next section, the PSD model can be used to augment the Kalman filter in a theoretically consistent manner.

2.4.2 Shaping Filters

A shaping filter is a system that models a coloured noise as a linear system driven by white noise, but whose output is the observed measurement error. The Kalman filter can be thus augmented with a shaping filter, as the assumption that all driving noises are white will no longer be violated.

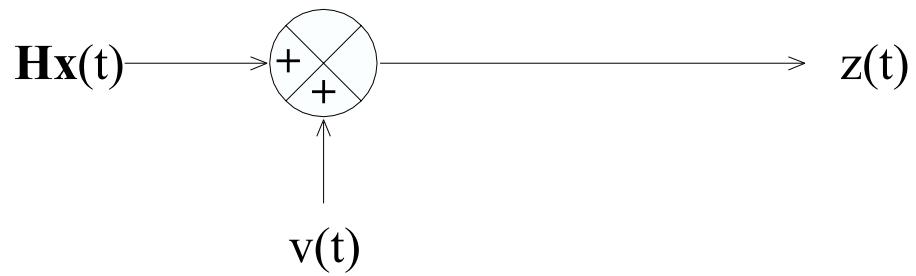


Figure 2.7: Sensor Model - Observation Corrupted by White Noise

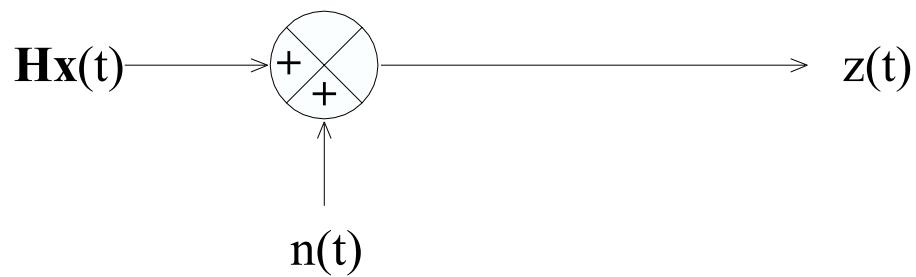


Figure 2.8: Sensor Model - Observation Corrupted by Coloured Noise

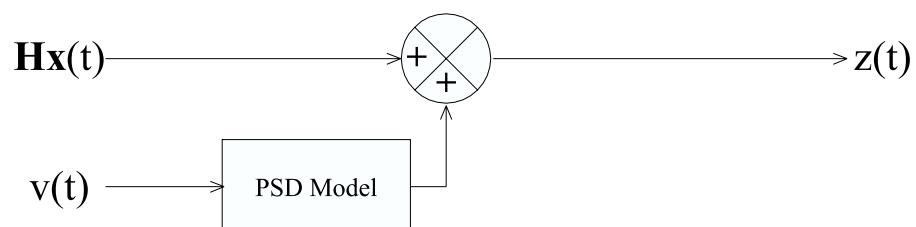


Figure 2.9: Sensor Model - Observation Corrupted by Coloured Noise Generated from a White Noise Source

The system model is the same as that for the unaugmented Kalman filter:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) \quad (2.31)$$

however the observation equation now has a coloured noise component, $\mathbf{n}(t)$, plus additive white noise, $\mathbf{v}(t)$.

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{n}(t) + \mathbf{v}(t) \quad (2.32)$$

The coloured noise $\mathbf{n}(t)$ is assumed to be the output of a linear system (often found as a state space representation of the sensor errors PSD) of the form

$$\begin{aligned} \mathbf{n}(t) &= \mathbf{H}_{sf}(t)\mathbf{x}_{sf}(t) \\ \dot{\mathbf{x}}_{sf}(t) &= \mathbf{F}_{sf}(t)\mathbf{x}_{sf}(t) + \mathbf{G}_{sf}(t)\mathbf{w}_{sf}(t) \end{aligned} \quad (2.33)$$

where the driving noise $\mathbf{w}_{sf}(t)$ is assumed white. The subscript sf denotes shaping filter specific matrices.

The filter may now be augmented with the shaping filter as follows

$$\mathbf{z}(t) = \begin{bmatrix} \mathbf{H}(t) & \mathbf{H}_{sf}(t) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_{sf}(t) \end{bmatrix} + \mathbf{v}(t) \quad (2.34)$$

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\mathbf{x}}_{sf}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{F}(t) & 0 \\ 0 & \mathbf{F}_{sf}(t) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_{sf}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{G}(t) & 0 \\ 0 & \mathbf{G}_{sf}(t) \end{bmatrix} \begin{bmatrix} \mathbf{w}(t) \\ \mathbf{w}_{sf}(t) \end{bmatrix} \quad (2.35)$$

giving system equations which are of the same form as those expressed in Equations 2.1 and 2.3. This augmented system may be seen graphically in Figure 2.10. Note that the *only* driving noises in the system are white sequences.

Figure 2.11 shows some commonly encountered non-white noise sources, together with their respective autocorrelation and power spectral density functions. The graphs are only intended to give an idea of the shape of the different statistics. The true autocorrelations and PSD's are dependent on the strength of the input noise (and in the case of the exponentially time correlated process, the correlation time also). A dashed entry in the figure indicates that the corresponding autocorrelation or PSD cannot be computed.

Similarly to the derivation in Section 2.2.2, the augmented system may be converted to the

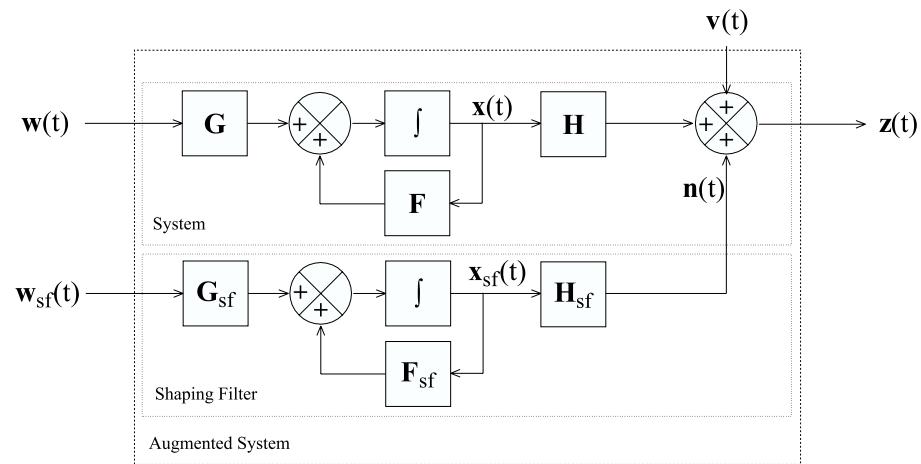


Figure 2.10: Augmented System Model with Observation Noise Shaping Filter

Process	Shaping Filter	Autocorrelation	Power Spectral Density
White Noise	$w(t) = x(t)$		
Random Constant or Bias	$x(t_0)$ 		
Brownian Motion	$w(t) \rightarrow \int \rightarrow x(t)$		
Brownian Motion Plus Random Constant	$x(t_0)$ $w(t) \rightarrow \int \rightarrow x(t)$		
Exponentially Time Correlated Process	$w(t) \rightarrow \text{summing junction} \rightarrow \int \rightarrow x(t)$ $1/T$		

Figure 2.11: Some Shaping Filters and their Corresponding Autocorrelation and PSD Functions

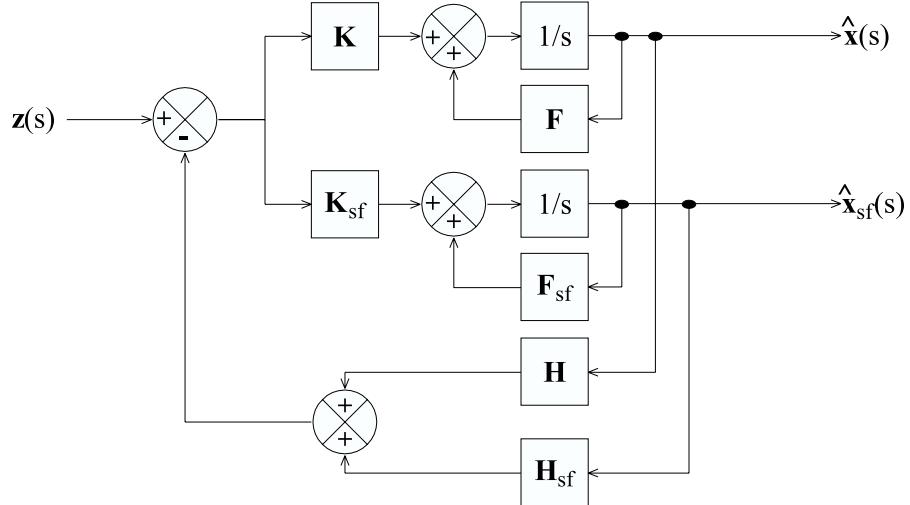


Figure 2.12: Block Diagram of Continuous Time Kalman Filter with Shaping State(s)

frequency domain via Laplace transform to give the frequency domain Kalman filter update equation as

$$s \begin{bmatrix} \mathbf{x}(s) \\ \mathbf{x}_{sf}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{F} & 0 \\ 0 & \mathbf{F}_{sf} \end{bmatrix} \begin{bmatrix} \mathbf{x}(s) \\ \mathbf{x}_{sf}(s) \end{bmatrix} + \begin{bmatrix} \mathbf{K} \\ \mathbf{K}_{sf} \end{bmatrix} \left[\mathbf{z}(s) - \begin{bmatrix} \mathbf{H} & \mathbf{H}_{sf} \end{bmatrix} \begin{bmatrix} \mathbf{x}(s) \\ \mathbf{x}_{sf}(s) \end{bmatrix} \right] \quad (2.36)$$

which is represented graphically by the block diagram shown in Figure 2.12. Comparing this figure to that shown in Figure 2.10, it can be seen that the Kalman filter keeps an internal representation of the system models for both the system states and the shaping states. The filter only couples the two systems through the \mathbf{H} matrices in the feedback path.

It should be noted here that the techniques described for estimating coloured measurement noise are equally applicable to coloured process noise. However, coloured process noise is usually the result of an inaccurate system model, and as will be seen in Section 2.7, this also may be dealt with in a systematic manner.

2.5 The Extended Kalman Filter

Many common processes cannot adequately be expressed using linear models such as those presented previously. In these cases, nonlinear models must be used. The Kalman filter,

however, is explicitly derived assuming that both the process and the observation models are linear. This section examines the extension of the Kalman filter estimator to nonlinear systems.

Many approaches have been used for the nonlinear estimation problem, virtually all of which use approximations of some kind, including both analytical[16, 88] and numerical[37, 40] methods. The most widely used approach, and perhaps the simplest, is the use of the extended Kalman filter (EKF).

In continuous time, the EKF is based on the premise that the dynamic system model is a nonlinear function of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{G}(t)\mathbf{w}(t) \quad (2.37)$$

where $\mathbf{f}(\cdot)$ is a nonlinear function mapping the control inputs $\mathbf{u}(t)$ and state $\mathbf{x}(t)$ to the state velocity $\dot{\mathbf{x}}(t)$.

The observation equation is also a nonlinear function given by

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}(t), t) + \mathbf{v}(t) \quad (2.38)$$

where $\mathbf{h}(\cdot)$ is a nonlinear function mapping the state to observation space.

The matrix $\mathbf{G}(t)$, and the noise sources $\mathbf{v}(t)$ and $\mathbf{w}(t)$ retain their definitions from Section 2.2.

The continuous time EKF for this nonlinear system is given by[38]

$$\hat{\dot{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t), t) + \bar{\mathbf{K}}(t)[\mathbf{z}(t) - \mathbf{h}(\hat{\mathbf{x}}(t), t)] \quad (2.39)$$

with covariance and gain equations as

$$\begin{aligned} \dot{\mathbf{P}}(t) &= \mathbf{F}'(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}'^T(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t) - \bar{\mathbf{K}}(t)\mathbf{H}'(t)\mathbf{P}(t) \\ \bar{\mathbf{K}}(t) &= \mathbf{P}(t)\mathbf{H}'^T(t)\mathbf{R}^{-1}(t) \end{aligned} \quad (2.40)$$

These equations can be seen to be the functional equivalent of the time varying linear Kalman filter equations given in Section 2.2. In fact, the EKF *is* a linear filter, in that it

weights predictions and observations in a linear manner in order to estimate the state. In the EKF, however, the prediction equations for state and observation have been replaced by the nonlinear functions $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ respectively evaluated at the state estimate $\hat{\mathbf{x}}$.

In the covariance and gain equations, the linear matrices $\mathbf{F}(t)$ and $\mathbf{H}(t)$ are replaced by $\mathbf{F}'(t)$ and $\mathbf{H}'(t)$ respectively, and are used to calculate the EKF gain, $\bar{\mathbf{K}}(t)$. The matrices $\mathbf{F}'(t)$ and $\mathbf{H}'(t)$ are calculated as linearisations about the state estimate of the functions $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ and are evaluated as

$$\mathbf{F}' = \left. \frac{\partial \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t)}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(t)} \quad (2.41)$$

$$\mathbf{H}' = \left. \frac{\partial \mathbf{h}(\mathbf{x}(t), t)}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(t)} \quad (2.42)$$

The Jacobians $\mathbf{F}'(t)$ and $\mathbf{H}'(t)$ are the result of truncating a Taylor series expansion of $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ after the linear terms.

As may be seen from these equations, the EKF essentially treats the estimation problem in the same framework as the linear Kalman filter. For prediction, the nonlinear equations are used, and these nonlinear functions are linearised about the operating point to provide the update of covariance and gain equations. By neglecting high order terms in this linearisation, the EKF only provides an approximation of the true gain and covariance for the nonlinear system. This approximation is valid for systems where the error between the true state and the estimated state remains small[38]. In practice, this means that the EKF will be valid for systems where the sample time is small compared to the rate of change of the nonlinear functions, such that any linearisations made are increasingly accurate.

2.6 The Discrete Kalman Filter

The Kalman filter is easily understood in the continuous time (and frequency) domain. Indeed, this is the domain in which most processes are modelled. Most Kalman filters, however, are implemented in discrete time to allow the incorporation of sampled data, and for implementation on digital computers.

The discrete time Kalman filter has a slightly different form to its continuous time counterpart. It consists of two distinct steps, the prediction, and the update. This may be seen

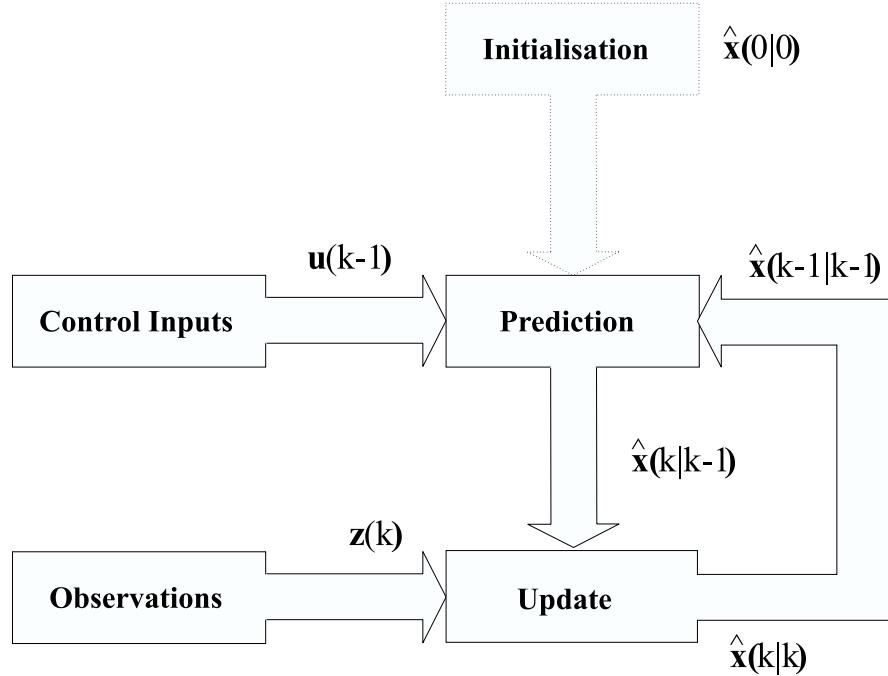


Figure 2.13: The Discrete Estimation Process

graphically in Figure 2.13.

In its discrete form, the Kalman filter uses the control inputs, the estimated state, and the process model at the discrete time step $k - 1$ to predict the state forward to the discrete time k . At time k , observations are made which are used to update the state estimate.

The prediction equations for state and covariance are given by [3, 12, 32]

$$\hat{\mathbf{x}}(k|k-1) = \mathbf{F}_d(k-1)\hat{\mathbf{x}}(k-1|k-1) \quad (2.43)$$

$$\mathbf{P}(k|k-1) = \mathbf{F}_d(k-1)\mathbf{P}(k-1|k-1)\mathbf{F}_d^T(k-1) + \mathbf{Q}(k-1) \quad (2.44)$$

where the matrices \mathbf{x} and \mathbf{P} retain their definitions from Section 2.2. The matrices \mathbf{F}_d and \mathbf{Q} are the state transition matrix and the discrete time noise covariance matrix respectively.

The estimate (or update) of state and covariance are given by

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1)] \quad (2.45)$$

$$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}(k)\mathbf{S}(k)\mathbf{K}^T(k) \quad (2.46)$$

where the matrix \mathbf{H} retains its definition from Section 2.2 and the discrete time gain is defined as

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{H}^T(k)\mathbf{S}(k)^{-1} \quad (2.47)$$

where \mathbf{S} is given by

$$\mathbf{S}(k) = \mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}^T(k) + \mathbf{R}(k) \quad (2.48)$$

and \mathbf{R} is the observation noise covariance matrix.

Appendix B presents the theory used to derive the discrete time system model \mathbf{F}_d from the equivalent continuous time model. In fact, there are many matrix manipulation packages such as Matlab which easily convert processes between the continuous and discrete time domains.

2.7 Modelling Process Faults

Typically, process faults are caused by inaccurate models. These model inaccuracies cause the process noise $v(t)$ to become non-white. This coloured process noise may be estimated and a shaping filter constructed in a method analogous to that for coloured sensor noise.

Many common sources of process error, however, may be reduced or eliminated simply by providing better models.

This section presents a process model considered representative of the process models encountered in AGV systems, giving an analysis as to potential sources of error, and suggestions for improvement. The insights provided are generalisable to many common processes.

2.7.1 The Tricycle Drive Robot

Example 2.7.1 Consider a simple three wheeled mobile robot as shown in figure 2.14. The equations of motion may be derived from the rigid body and rolling motion constraints on the vehicle. For this system, the kinematic equations of motion are

$$\begin{aligned} \dot{x}(t) &= V(t) \cos(\phi(t) + \gamma(t)) \\ \dot{y}(t) &= V(t) \sin(\phi(t) + \gamma(t)) \\ \dot{\phi}(t) &= \frac{V(t) \sin(\gamma(t))}{B} \end{aligned} \quad (2.49)$$

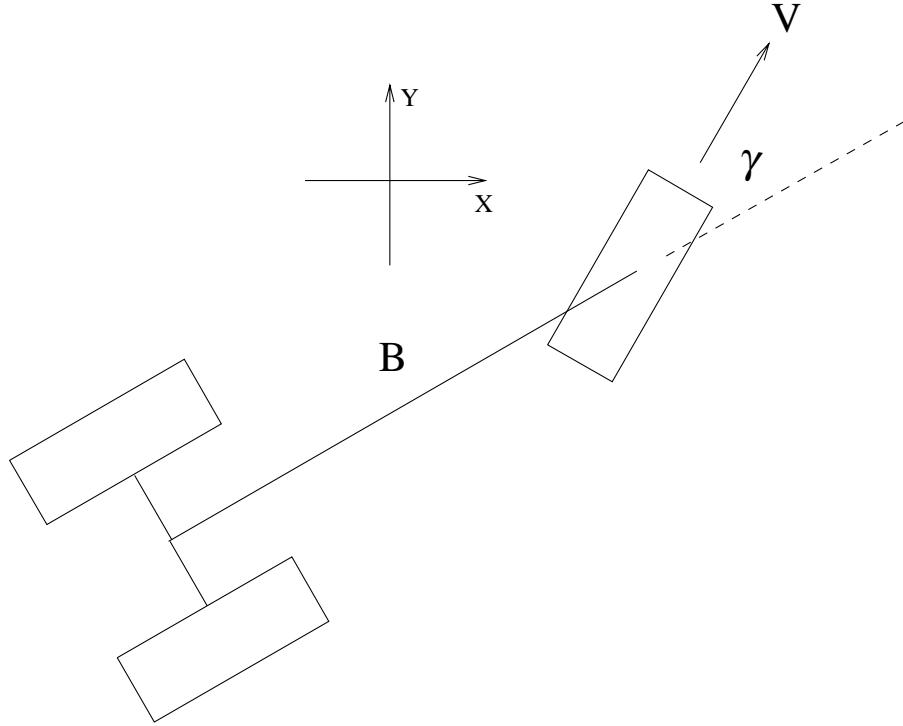


Figure 2.14: Tricycle Drive Mobile Robot Geometry

where $x(t)$ and $y(t)$ denote the position of the vehicle, the angle $\phi(t)$ is the orientation of the vehicle with respect to the x axis, and $V(t)$ represents the linear velocity the front wheel. The angle γ is defined as the steer angle of the vehicle, and B is the base length of the vehicle. These equations are reported widely in the literature[23, 26, 48, 64].

Measurements are made of the wheel's angular velocity ω and steer angle γ .

It should be noted that for nonlinear models such as this, a closed form solution for the discrete time vehicle model can not in general be found by integrating Equation 2.49 between the time intervals t_k and t_{k-1} as is possible in the linear case (see Appendix B). For nonlinear models, an Euler approximation of the integral can be used, assuming the control inputs are approximately constant over this interval. This approximation yields a discrete time

vehicle model of the form

$$\begin{aligned} x(k) &= x(k-1) + \Delta T V(k-1) \cos(\phi(k-1) + \gamma(k-1)) \\ y(k) &= y(k-1) + \Delta T V(k-1) \sin(\phi(k-1) + \gamma(k-1)) \\ \phi(k) &= \phi(k-1) + \Delta T \frac{V(k-1) \sin(\gamma(k-1))}{B} \end{aligned} \quad (2.50)$$

where ΔT is defined as the sample time from t_{k-1} to t_k .

2.7.2 Error Models

Examining equation 2.49 for possible sources of model error, that is, parameters that may be sensed or measured incorrectly, a list can be formed as follows:

- The calculated linear velocity $V(t)$ may be incorrect
- The vehicle may not be heading in the direction indicated by the steer angle $\gamma(t)$
- The measured base length B may be incorrect.

These errors are illustrated graphically (although greatly exaggerated) in Figure 2.15.

The first error, where $V(t)$ differs from the true linear velocity by $\delta V(t)$, can be modelled as an error in wheel radius. As $V(t) = \omega(t)R(t)$, and $\omega(t)$ is a measured quantity (assumed to be accurate), the only parameter that can cause a difference in the magnitude of $V(t)$ is $R(t)$. Therefore an error in the wheel radius is a multiplicative error with respect to the vehicle velocity.

The second error source occurs when the instantaneous velocity vector is not parallel with the front wheel. This effect may be modelled as an angle $\alpha(t)$ that the true vector deviates from the measured steer angle $\gamma(t)$. This is an additive error, as may be seen from Figure 2.15.

The final effect, of incorrect base length measurement, may simply be modelled by an error $\delta B(t)$ which describes the difference in length between the true base length and the measured base length. This is also an additive error.

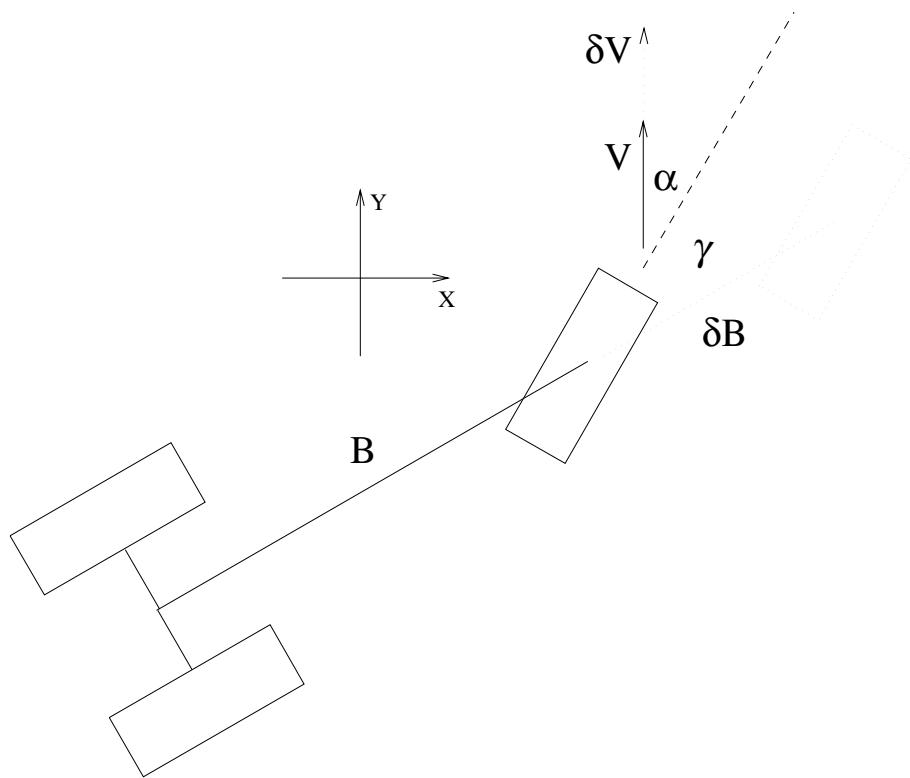


Figure 2.15: Tricycle Drive Mobile Robot Error Source Geometry

There are now a further three state variables that need to be added to the model: $R(t)$, $\alpha(t)$ and $\delta B(t)$. Assuming that these parameters at any instant in time are unknown constants, the continuous time vehicle model becomes

$$\begin{aligned}\dot{x}(t) &= \omega(t)R(t)\cos(\phi(t) + (\gamma(t) + \alpha(t))) \\ \dot{y}(t) &= \omega(t)R(t)\sin(\phi(t) + (\gamma(t) + \alpha(t))) \\ \dot{\phi}(t) &= \frac{\omega(t)R(t)\sin(\gamma(t) + \alpha(t))}{B + \delta B(t)} \\ \dot{R}(t) &= 0 \\ \dot{\alpha}(t) &= 0 \\ \dot{\delta B}(t) &= 0\end{aligned}\tag{2.51}$$

Note that $V(t)$ has now been replaced by $\omega(t)R(t)$.

An interesting and useful side effect of these extra states is that they act as lumped parameters, acting to compensate for unmodelled sources of error. For example, if the vehicle's drive wheel slips, resulting in the predicted vehicle position differing from the true position, the resulting position error will be estimated as a change in the parameters R , α and δB . In the case where a simple kinematic model is used, there is no way for the filter to compensate for errors such as this, resulting in the vehicle model continuously giving a biased estimate of pose to the filter which will almost certainly cause divergence.

Having the extra states in the vehicle model does pose problems however. The system has sensor(s) which supply only a finite amount of information, and, if the vehicle model contains too many states, this information may be spread between the states 'too thinly'. Also, some or all of the extra states may be unobservable. For example the wheel radius cannot be estimated when the vehicle is stationary. For a detailed analysis and discussion of the issues involved refer to [49].

Given that the model may become too complicated and/or unobservable, parameters may have to be omitted from the state vector. The obvious decision is to leave out those parameters which least effect the states of interest, which in this case are $x(t)$, $y(t)$ and $\phi(t)$.

2.7.3 Sensitivity Analysis

In order to simplify the model presented in equation 2.51, a simple sensitivity analysis can be performed. The analysis shows how variations in each of the parameters affects each of the states. Once the dominant error sources have been determined, the order of the model may be reduced by removing those states which are determined to have little effect on estimation accuracy.

Example 2.7.2 Continuing with the example given in Equation 2.49, take the Jacobian of the states $\dot{x}(t)$, $\dot{y}(t)$, and $\dot{\phi}(t)$ with respect to the error sources $V(t)$, $\gamma(t)$, and B as

$$\frac{\partial(\dot{x}(t), \dot{y}(t), \dot{\phi}(t))}{\partial(V(t), \gamma(t), B)} = \begin{bmatrix} \cos(\phi(t) + \gamma(t)) & -V(t) \sin(\phi(t) + \gamma(t)) & 0 \\ \sin(\phi(t) + \gamma(t)) & V(t) \cos(\phi(t) + \gamma(t)) & 0 \\ \frac{\sin(\gamma(t))}{B} & \frac{V(t) \cos(\gamma(t))}{B} & \frac{-V(t) \sin(\gamma(t))}{B^2} \end{bmatrix} \quad (2.52)$$

This gives a measure of the sensitivity of the states with respect to the error sources. It is evident from the third column of the Jacobian that errors in B do not affect the states $\dot{x}(t)$ and $\dot{y}(t)$ at all, and only affect $\dot{\phi}(t)$ in the order $\frac{1}{B^2}$. The other error sources, however, affect all of the states, with greater magnitude.

Therefore, if any error sources may be neglected, then the first choice should be the error in B as it has the least effect of all the modelled error sources. It should also be noted that the error in B may be considered a static error, that is, an error which will not significantly change in time. It is not unreasonable to assume that better calibration will reduce the size of this mismeasurement. As opposed to the error in B (which is internal to the vehicle), the errors in $V(t)$ and $\gamma(t)$ may be considered to be dynamic, being mainly caused by forces acting external to the vehicle, and not being constant in time. No amount of calibration is sufficient to detect slip for example, a dynamic source of error, although the slip may be estimated if enough information exists from sensor measurements.

For this example then, to reduce the complexity of the model, the parameter $\dot{\delta}B(t)$ will be neglected. The continuous time model then becomes

$$\begin{aligned}\dot{x}(t) &= \omega R(t) \cos(\phi(t) + (\gamma(t) + \alpha(t))) \\ \dot{y}(t) &= \omega R(t) \sin(\phi(t) + (\gamma(t) + \alpha(t))) \\ \dot{\phi}(t) &= \frac{\omega R(t) \sin(\frac{\gamma(t)+\alpha(t)}{2})}{B} \\ \dot{R}(t) &= 0 \\ \dot{\alpha}(t) &= 0\end{aligned}\tag{2.53}$$

This model may now be used in an EKF, and providing sufficient sensing is present, will be more accurate than the conventional kinematic model. The detailed design and implementation of a vehicle model using these techniques can be seen in Section 5.3.

2.8 Characteristic Sensor Errors

It is important to understand the typical types of errors that navigation sensors exhibit. The sensor error characteristics may influence the system designer's choice of sensor or combination of sensors. This section examines some commonly used sensing technologies and summarises those sensors' error characteristics, motivating the study of three 'characteristic' sensor error models.

Following is a short survey of navigation sensors, paying close attention to the types of errors one should expect from the various sensing technologies.

2.8.1 Absolute Sensors

Absolute sensors may be characterised as providing a measurement of the state(s) of interest.

The following list summarises the absolute sensors commonly used in navigation systems, together with a brief discussion of the error characteristics.

- **Odometry Sensors** - Sensors such as encoders, resolvers and potentiometers are routinely modelled as having white noise type error on an absolute position measurement.

- Sonar Sensors - According to Schenkel[87], in-air sonar sensors display range measurement error which is approximately normally distributed with a mean of zero. Of course this is only applicable when the sensor is perpendicular to the surface being measured and there are no nearby obstructions.
- Radar - Millimetre wave radar, in the form most useful to navigation systems, measures range and bearing to targets at known locations. Bearing information is usually supplied via an optical encoder which measures the angle of the swash plate. Durrant-Whyte[28] describes a navigation system based on millimetre wave radar that achieves accuracies of better than 3cm . In this paper, the radar bearing and range data are adequately modelled as being in white noise only.
- Laser Sensors - Similarly to radar, bearing measurements for lasers are typically achieved using optical encoders. Therefore, the bearing measurement is corrupted by white noise. Kweon[53] shows that the range measured by an infrared laser is Gaussian distributed with a variance that scales with the range measured. This means that for any fixed distance, the laser range measurement noise is also white.

These sensors are all characterised as having measurements corrupted by white noise only. There are other absolute sensors which cannot be characterised in this way as described in the following:

- Global Positioning System (GPS) - When using CA code (the degraded form of GPS available to civilians), GPS measurement error displays a second order time correlation[22, 67]. It is interesting to note that the measurement error for each satellite is nearly identical, suggesting that the algorithm used to generate the noise is the same for each satellite, and that the remaining small differences are due to atmospheric effects. Differential GPS effectively decorrelates the coloured noise from the position solution improving accuracy to the point where it is possible to buy a commercial unit capable of 2cm precision. Nebot[69] uses white noise to model the error of a high quality (around 37cm accuracy) differential GPS unit.

As described, conventional GPS has a time correlated measurement error, and unless this is modelled will cause the Kalman filter to fail.

2.8.2 Rate Sensors

Rate sensors are characterised as measuring the rate of change of the states of interest. The following list summarises common rate sensors and their error characteristics.

- Inertial Measurement Unit (IMU) - Inertial sensors (accelerometers and gyros) are well known to exhibit undesirable error characteristics. As well as displaying bias, inertial sensors are prone to drift[30] which can be adequately modelled as either Brownian motion or a time correlated error[63]. In sampled systems, the signals from inertial sensors will often display a white noise component as well.
- Doppler Radar - Doppler radar systems measure the velocity (inferred from Doppler shift) of a surface relative to the sensor. To obtain position information, this velocity is integrated. This integration process introduces Brownian process noise to the sensor observations[66].

These rate sensors are characterised as having drift type measurement error. In some cases, these sensors will also display time correlated errors[63].

2.8.3 Sensor Error Models

From the survey presented in this section, it can be seen that a sensor's measurement error typically falls into one of three categories. These are

- Measurements corrupted by white noise
- Measurements corrupted by time-correlated noise
- Measurements corrupted by random walk, drift or bias

Sensors corrupted by white noise only are a subset of the class of absolute sensors. Those corrupted by Brownian motion are a subset of the class of rate sensors. Time correlated errors may affect sensors from either class.

Figure 2.11 showed that each of these types of measurement error may be represented by a shaping filter and thus incorporated into the Kalman filter framework. The behaviour

of the Kalman filter under the assumption that measurement errors are white Gaussian sequences has been shown. Sensors with coloured measurement noise will exhibit different frequency domain behaviour, as will be seen in the next section.

To simplify things, sensors with white measurement noise will be referred to as type I sensors, those sensors with time correlated measurement noise as type II, and sensors subject to brownian motion or random walk as type III.

2.9 Sensor Error and Filter Transfer

This section examines the effect of the different characteristic sensor error models identified in Section 2.8 on the Kalman filter transfer functions.

The examples that follow use the constant velocity model introduced in Section 2.2.2, as this is perhaps the simplest practical process model used in estimation. This simplicity will aid in the understanding of the Kalman filter, and the way in which sensor information is used.

2.9.1 Type I Sensors

As discussed in Section 2.8, type I sensors are absolute sensors corrupted by white noise only. In frequency terms, these sensors supply a band limited signal corrupted by broad band noise.

Example 2.9.1 Consider the constant velocity process model and observation model for a type I sensor as follows

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \quad (2.54)$$

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + v(t) \quad (2.55)$$

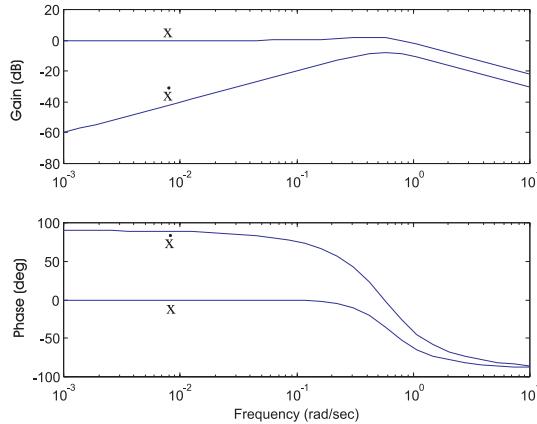


Figure 2.16: Bode Diagram of the Transfer Function of a Single Type I Sensor

with system with process and observation noise covariance matrices given by

$$\mathbb{E} [w(t)w^T(t + \tau)] = [0.0001] \quad (2.56)$$

$$\mathbb{E} [v(t)v^T(t + \tau)] = [0.001] \quad (2.57)$$

The Bode plot of the transfer functions from the sensor to the states is shown in Figure 2.16.

The response of this system is exactly as predicted by the example in Section 2.2.2.

The sensor has a low pass response to the position state and a band pass response to the velocity state.

2.9.2 Type II Sensors

Type II sensors are those corrupted by time correlated measurement errors. Conventional GPS is an example of a sensor of this type. These sensors supply a band limited signal corrupted by predominantly low frequency noise, as can be seen in Figure 2.11.

Example 2.9.2 *The constant velocity process model (including shaping state) and obser-*

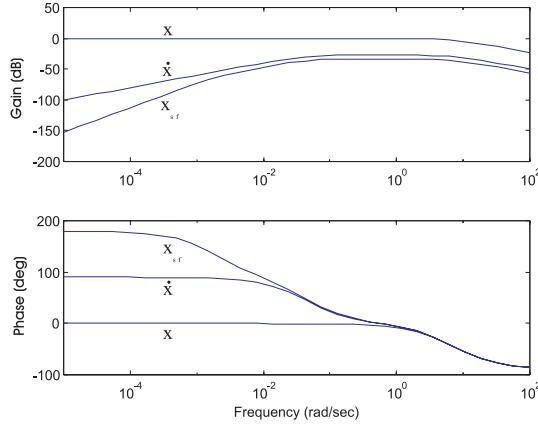


Figure 2.17: Bode Diagram of the Transfer Function of a Single Type II Sensor

vation model for a type II sensor are

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \\ \dot{x}_{sf}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\omega \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ x_{sf}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ w_1(t) \\ w_2(t) \end{bmatrix} \quad (2.58)$$

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ x_{sf}(t) \end{bmatrix} + v(t) \quad (2.59)$$

To show the effect of coloured sensor noise on the filter transfer, the sensor and process noise variance has been kept the same as the previous example. The system's process and observation noise covariance matrices and correlation frequency are given by

$$\mathbb{E} [\mathbf{w}(t)\mathbf{w}^T(t+\tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.05 \end{bmatrix} \quad (2.60)$$

$$\mathbb{E} [v(t)v^T(t+\tau)] = [0.001] \quad (2.61)$$

$$\omega = 0.002 \quad (2.62)$$

This Bode plot of the transfer functions is shown in Figure 2.17. The transfer function for a single type II sensor appears to be very similar to that of a single type I sensor.

As the filter has no additional information to decorrelate the true type II signal from the coloured noise, the best use of the information is to simply use all of it. Even though it is known that type II sensors have high noise power at low frequencies, it is better to use this information than not to use it and have no information at all.

The cutoff frequency for this system appears to be higher than for the equivalent type I sensor. This may be accounted for by noting that additional uncertainty has been added in the \mathbf{Q} matrix. The tradeoff is that by introducing extra noise terms, the associated covariance matrix will also increase.

This means that when using a single type II sensor, the filter will exhibit almost identical behaviour regardless of whether the sensor error has been modelled. Therefore, further sensing is needed to aid in the decorrelation of the coloured noise to improve system accuracy.

2.9.3 Type III Sensors

Type III sensors are those corrupted by drift type measurement errors. These sensors are characterised as measuring the rate of change of the state(s) of interest.

Example 2.9.3 *The constant velocity process model (including shaping state) and observation model for a type III sensor are*

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \\ \dot{x}_{sf}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ x_{sf}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ w_1(t) \\ w_2(t) \end{bmatrix} \quad (2.63)$$

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ x_{sf}(t) \end{bmatrix} + v(t) \quad (2.64)$$

However, when the observability grammian for this system is constructed[63], it is found to be rank deficient. The system is unobservable, due mainly to the fact that there is no way to differentiate between drift and signal in a system with a single type III sensor. For this reason, a Bode diagram cannot be constructed.

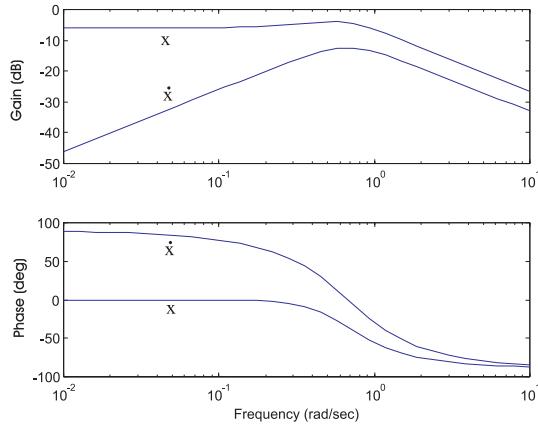


Figure 2.18: Bode Diagram of the Transfer Function of Dual Type I Sensors

2.9.4 Dual Type I Sensors

Arguably the simplest multi-sensor system to analyse is one with exactly matched sensors. In this section, the effect of using two identical type I sensors is discussed.

Example 2.9.4 Consider the process model given for the type I sensor combined with an observation equation given by

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \quad (2.65)$$

which is, of course, two type I sensors both observing the position state.

The process and observation noise covariance matrices are

$$\mathbb{E}[w(t)w^T(t+\tau)] = [0.0001] \quad (2.66)$$

$$\mathbb{E}[\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (2.67)$$

The transfer functions for the system are shown in Figure 2.18. Note that the transfer functions for each sensor are identical (therefore only one is shown). The main difference between this system and the system with only a single type I sensor is in the magnitude of the Bode plot.

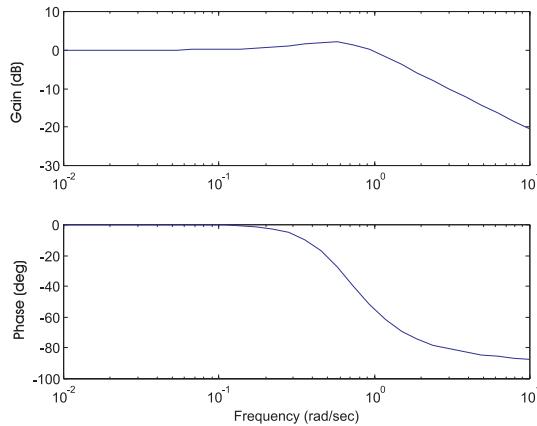


Figure 2.19: Bode Diagram of the Combined Transfer Function of Dual Type I Sensors

The magnitude of the pass band of the filter shown in Figure 2.18 has fallen from $0dB$ to $-6dB$. This means that each sensor is contributing 50% of the information used by the filter.

This can be seen in Figure 2.19, which is the combined transfer function (to the position state) of each of the type I sensors. This figure is identical to that shown in Figure 2.16.

In this thesis, a combined transfer function will refer to the vector sum of the responses of more than one sensor to a single state, which, because the system is linear, represents the overall transfer of information.

The Kalman filter, therefore, weights the information from the sensors according to their accuracy.

In this example, if one type I sensor were slightly better than the other (measured by the noise covariance matrix), it would be expected that it will provide more information to the filter. In the limit, if one sensor has very much smaller noise covariance than the other, the filter will essentially only use information from the better sensor, and attenuate all the information from the other.

2.9.5 Type I and Type II Sensors

This section discusses the effects of aiding a type II sensor with information supplied from a type I sensor. As previously stated, type II sensors must be aided in order that the coloured

measurement noise may be decorrelated from the signal.

Example 2.9.5 A constant velocity system with type I and type II information combined would have a process model as given for a single type II sensor and an observation model of the form

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ x_{sf}(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \quad (2.68)$$

The process and observation noise covariance matrices are

$$\mathbf{E} [\mathbf{w}(t)\mathbf{w}^T(t+\tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.05 \end{bmatrix} \quad (2.69)$$

$$\mathbf{E} [\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (2.70)$$

$$\omega = 0.002 \quad (2.71)$$

The transfer functions from the type I sensor to the output and the type II sensor to the output are given by Figures 2.20 and 2.21 respectively. It can be seen from these Bode plots that the information from the type II sensor is of such poor quality when compared to the type I sensor that its entire signal spectrum is attenuated with respect to the states of interest x and \dot{x} . The type II sensor in this case is only useful for estimating its own coloured noise (shaping state). Figure 2.22 shows the combined transfer function for this system to the position state.

Because of the poor quality of the type II sensor information in this example, the combined transfer function is essentially the same as for the type I sensor alone.

Also note that in figure 2.20 the state x and the state x_{sf} are approximately 180 degrees out of phase. This means that the type I sensor is subtracting the coloured noise from the position state over the entire operating frequency range.

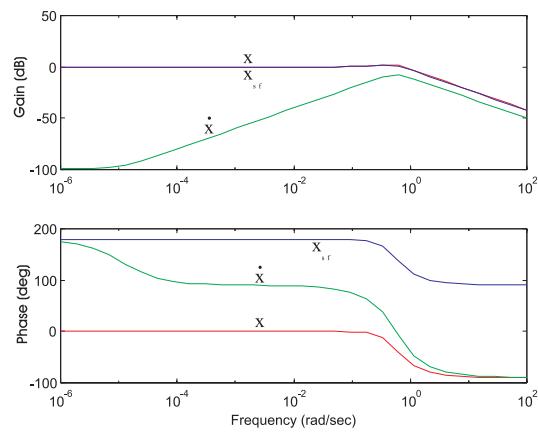


Figure 2.20: Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type II Sensors

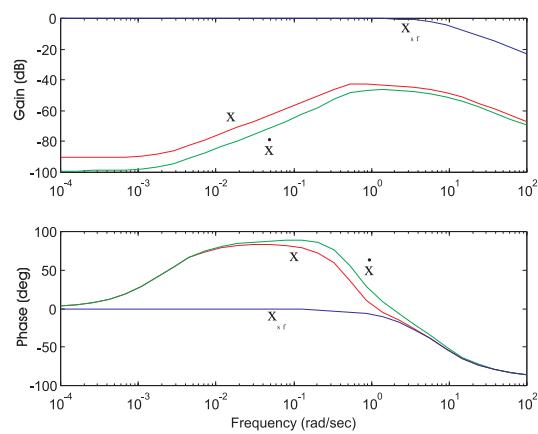


Figure 2.21: Bode Diagram of the Transfer Functions from a Type II Sensor for a System Consisting of Type I and Type II Sensors

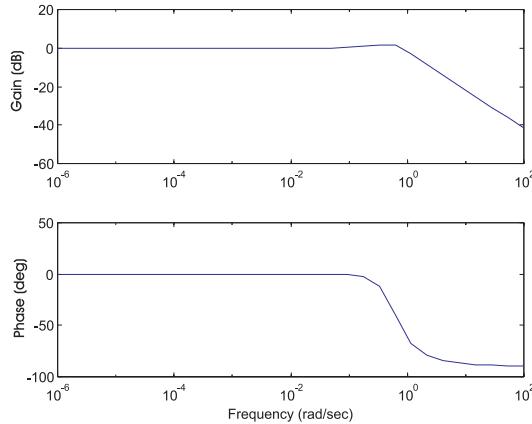


Figure 2.22: Bode Diagram of the Combined Transfer Function of Type I and Type II Sensors to the Position State

Example 2.9.6 Consider now the same example with the following matrices

$$\mathbb{E} [\mathbf{w}(t)\mathbf{w}^T(t + \tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.0001 \end{bmatrix} \quad (2.72)$$

$$\mathbb{E} [\mathbf{v}(t)\mathbf{v}^T(t + \tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (2.73)$$

$$\omega = 0.002 \quad (2.74)$$

which represents a decrease in the driving noise for the shaping filter, and an increase in the quality of the type II sensor.

As can be seen from Figures 2.23 and 2.24, this improvement in the quality of the type II sensor means that the Kalman filter now uses it for state estimation in the higher frequency range.

The combined transfer function for the position state (see Figure 2.25) now has a cutoff frequency that is significantly higher than that allowed for by the single type I sensor case.

The sensors in this example are complementary. The type I sensor provides good low frequency information to the filter, while the type II sensor now provides good information in the mid to high frequency range.

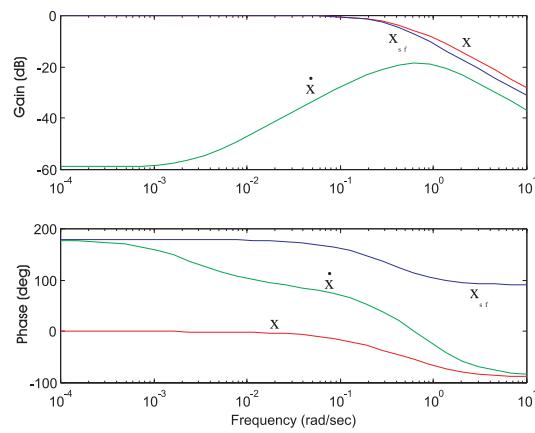


Figure 2.23: Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type II Sensors

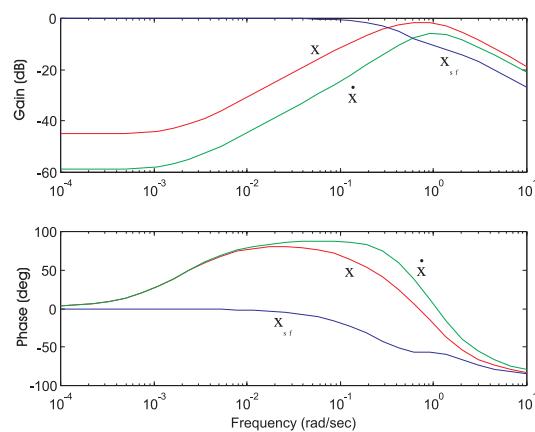


Figure 2.24: Bode Diagram of the Transfer Functions from a Type II Sensor for a System Consisting of Type I and Type II Sensors

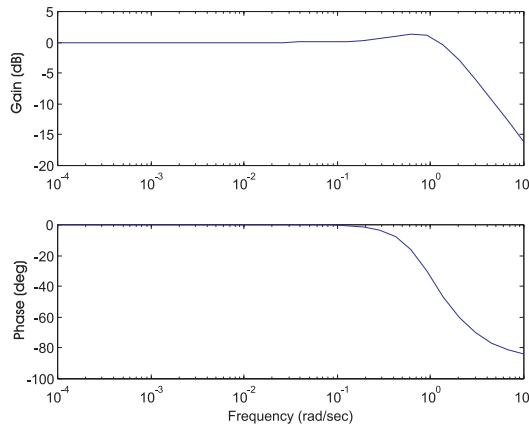


Figure 2.25: Bode Diagram of the Combined Transfer Function of Type I and Type II Sensors to the Position State

In the extreme case, when the type II sensor is significantly better quality than the type I sensor, as in the case given by the following matrices

$$\mathbb{E} [\mathbf{w}(t)\mathbf{w}^T(t + \tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.0001 \end{bmatrix} \quad (2.75)$$

$$\mathbb{E} [\mathbf{v}(t)\mathbf{v}^T(t + \tau)] = \begin{bmatrix} 0.00001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (2.76)$$

$$\omega = 0.002 \quad (2.77)$$

the information supplied by the sensors may have ‘gaps’ in the frequency band, as can be seen in Figure 2.26.

When a frequency gap in the combined transfer functions occurs, further sensors must be added such that the entire spectrum of operating manouevres is able to be tracked.

This is described in more detail in Chapter 4.

2.9.6 Type I and Type III Sensors

This section discusses the effects of aiding type III sensors with information from type I sensors. Type III sensors must be aided simply to make the system observable.

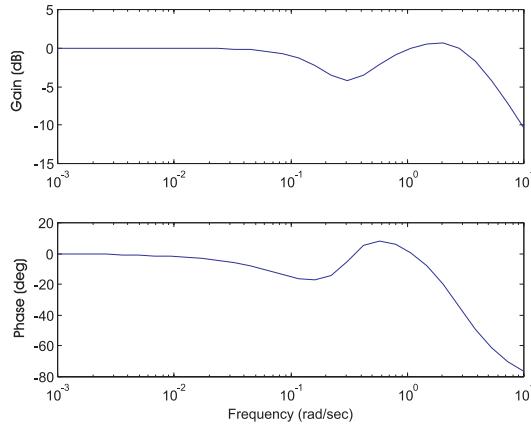


Figure 2.26: Bode Diagram of the Combined Transfer Function of Type I and Type II Sensors to the Position State, Showing Frequency ‘Gap’

Example 2.9.7 A constant velocity system with type I and type III information combined would have a process model as given for a single type III sensor and an observation model of the form

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ x_{sf}(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \quad (2.78)$$

Note that this system is observable.

The system has noise covariance matrices given by

$$\mathbb{E} [\mathbf{w}(t)\mathbf{w}^T(t+\tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.05 \end{bmatrix} \quad (2.79)$$

$$\mathbb{E} [\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (2.80)$$

The transfer functions from the type I sensor to the output and the type III sensor to the output are given by Figures 2.27 and 2.28 respectively.

The system responses are very similar to those in the type I and type II system, which is not surprising, as in the limit as the correlation frequency approaches zero, the time correlated noise model approaches the Brownian model.

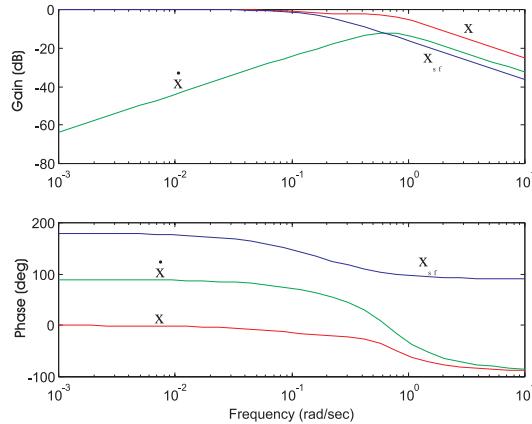


Figure 2.27: Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type III Sensors

The main difference that can be seen here is that a type III sensor with identical noise covariance matrices as a type II sensor provides much more information. Figure 2.29 shows the combined transfer function (to the position state) for this system.

When the type III sensor is very much better quality than the type I sensor, as in the case given by the matrices

$$E [\mathbf{w}(t)\mathbf{w}^T(t+\tau)] = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.00005 \end{bmatrix} \quad (2.81)$$

$$E [\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.00001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (2.82)$$

the transfer functions from the type I and type III sensors are given in Figures 2.30 and 2.31 respectively.

These sensors are complementary, causing the combined transfer function to the position state to track high frequencies due to the type III sensor and low frequencies due to the type I sensor, giving an overall transfer function as seen in Figure 2.32.

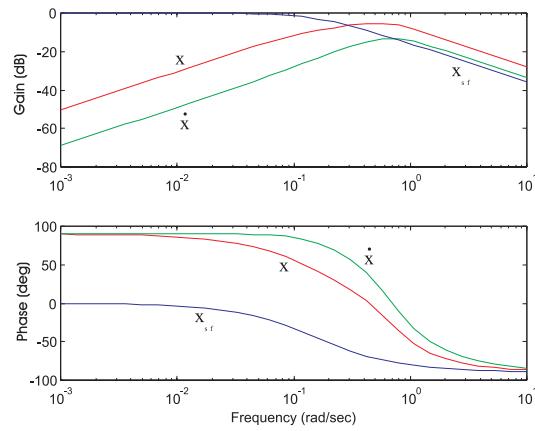


Figure 2.28: Bode Diagram of the Transfer Functions from a Type III Sensor for a System Consisting of Type I and Type III Sensors

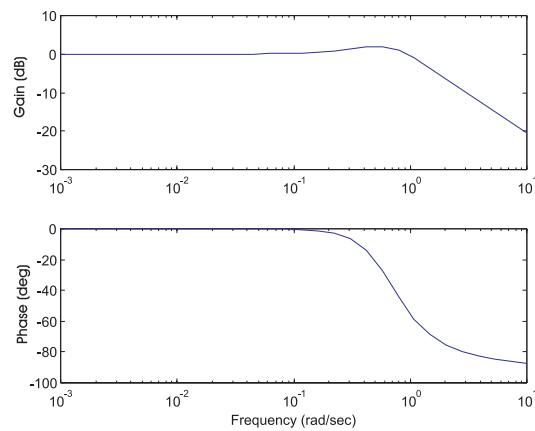


Figure 2.29: Bode Diagram of the Combined Transfer Function of Type I and Type III Sensors to the Position State

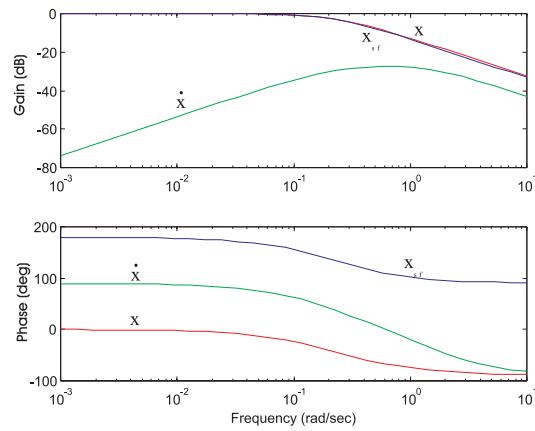


Figure 2.30: Bode Diagram of the Transfer Functions from a Type I Sensor for a System Consisting of Type I and Type III Sensors

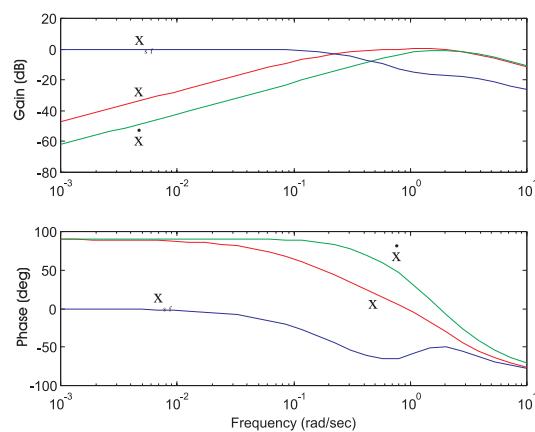


Figure 2.31: Bode Diagram of the Transfer Functions from a Type III Sensor for a System Consisting of Type I and Type III Sensors

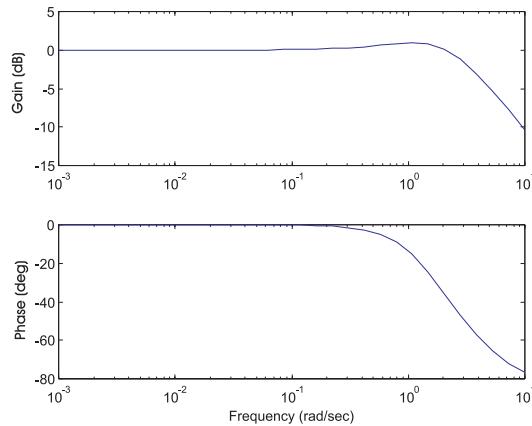


Figure 2.32: Bode Diagram of the Combined Transfer Function of Type I and Type III Sensors to the Position State

2.9.7 Summary

The following is a summary of the insights gained by studying sensor error models in the frequency domain.

- Regardless of the error model, without additional sensing, correlated measurement noise cannot be accurately estimated.
- Sensor information is weighted according to its accuracy (in multi-sensor systems), such that the sum of the weights provides 100% of the required information, except in the case where the sensors simply do not provide enough information at a particular frequency.
- Time correlated errors (Type II), and brownian motion (Type III) are both predominantly low frequency effects. These sensors provide good high frequency information, however, if their quality is significantly better than the Type I sensor.
- Type III sensors provide more information than type II sensors with the same noise power.

These insights allow an AGV navigation system designer to choose sensors based solely on their noise characteristics to achieve a desired frequency response.

For example it was shown that type I and type III sensors have complementary frequency domain behaviour.

Chapter 4 will examine methods for determining the *quality* of the sensors needed for reliable navigation.

Further, as will be shown in Chapter 3, the frequency domain behaviour of the various sensors has a direct impact on the detectability of unmodelled sensor faults.

2.10 Frequency Interpretation of the EKF

In Section 2.2.2, the transfer function for the Kalman filter was derived. There is no such closed form transfer function for the EKF due to the nonlinear process and observation models.

The transfer function of the EKF can, however, be evaluated numerically at any instant in time. This means that there will be a ‘family’ of transfer functions dependent on the evaluation time and the estimated state at that time. In general there may be an infinite number of members of this family.

Nonlinear vehicle models, however, such as the one presented in Section 2.7.1, tend to be smoothly nonlinear. The vehicle geometry constrains the motion of the vehicle to follow splines, the rate of change of which is determined by the response of the steering mechanism. Further, the parameters in vehicle models tend to be bounded. For example, the steer angle of most conventional vehicles will never exceed ± 90 degrees.

So, if the variation of the nonlinear functions is limited, it follows that the family of transfer functions of the system will be bounded.

It should be noted that due to the correspondence between the Kalman filter for time varying linear systems and the EKF shown in Section 2.5, the analysis provided here for the EKF applies equally well to time varying linear systems.

2.10.1 A Tricycle Drive Vehicle Example

Example 2.10.1 Consider the tricycle drive vehicle model presented in Section 2.7.1. The vehicle is equipped with a sensor that measures range and bearing to a beacon whose position

parameter	minimum	maximum
V (velocity)	0 ms^{-1}	4 ms^{-1}
ϕ (orientation)	$-\pi \text{ rad}$	$\pi \text{ rad}$
γ (steer angle)	$-\frac{\pi}{2} \text{ rad}$	$\frac{\pi}{2} \text{ rad}$
X_b (beacon position)	1 m	10 m
Y_b (beacon position)	1 m	10 m

Table 2.1: Parameter Minimum and Maximum Values

is (X_b, Y_b) . The nonlinear process model for the vehicle is given in Section 2.7.1, and the nonlinear observation model is given by

$$\begin{bmatrix} z_r(t) \\ z_\theta(t) \end{bmatrix} = \begin{bmatrix} \sqrt{(X_b - x(t))^2 + (Y_b - y(t))^2} \\ \arctan\left(\frac{Y_b - y(t)}{X_b - x(t)}\right) - \phi(t) \end{bmatrix} \quad (2.83)$$

The system parameters were varied between the minimum and maximum values summarised in Table 2.1. Without loss of generality, the vehicle was assumed to be situated at the point $(0, 0)$ at the moment of interest. As the variation in orientation is functionally equivalent to a variation in bearing measurement, the bearing to the beacon was fixed constant. Similarly, the variation in the beacon position is equivalent to varying the range to the beacon. The vehicle base length B was also assumed to be constant.

Figure 2.33 graphs the family of transfer functions from the range measurement to the position states x and y under the variation of parameters listed in Table 2.1. Similarly, Figure 2.34 graphs the family of transfer functions from the bearing measurement to position states, and Figure 2.35 graphs the family of transfer functions from the range and bearing measurements to the orientation state ϕ .

These graphs show that for this simple vehicle model, the frequency response is bounded.

A technique common in \mathcal{H}_∞ control[8, 91] is to calculate the upper and lower bounds of the family of transfer functions to perform a worst case analysis. If either the upper or lower bound are found to fall outside the desired frequency response of the system, the system will have to be redesigned until it meets the specification.

An important aspect of this technique is that it essentially allows a linear analysis to be performed on a nonlinear system. The nonlinear family of transfer functions is simply

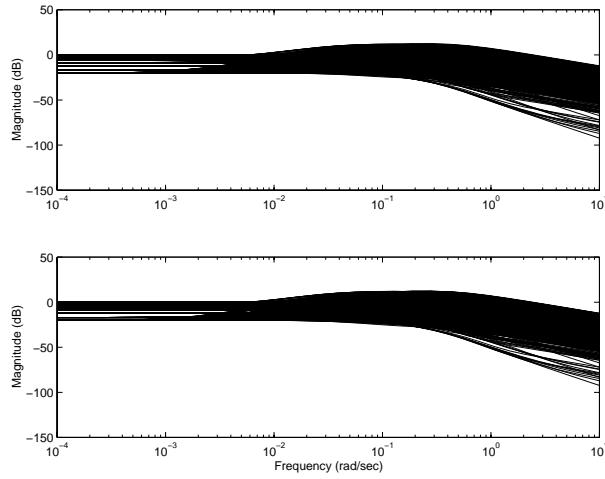


Figure 2.33: Bode Magnitude Plot of the Transfer Function From the Range Measurement to the Position States x (top) and y (bottom)

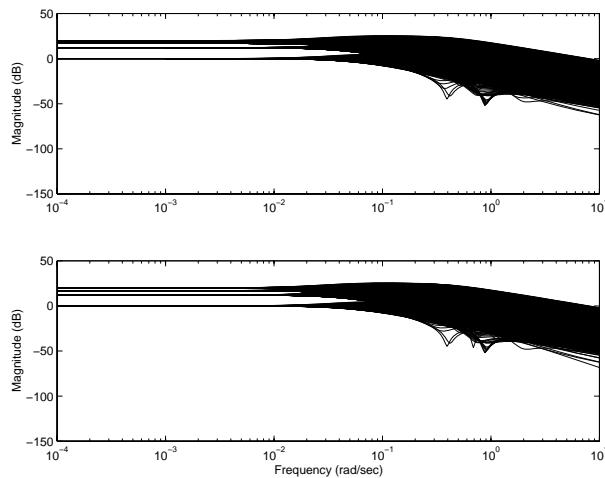


Figure 2.34: Bode Magnitude Plot of the Transfer Function From the Bearing Measurement to the Position States x (top) and y (bottom)

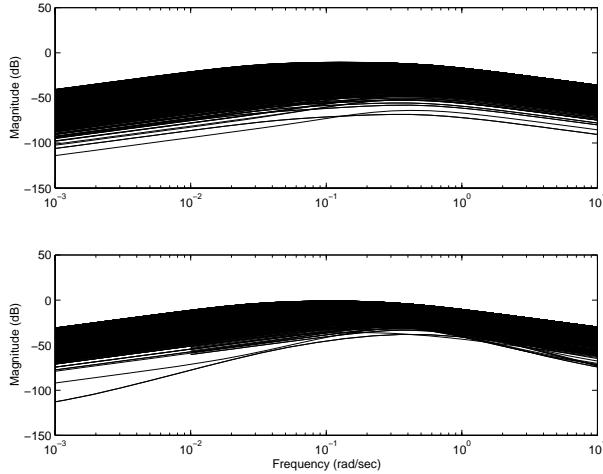


Figure 2.35: Bode Magnitude Plot of the Transfer Function From the Range Measurement (top) and the Bearing Measurement (bottom) to the Orientation State ϕ

replaced by the linear upper and lower bound, allowing linear techniques to be used to generate the worst case response.

2.11 Conclusion

This chapter introduced the theory essential for understanding the operation of a Kalman filter, and more importantly an understanding of what constitutes a navigation failure, through the use of frequency domain analysis techniques.

The continuous time Kalman filter was presented in Section 2.2, and an analysis given describing the filters's operation in the frequency domain. Alternate filter structures were examined which exploit the frequency domain characteristics of sensors.

In Section 2.3, it was concluded that the only possible causes of navigation failure arise from the process model and the observation model, motivating the study of modelling techniques which improve the accuracy of the process and system models.

Section 2.4 introduced a technique for modelling sensor error. Section 2.5 presented the theory required to extend the applications of Kalman filtering to nonlinear systems, and Section 2.6 introduced techniques used to convert between continuous and discrete time system models.

In Section 2.7, a method for determining and modelling the sources of error in process models was presented. Section 2.8 provides a survey of common navigation sensors, concluding that there are typically only three types of measurement error: white noise, time correlated noise and brownian motion. Section 2.9 then provided an analysis of the effects of these sensor error models on the Kalman filter transfer characteristics.

Finally, it was shown in Section 2.10 that for a simple AGV, the frequency response of the nonlinear system is bounded.

Chapter 3

Fault Detection

3.1 Introduction

Chapter 2 examined the issue of navigation including the possible causes of navigation failure, drawing the conclusion that the only things that can cause the Kalman filter to diverge (and thus fail) are process faults, observation faults, and nonlinearities. It was shown that many common failures may be modelled, and incorporated into the Kalman filter framework.

This chapter examines the effect on navigation system performance of those failure modes that are not modelled. Specifically, metrics are developed which give a quantifiable measure of the detectability of faults within the Kalman filter framework.

In Section 3.2, arguments are presented to establish the need for fault detection in a high integrity navigation system. Section 3.3, examines the literature on fault detection, and motivation is given for the detection of faults using innovations testing. Section 3.4 develops metrics which describe the detectability of process faults and sensor faults, including the introduction of fault plots, which graphically display these metrics. In Section 3.5, a method is given for guaranteeing that all process faults may be rendered detectable. Similarly, in Section 3.6 it is argued that redundancy must be employed in order that all sensor faults are detectable. In Section 3.7 it is argued that bounded nonlinear systems can be treated similarly to linear systems in terms of fault detection. Finally in Section 3.8, a simple recursive method for checking filter consistency is given.

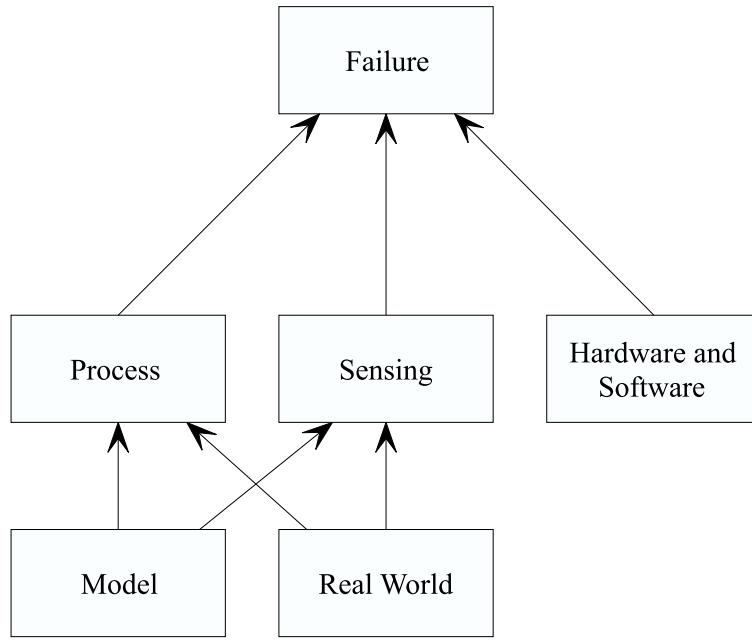


Figure 3.1: The Factors that may Lead to Navigation Failure

3.2 The Need for Fault Detection

In Chapter 1, it was argued that to be commercially viable, AGVs must be considered to be of high integrity. A vital component of high integrity systems is fault detection. By including fault detection, the navigation systems designer has acknowledged that the system will fail at some point.

Consider Figure 3.1. This figure describes the set of possible events that can potentially cause failure in a navigation system. Chapter 2 identified the process and sensor models used in the Kalman filter as potential sources of error. When implementation is also considered, one must assume that computer hardware and software may also potentially cause navigation failure.

Figure 3.1 also shows that the failures caused by the process and sensing blocks may be due to one or both of two influences: The model of the sensor or process, and the real world.

This idea is further expanded in Figure 3.2. This figure shows graphically the error ϵ due to the discrepancies between the model of a process, and the real process. The Kalman filter derivation dictates that ϵ must be a white noise source to ensure correct filter operation, as was identified in Chapter 2.

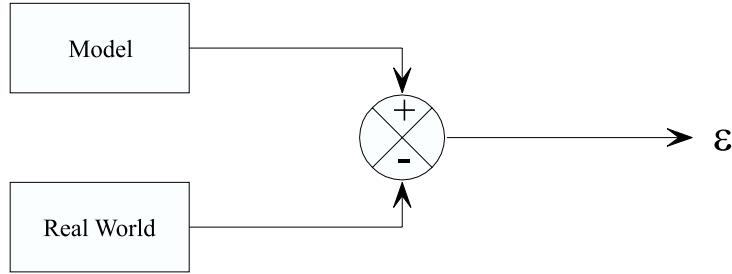


Figure 3.2: The Nature of Fault Processes

Figure 3.2 indicates that a model fault will have the same effect on the error ϵ as a fault occurring in the real world.

This means that to detect all faults, the detection method cannot discriminate as to the source of the fault whether it be a modelling error, or a fault occurring in the real world, such as a sensor breakdown.

The essence of fault detection therefore is to determine when the model of a process differs markedly from the true system. The actual source of the fault cannot be assumed a priori.

Simply having a fault detection scheme does not make a navigation system have high integrity.

For high integrity, one must provide a quantifiable guarantee that all faults are indeed detectable.

This chapter seeks to provide this guarantee.

3.3 Fault Detection

To ascertain whether a navigation filter is indeed performing correctly, one must simply verify that the estimated state lies close to the true state. Or, more specifically, the filter must be unbiased

$$\hat{\mathbf{x}}(k|k) = \mathbb{E} [\mathbf{x}(k)|\mathbf{Z}^k] \quad (3.1)$$

and its state estimation error should satisfy

$$\mathbb{E} [(\mathbf{x}(k) - \hat{\mathbf{x}}(k|k))(\mathbf{x}(k) - \hat{\mathbf{x}}(k|k))^T | \mathbf{Z}^k] = \mathbf{P}(k|k) \quad (3.2)$$

However, the true state is usually not known. In fact if the true state *were* known, then using a Kalman filter to estimate the state would be entirely unnecessary.

To overcome this fundamental problem in fault detection, it is necessary to find and test some *internal* measure of filter consistency, as the only external information available comes from the sensors (which may be faulty).

A great many researchers, over the course of the last thirty years or so, have tried to solve the problem of fault detection. All use some internal measure of system performance, using approaches developed in many varied disciplines. One researcher has optimistically written[82]

...the sensor failure problem has been solved.

However, the number of practical systems using fault detection techniques is small compared to the amount of research being done in this area. Why, if the problem (at least for sensors) is believed to be solved, is there a comparative lack of interest from industry?

There have been numerous survey papers[11, 33, 43–45, 58, 70, 95] which describe various fault detection techniques. These techniques may be summarised as

- ‘Failure Sensitive’ Filters[24, 72, 77] - This set of techniques is based upon two main ideas. The first is the desire to keep the filter sensitive to new data so that sudden changes can be more easily detected. The second is to design a ‘detection’ filter which is designed such that faults have a known effect on the filter innovations. The drawback in this case is that filter optimality is sacrificed in order to achieve these desired characteristics
- Voting Systems[36] - Standard voting schemes utilise at least three identical instruments. Simple techniques may then be used to determine which sensor is at fault. The problem here is that multiple faults cannot be adequately arbitrated. Further, it is not unreasonable to assume that identical sensors will have similar failure modes, which cannot be detected by voting.
- Multiple Hypothesis Filters[56, 65] - This technique utilises a ‘bank’ of filters, each hypothesising that a different failure event has taken place. The conditional probability that each filter is ‘correct’ is used to choose the filter that is most consistent,

and therefore signal whether or not a failure has occurred. By constraining the set of possible faults (in N filters, a maximum of N faults may be detected), this method ignores possible faults. In fact to accurately estimate and identify all possible faults an infinite bank of filters must be constructed.

- Jump Process Filters - The set of possible faults to be detected is severely constrained to those which change abruptly.
- Innovations Based Detection[68] - The innovations (or residuals) of a filter are monitored based on the hypothesis that the filter is in its normal mode of operation. These techniques have the advantage of being relatively easy to implement and do not affect filter performance; however, slowly occurring faults are difficult to detect.
- Analytic Redundancy[34, 73, 76, 94] - Analytic redundancy exploits the designer's knowledge of the system under consideration. In some cases (notably in avionic and subsea applications), several equations may be derived which map a set of inputs to a set of outputs. Redundancy occurs when there exists more than one mapping from the inputs to outputs. Fault detection is implemented by comparing the outputs generated by different sets of inputs. Analytic redundancy is not, however, readily extensible to any possible system, but must be treated on a case by case basis.
- Validation Gates[1] - This technique (which is applicable to sensor faults only), attempts to validate sensor data before it reaches the filter by using a 'gate' determined by some statistical measure. Under the assumption that the filter is performing correctly, this approach deletes useful and valid data. When the filter fails, the sensor information will not be used at all.
- Geometric Methods - Zolghadri[98] proposes a detection scheme that does the following. Two estimates of state and covariance are held. One is based purely on prediction. The other is found from a Kalman filter using all observations. The detection scheme is simple: if the uncertainty regions (covariances) overlap, then the filter is not at fault. The main problem with this technique is that a system running on prediction alone will eventually have infinite uncertainty. In this case, it doesn't matter what the Kalman filter estimates, because it will always overlap the prediction.
- Sensor Self-Validation[20] - There is an increasing push from industry for sensor man-

ufacturers to create sensors and actuators which diagnose themselves for faults. Although this is an admirable goal, it does not solve the problem of model faults.

- Neural Nets, Genetic Algorithms and Fuzzy Logic[18, 25] - These systems seek to ‘learn’ the patterns caused by faults. Typically, though, these methods are only used to classify or identify the fault after a fault event has been signalled by some other algorithm.

All of these techniques have some drawbacks. Many make assumptions that are untenable in the real world.

To create a *consistent* methodology for automated fault detection, the following criteria must be observed:

- The filter’s optimality (or ‘correct’ performance) should not be compromised under normal operation;
- The fault detection scheme should not be ‘ad hoc’. It must have a firm theoretical foundation; and
- Faults should not be modelled. By using a finite set of models, an infinite set is ignored.

The only scheme that satisfies all three of these criteria is innovations based detection.

As will be shown in Section 3.8, the innovations have certain statistical properties which may be tested for. These properties are voided if a fault is transmitted to the innovations. However, it is possible that certain faults may never be transmitted to the innovations (a fact that has perhaps caused the decline of research into this area).

Therefore, in the analyses that follow, faults are defined by their transmittance to the innovation. If a fault is not transmitted to the innovation sequence it will be defined as undetectable. Conversely, if a fault *is* transmitted to the innovations it will be defined as detectable. The following sections examine the detectability of faults within the Kalman filter with no assumptions of the type or even the source of the fault. Methods are examined by which *all* faults may be rendered detectable.

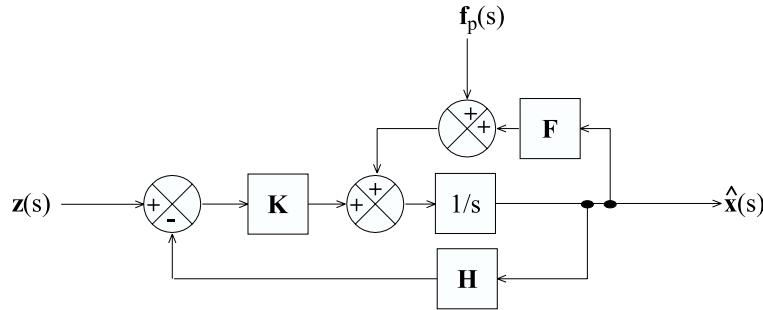


Figure 3.3: Kalman Filter Block Diagram with Injected Process Fault

3.4 Detecting Unmodelled Faults

3.4.1 Process faults

The case where faults are propagated through the Kalman filter via the process model[19, 46, 47, 79] will be examined first. As shown in Section 2.7.1, most vehicle models have the control inputs embedded in them as part of a nonlinear function. For this reason faults due to actuator failure may be thought of as propagating in the same way as faults due to the process (vehicle) model (of which they are a part).

3.4.2 Process Fault Detectability

Figure 3.3 shows the continuous time Kalman filter with an additive fault vector (of the same dimension as the state) $\mathbf{f}_p(s)$ injected at the process model.

To analyse this system, the injected fault is considered a disturbance, and the analysis is performed using classical control techniques. The input $\mathbf{z}(s)$ is therefore omitted in the analysis that follows.

The transfer function from the disturbance input to the output may be derived as

$$\begin{aligned} \frac{\hat{\mathbf{x}}(s)}{\mathbf{f}_p(s)} &= [s\mathbf{I} - \mathbf{F} + \mathbf{K}\mathbf{H}]^{-1} \\ &= \mathbf{T}_{pf}(s) \end{aligned} \quad (3.3)$$

where the subscript pf denotes the transfer function with a process fault.

Therefore, under process fault conditions

$$\hat{\mathbf{x}}(s) = \mathbf{T}_{pf}(s)\mathbf{f}_p(s) \quad (3.4)$$

The innovations with a process fault (and zero input) may now be defined as

$$\begin{aligned} \nu_{pf}(s) &= -\mathbf{H}\hat{\mathbf{x}}(s) \\ &= -\mathbf{H}\mathbf{T}_{pf}(s)\mathbf{f}_p(s) \end{aligned} \quad (3.5)$$

The term $-\mathbf{H}\mathbf{T}_{pf}(s)$ is the transfer function from the fault to the innovations

Now, in normal operation the input $\mathbf{z}(s)$ will be non-zero. According to the linear superposition principle, the response of the system (or in this case the response of the innovations) will simply sum. So, to find the response of the innovations subject to normal sensor operation together with a process fault (disturbance), the separate responses are simply added together. So the innovations under process fault conditions become the no-fault innovations summed with the innovations due to a process fault as

$$\nu_{pf}(s) = \nu_{nf}(s) - \mathbf{H}\mathbf{T}_{pf}(s)\mathbf{f}_p(s) \quad (3.6)$$

where the subscript nf denotes the no-fault condition.

Therefore, a process fault is detectable in the innovation sequence when the term $\mathbf{H}\mathbf{T}_{pf}(s)\mathbf{f}_p(s)$ is non-zero.

3.4.3 Sensor Faults

Sensor hardware faults have been given far more attention in the literature[35, 50, 78, 92] than actuator and process faults combined. There is a general assumption in most fault detection techniques that the only sensor faults that will occur arise from faulty hardware. This assumption is untenable in the real world. Sensor hardware faults may indeed make up a high proportion of the possible set of faults, but it is undeniable that sensor faults due to say bad design or modelling error also exist, and make up the remainder.

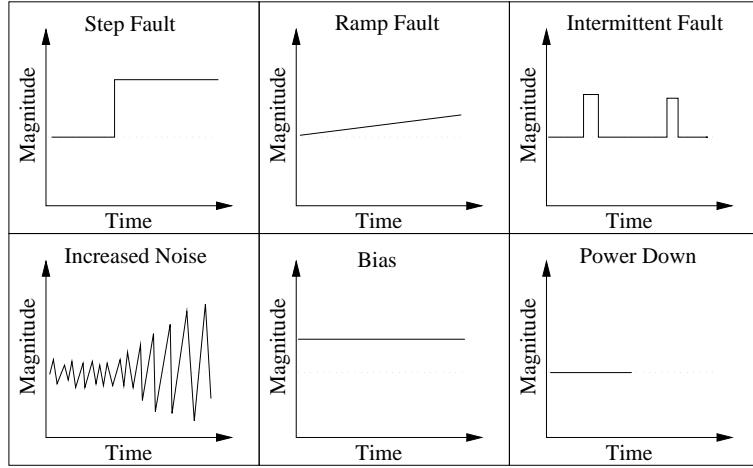


Figure 3.4: Some Common Sensor Hardware Failures

In the literature, sensor hardware faults are typically divided into the following categories[44, 81];

- step fault
- ramp fault
- intermittent fault
- increased measurement noise
- bias
- power down fault (no signal)

which may be seen graphically in Figure 3.4. The dotted line indicates the nominal (error free) signal.

Section 2.4 showed that the faults due to nonwhite noise processes such as the bias and ramp (or random walk) faults may be incorporated into the sensor model and thus be estimated by the Kalman filter. The other faults such as the step, power down and increased measurement noise faults tend to be more random in nature. Any of these fault models may closely resemble a valid sensor input, and indeed many sensor faults may have combinations of fault characteristics, and therefore not match any single model closely. Again, this is justification that the fault detection scheme should be model free.

3.4.4 Sensor Fault Detectability

Equation 2.13 gave the transfer between the estimate and the output as

$$\begin{aligned}\frac{\hat{\mathbf{x}}(s)}{\mathbf{z}(s)} &= [s\mathbf{I} - \mathbf{F} + \mathbf{K}\mathbf{H}]^{-1}\mathbf{K} \\ &= \mathbf{T}(s)\end{aligned}\quad (3.7)$$

which is constant once the filter has reached steady state.

Therefore, under no-fault conditions,

$$\hat{\mathbf{x}}_{nf}(s) = \mathbf{T}(s)\mathbf{z}_{nf}(s) \quad (3.8)$$

where the subscript nf again denotes the no-fault condition, and the subscript f denotes a fault condition.

Now, consider an observation corrupted by a fault vector $\mathbf{f}_o(s)$,

$$\mathbf{z}_f(s) = \mathbf{z}_{nf}(s) + \mathbf{f}_o(s) \quad (3.9)$$

From Equation 3.8, the state estimate becomes

$$\hat{\mathbf{x}}_f(s) = \mathbf{T}(s)[\mathbf{z}_{nf}(s) + \mathbf{f}_o(s)] \quad (3.10)$$

The no-fault filter innovations are defined as

$$\nu_{nf}(s) = \mathbf{z}_{nf}(s) - \mathbf{H}\hat{\mathbf{x}}_{nf}(s) \quad (3.11)$$

so, under fault conditions,

$$\begin{aligned}\nu_f(s) &= \mathbf{z}_f(s) - \mathbf{H}\hat{\mathbf{x}}_f(s) \\ &= [\mathbf{z}_{nf}(s) + \mathbf{f}_o(s)] - \mathbf{H}[\mathbf{T}(s)[\mathbf{z}_{nf}(s) + \mathbf{f}_o(s)]] \\ &= [\mathbf{z}_{nf}(s) - \mathbf{H}\hat{\mathbf{x}}_{nf}(s)] + [\mathbf{f}_o(s) - \mathbf{H}\mathbf{T}(s)\mathbf{f}_o(s)] \\ &= \nu_{nf}(s) + [\mathbf{I} - \mathbf{H}\mathbf{T}(s)]\mathbf{f}_o(s)\end{aligned}\quad (3.12)$$

Therefore, a sensor fault is detectable in the innovation sequence when the term $[\mathbf{I} - \mathbf{HT}(s)]\mathbf{f}_o(s)$ is non-zero.

It is interesting to note that this has a direct analogy with the no-fault innovation which may be written as $[\mathbf{I} - \mathbf{HT}(s)]\mathbf{z}(s)$. It may therefore be thought of as representing the effects of the sensors' frequency content on the innovations - particularly when the frequency content is dissimilar as in the case when a sensor is in fault.

The term $[\mathbf{I} - \mathbf{HT}(s)]$ is the transfer function from the sensor fault to the innovations.

3.4.5 Quantifying Process Fault Detectability

Example 3.4.1 Consider again the constant velocity model with an observation of position only (a Type I sensor), given by the following equations.

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \quad (3.13)$$

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + v(t) \quad (3.14)$$

The continuous time Kalman filter for this system is given by

$$\begin{aligned} \begin{bmatrix} \hat{\dot{x}}(t) \\ \hat{\ddot{x}}(t) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}(t) \\ \hat{\dot{x}}(t) \end{bmatrix} \\ &+ \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \left[\mathbf{z}(t) - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}(t) \\ \hat{\dot{x}}(t) \end{bmatrix} \right] \end{aligned} \quad (3.15)$$

giving the process fault transfer function (from Equation 3.3)

$$\frac{\hat{\mathbf{x}}(s)}{\mathbf{f}_p(s)} = \begin{bmatrix} \frac{s}{s^2 + K_1 s + K_2} & \frac{1}{s^2 + K_1 s + K_2} \\ \frac{-K_2}{s^2 + K_1 s + K_2} & \frac{s + K_1}{s^2 + K_1 s + K_2} \end{bmatrix} \quad (3.16)$$

Consider the case where the process model and observation covariance matrices are given

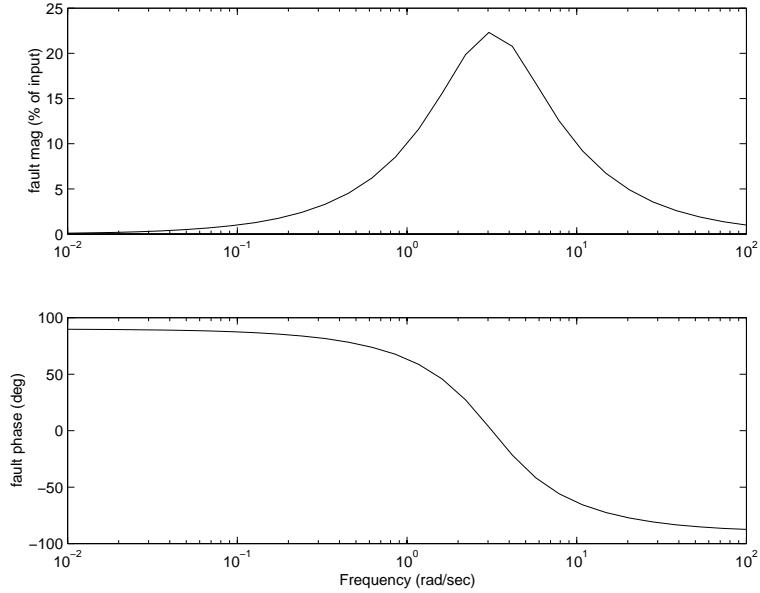


Figure 3.5: Fault Plot - Showing Percentage of Process Fault Transmitted to the Innovations and the Corresponding Phase for a Process Fault Injected into the Position State

by

$$\mathbf{Q}(t) = \mathbb{E}[w(t)w^T(t + \tau)] = [0.1]$$

and

$$\mathbf{R}(t) = \mathbb{E}[v(t)v^T(t + \tau)] = [0.001]$$

The transmittance of process faults to the innovation sequence (the term $\mathbf{HT}_{pf}(s)\mathbf{f}_p(s)$) is shown graphically in Figures 3.5 and 3.6. In this thesis, this type of figure will be referred to as a fault plot. The fault plot graphically displays the transmittance of a fault to the innovations as a percentage of the magnitude of the input fault and the corresponding phase.

Note that in both figures, there are regions where the fault is not transmitted to the innovations, and is therefore *undetectable*. This is obviously undesirable. However, for the *observed* state, the fault's transmittance to the *state* has exactly the same form as the transmittance to the innovations. This means that faults that are undetectable in the innovation sequence are *not* transmitted to the observed state and may therefore be ignored. This may be deduced from the error term $\mathbf{HT}_{pf}(s)\mathbf{f}_p(s)$ which simply maps the fault to the states, before being mapped to observation space by the \mathbf{H} matrix.

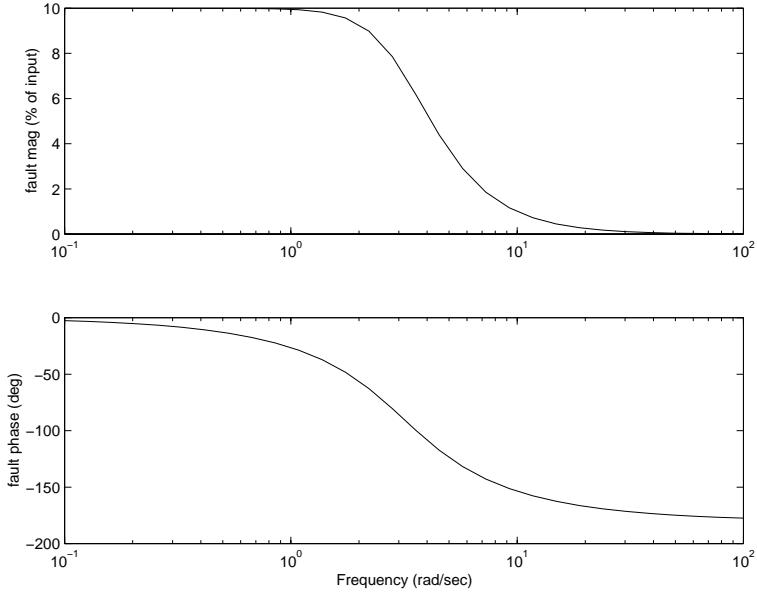


Figure 3.6: Fault Plot - Showing Percentage of Process Fault Transmitted to the Innovations and the Corresponding Phase for a Process Fault Injected into the Velocity State

It can be seen from Equation 3.3 that the transfer functions from the process fault to the estimated state all have the same poles (following the argument presented in Section 2.2.2) and that the transfer functions are all low pass. This means that process faults at frequencies above the cutoff ω_c , are increasingly attenuated. So for faults at frequencies much greater than the cutoff, faults are neither transmitted to the innovations nor to the states. So high frequency process faults have *no* effect on filter performance.

In the example presented above, however, there is a class of process faults which *cannot* be detected. Consider the first case presented where the fault is introduced into the position state only. It can easily be seen that the transfer from the fault to the innovations is identical to the transfer from the fault to the position state. This means that a process fault affecting the position state is either detectable in the innovation sequence, or when it is not, has no effect on the position state and is therefore not a threat to filter robustness. However, the transfer from the fault to the velocity state is not the same (in this case it is low pass rather than band pass), and therefore faults at low frequencies are transmitted to the velocity state. This may be conceptualised as being due to the velocity state being *unobserved*. There is no way for the filter to check the consistency of the velocity estimate, since it is inferred from a position observation.

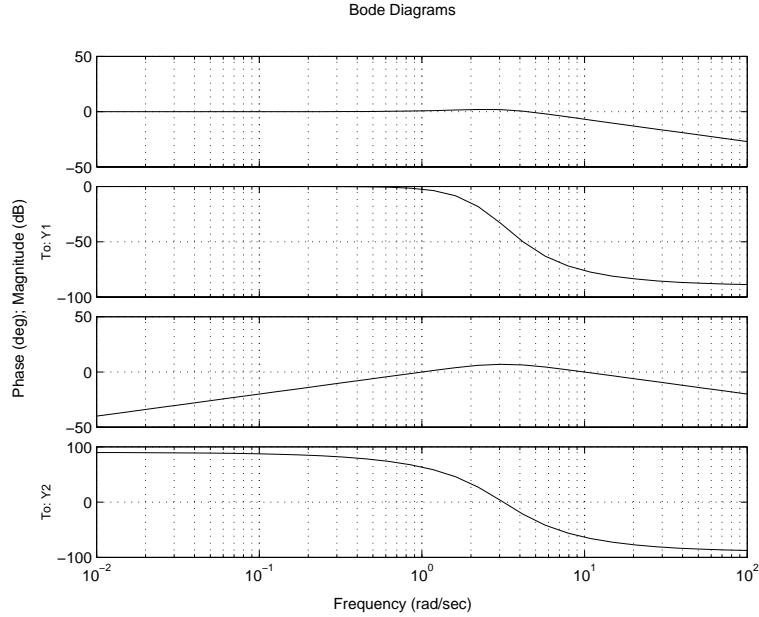


Figure 3.7: Bode Plot of the Filter Transfer Function - Top Two Subplots are the Magnitude and Phase Plots from the Sensor to the Estimate \hat{x} - The Bottom Two Subplots are the Magnitude and Phase Plots from the Sensor to the Estimate $\hat{\hat{x}}$

All process faults that affect the estimate of observed states are detectable. However process faults that occur below the filter cutoff frequency ω_c may be transmitted to the unobserved states without detection.

3.4.6 Quantifying Sensor Fault Detectability

Example 3.4.2 Continuing the example presented in Section 3.4.5 of a constant velocity process model, with an observation of position only, the transfer function for the system is (from Equation 3.7)

$$\frac{\hat{\mathbf{x}}(s)}{\mathbf{z}(s)} = \begin{bmatrix} \frac{sK_1+K_2}{s^2+K_1s+K_2} \\ \frac{K_2}{s^2+K_1s+K_2} \end{bmatrix} \quad (3.17)$$

This example will use the process model and observation covariance matrices given in Section 3.4.5. The Bode plot of the transfer function (at steady state) is shown in Figure 3.7.

Figure 3.8 is a graphical representation of the function $[\mathbf{I} - \mathbf{HT}(s)]\mathbf{f}_o(s)$ as a percentage of the sensor fault transmitted to the innovations along with the transmitted fault's phase angle

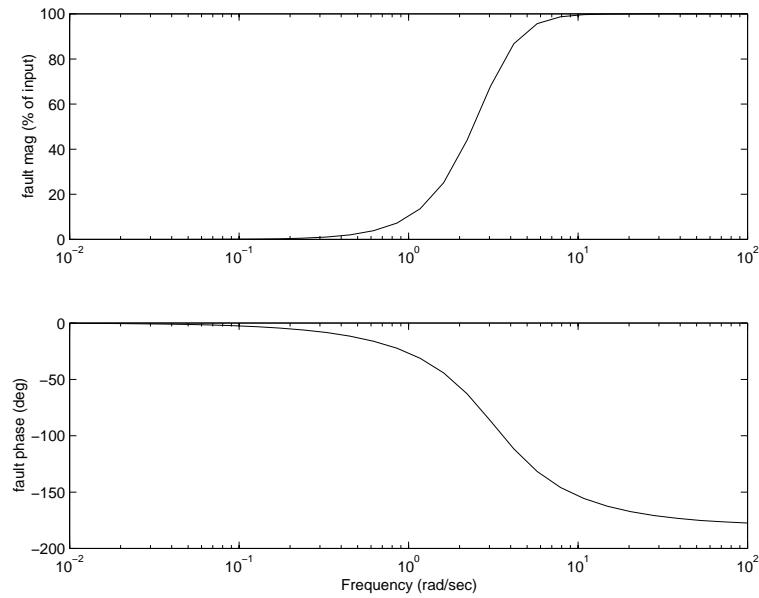


Figure 3.8: Fault Plot - Showing Percentage of Sensor Fault Transmitted to the Innovations and the Corresponding Phase

versus frequency. Note that at frequencies below about 0.2 rad/s the fault is *undetectable* in the innovation sequence, but is transmitted *without attenuation* to the state estimate.

In this case, if the fault can be characterised as a predominantly low frequency event (e.g. drift, bias, random walk), then the filter will *track* the fault and eventually diverge. If the fault is a predominantly high frequency event (e.g. step fault) then the fault will be transmitted to the innovations and is thus detectable.

In general, the sensor fault transmittance to the innovations is the *complement* of the transfer functions from the sensors to the observed states.

3.4.7 Summary

To summarise, this section has developed quantifiable measures of fault detectability. The detectability is defined in terms of a faults transmittance to the Kalman filter innovations.

The following boxes summarise the results for process faults. The innovations at fault are given together with the transfer function from the fault to the innovations. The box following summarises the physical significance of the derived equations.

Process Faults

The Innovations at Fault

$$\nu_{pf}(s) = \nu_{nf}(s) - \mathbf{HT}_{pf}(s)\mathbf{f}_p(s)$$

The Fault Transfer Function

$$-\mathbf{HT}_{pf}(s)$$

All process faults that affect the estimate of observed states are detectable. However process faults may be transmitted to the unobserved states without detection.

The following boxes summarise the results for sensor faults.

Sensor Faults

The Innovations at Fault

$$\nu_f(s) = \nu_{nf}(s) + [\mathbf{I} - \mathbf{HT}(s)]\mathbf{f}_o(s)$$

The Fault Transfer Function

$$[\mathbf{I} - \mathbf{HT}(s)]$$

The sensor fault transmittance to the innovations is the complement of the transfer functions from the sensors to the observed states.

This means that the general form of sensor fault transmittance to the innovations may be predicted from the insights gained in Section 2.9. i.e if a sensor has a low pass transfer function, then that sensors fault transmittance to the innovations will be high pass in nature.

It should be noted here that this technique makes no assumptions as to the source of the fault. This technique will work regardless whether the fault is a modelling error or a ‘hard’

sensor failure. The only classification that is made about a fault is its frequency content. The frequency content of a fault may be thought of as a fault signature.

Also, due to the fault representation as a vector, the techniques developed in this thesis allow the study of multiple faults by simply adding or subtracting fault elements to the fault vector.

3.5 Guaranteeing Process Fault Detection

Example 3.5.1 Consider again the constant velocity model example. It was shown that process faults that affect the state estimates can only be guaranteed to be detectable in states that are directly observed. If the observation model described in the Section 3.4 is now augmented to contain a velocity sensor,

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \mathbf{v}(t) \quad (3.18)$$

with observation covariance

$$\mathbf{R}(t) = \mathbb{E}[\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.01 \end{bmatrix}$$

then both of the estimated states are directly observed.

Figures 3.9 and 3.10 show the transmittance of position state and velocity state process faults to the innovations.

The figures show a general low pass nature for the fault transmittance to each of the states. This means that all low frequency faults are transmitted to the innovations, regardless of the state that is affected by the fault. However, when the arguments of Section 3.4 are considered, that is, that process faults above the cut-off frequency are not transmitted to the state estimate *or* to the innovations, then the figures show that in this case *all* process faults that affect the state estimate are detectable as desired.

Therefore, to guarantee process fault detectability, the system must be observable

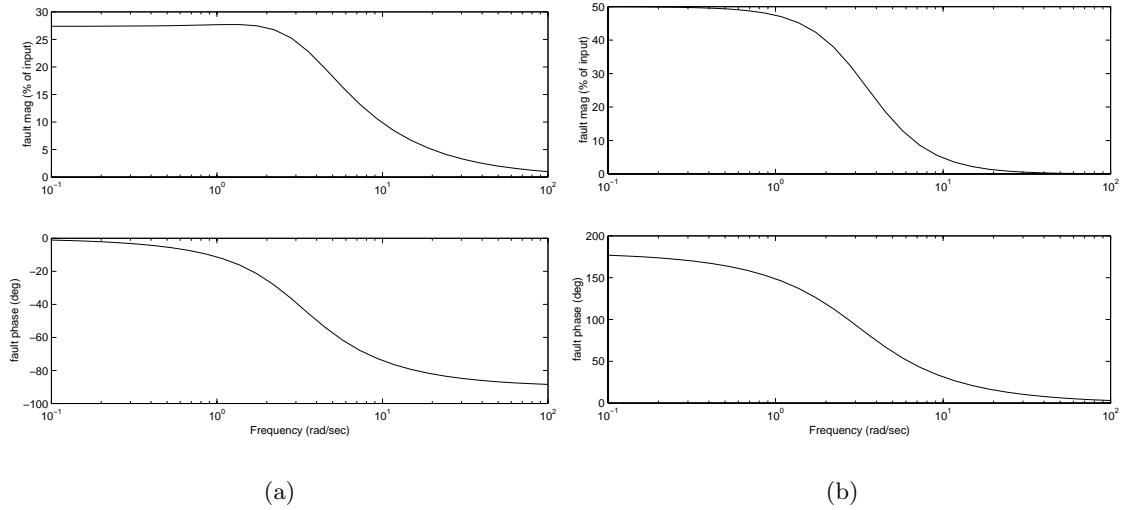


Figure 3.9: Fault Plot - Showing Percentage of Process Fault Injected into the Position State Transmitted to the (a) Position Innovation and (b) the Velocity Innovation

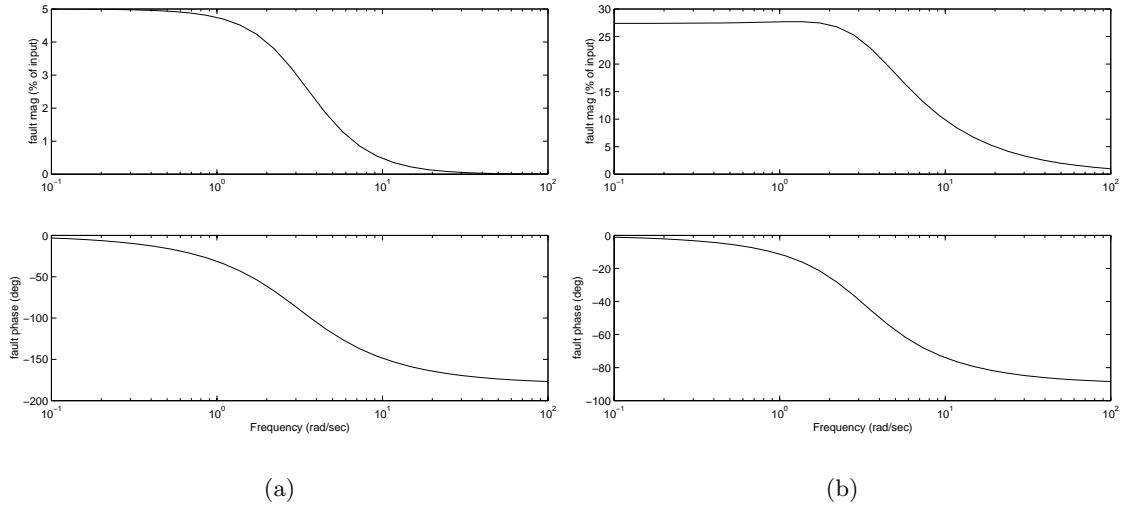


Figure 3.10: Fault Plot - Showing Percentage of Process Fault Injected into the Velocity State Transmitted to the (a) Position Innovation and (b) the Velocity Innovation

3.6 Guaranteeing Sensor Fault Detection

3.6.1 Redundancy

To reduce or eliminate the frequency band where sensor faults are undetectable, redundancy must be employed. The most common form of redundancy involves the use of matched sensors. The idea is that if one sensor fails, the other (one or more) sensor readings will not agree with the failed sensor.

Example 3.6.1 Consider the case where two exactly matched sensors take measurements of the state x (using the example from Section 3.4), giving an observation equation of the form

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \mathbf{v}(t) \quad (3.19)$$

with observation covariance

$$\mathbf{R}(t) = \mathbb{E}[\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} \quad (3.20)$$

The bode plot from the sensors to the states is shown in Figure 3.11. Note that the plot is representative of the transfer from both sensors to the states.

If a single sensor is in fault, the fault magnitude and phase transmitted to the innovations is shown in Figure 3.12. This clearly shows that a fault in one sensor only is *always* detectable in the innovation sequence. However, if the sensors display *identical* fault characteristics, the fault(s) *cannot* be detected if they contain only low frequency information. This can be seen in Figure 3.13.

It is easy to see from these results that if one assumes that only one sensor will fail in a particular way at any time, then with matched sensors the fault will be detectable in the innovations. If, however, the sensors exhibit the same fault behaviour at the same time, which it is not unreasonable to expect exactly matched sensors to do, then low frequency faults become undetectable.

Considering this example, sensor redundancy is clearly a good idea if the entire spectrum of fault types is to be detectable in the filter innovations.

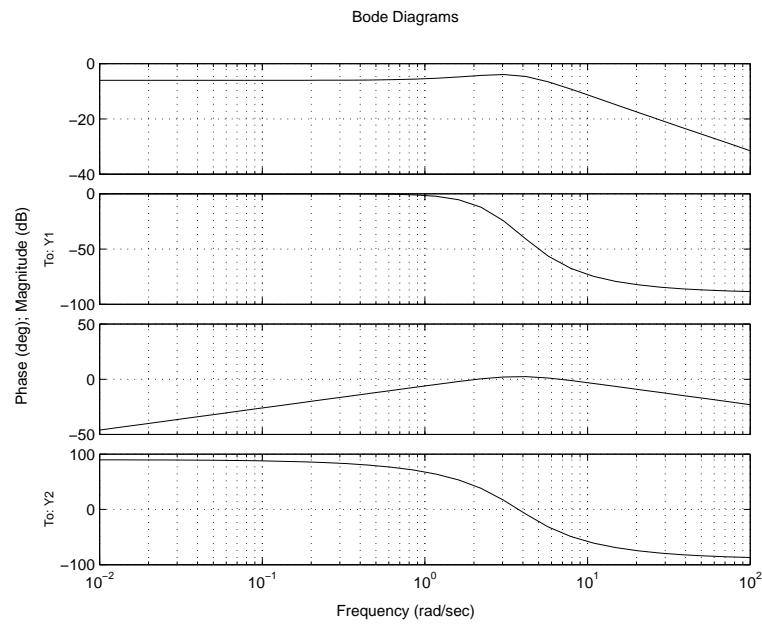


Figure 3.11: Bode Plot of the Filter Transfer Function - Identical for Both Sensors

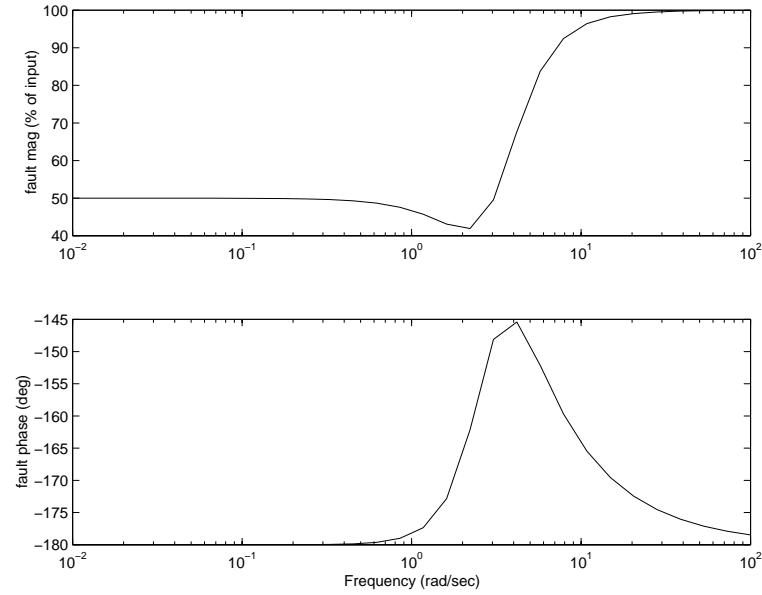


Figure 3.12: Fault Plot - Fault in One Sensor Only

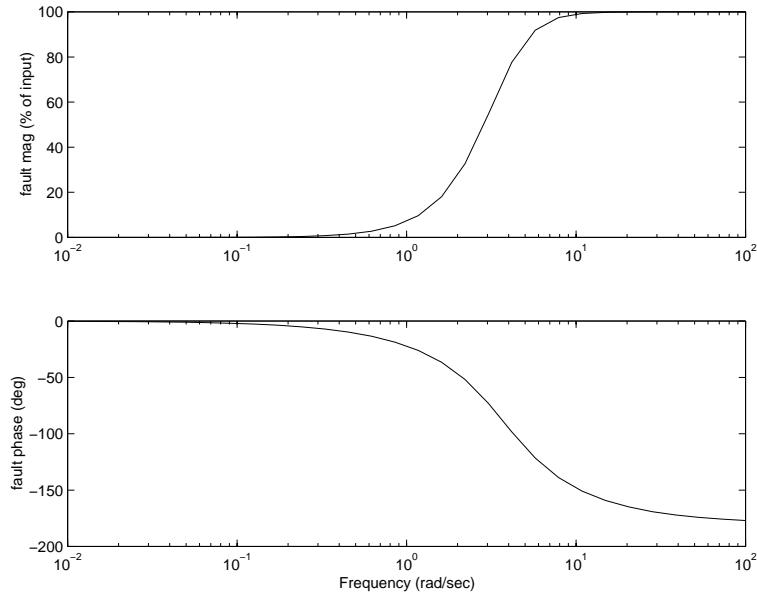


Figure 3.13: Fault Plot - Identical Faults in Both Sensors

The choice of identical sensors for redundancy is *not* encouraged however, as it is likely that identical sensors will exhibit identical failure modes, and thus render certain faults undetectable, negating the designer's ability to guarantee that faults will be detected.

This prompts a rethink of the design of redundant systems.

3.6.2 Frequency Domain Redundancy

The previous example shows that redundant matched sensors guarantee sensor fault detectability if only one sensor is in fault. The reason for this is that the two sensors supply redundant information, as the frequency spectrum of each sensor is duplicated. Therefore, if a fault occurs at any frequency or more usually at any group of frequencies, then the information supplied by the sensors is contradictory and the fault is thus detectable. Contrast this with the case where, because there is no redundant information, there is no way to tell if the sensor is performing correctly or not (at least in the low frequency range).

From this insight, it should be clear that to provide redundancy, which is the *only* way to guarantee sensor fault detectability, one only needs to use sensors that provide information over similar frequency ranges. Identical sensors are not needed for fault detection.

Indeed, as highlighted by the previous example, the redundant sensors should ideally be based on different physical principles and technologies so as to minimise the chance of any group of sensors displaying the same fault characteristics.

This reduces the need to place the restriction that fault detection is guaranteed if only one sensor is at fault. With the technique presented in this thesis, provided that the sensor fault modes are different, the detection of *multiple* sensor faults is guaranteed.

3.6.3 An Example of Frequency Domain Redundancy

By using sensors based on very different physical principles (for this example a sonar and a laser) to provide frequency domain redundancy, the fault modes of the sensors are likely to be very different. This means that for the unlikely case where both sensors are at fault, the nature of the faults will most likely be different enough that a fault will still be detected. This example seeks to illustrate this concept.

This example will also show that a sensor fault need not be a physical breakdown of the sensor, but that it is just as likely that inadequacies in the sensor model result in apparent faults.

Example 3.6.2 Consider the case where a laser range finder and a sonar sensor are used to map the room shown in Figure 3.16. If the sensors are modelled as simply returning the distance to the nearest object in the room in the direction they are pointing, then the state of interest x is just the distance from the sensor to the wall. As the room is assumed static, the constant velocity model described in the previous example may be used (see Equation 3.13). The sensor model for this example is given by;

$$\mathbf{z}(t) = \begin{bmatrix} z_{sonar} \\ z_{laser} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \mathbf{v}(t) \quad (3.21)$$

with observation covariance (in mm²)

$$\mathbf{R}(t) = \mathbb{E}[\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = \begin{bmatrix} 0.1 & 0 \\ 0 & 1 \end{bmatrix} \quad (3.22)$$

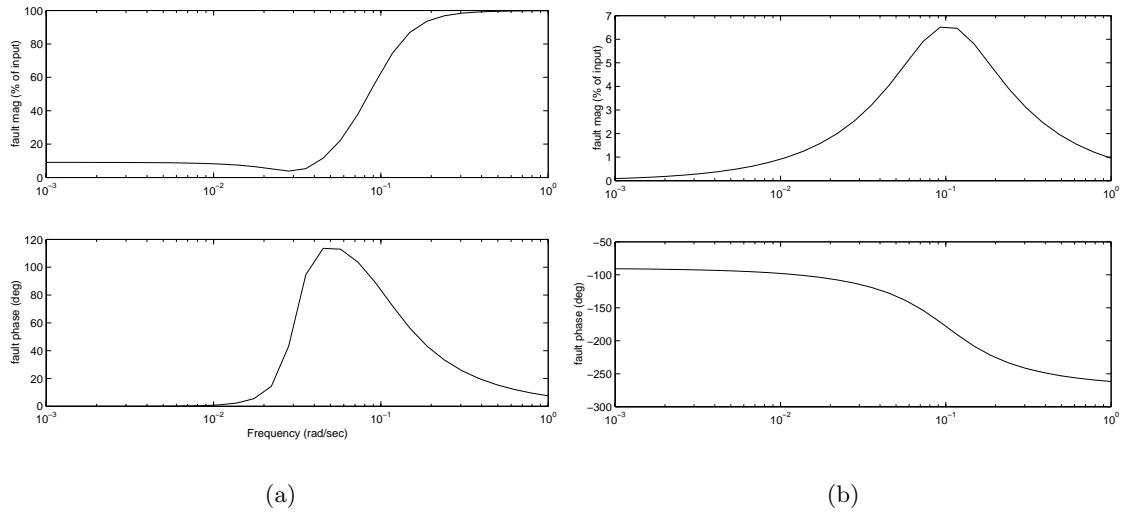


Figure 3.14: Fault Plot - Percentage of Sonar Fault Transmitted to (a) Sonar Innovations and (b) the Laser Innovations

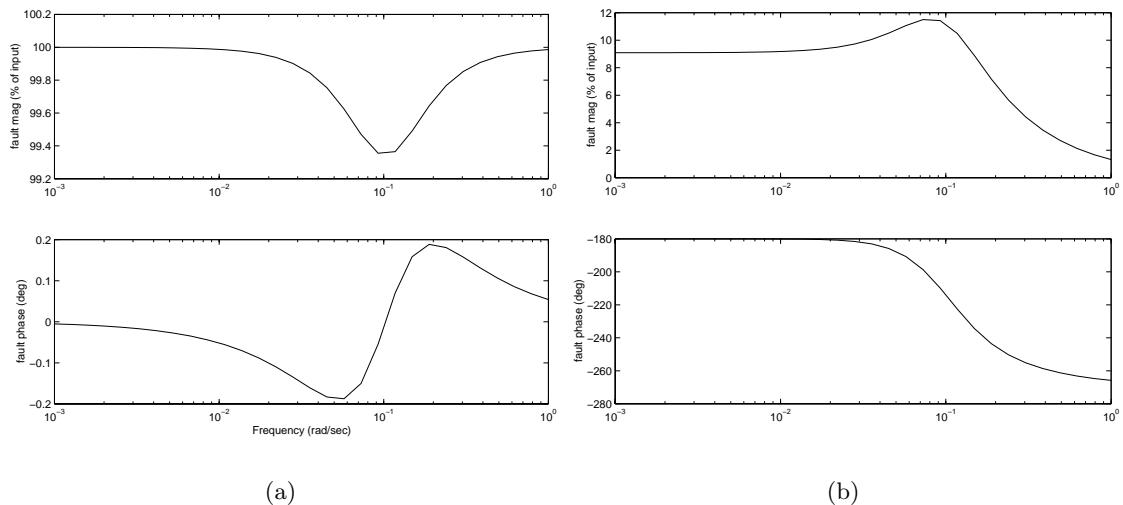


Figure 3.15: Fault Plot - Percentage of Laser Fault Transmitted to (a) Laser Innovations and (b) the Sonar Innovations

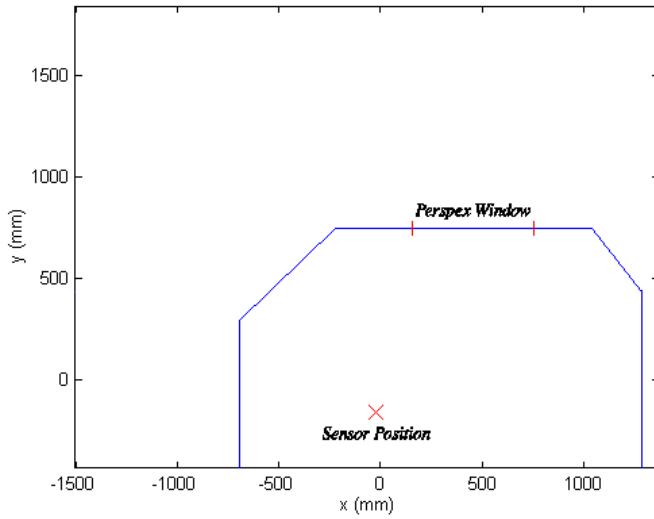


Figure 3.16: Map of the Sensor Trial Area

The fault detectability plots for this system are shown in Figures 3.14 and 3.15.

It can be seen that as in the case of exactly matched sensors, *all* faults can be detected in this system. Note that faults in the poorer quality sensor (in this case the laser) have a much higher percentage of the fault transmitted to the innovations. This is due to the filter placing significantly higher weight on the information from the better quality sensor.

Now, the models employed for the two sensors (Equation 3.21), are quite naive given that the sensors, the sonar in particular, do not simply return the distance to the nearest obstacle. The sonar will often display a bias when scanning walls, as the edge of the sonar beam reflects from the wall before the central part. This is known as a Region of Constant Depth (RCD) [57]. Other sonar properties include specular and diffuse reflections. These effects may be seen in Figure 3.17 which is a plot of sonar distance measurements as the sensor rotates through 180 degrees in 0.9 degree steps. The laser range sensor fits the environment model much better than the sonar, although complex environmental interactions such as multiple path reflections and laser transparent surfaces such as glass and perspex are not modelled. The laser map of the room can be seen in Figure 3.18: note that the laser data more accurately reflect the true state of the room except in the area of the perspex window.

If Figures 3.16, 3.17 and 3.18 are considered, it should be evident that if the sonar and laser data are fused with a Kalman filter described by the observation equations given above,

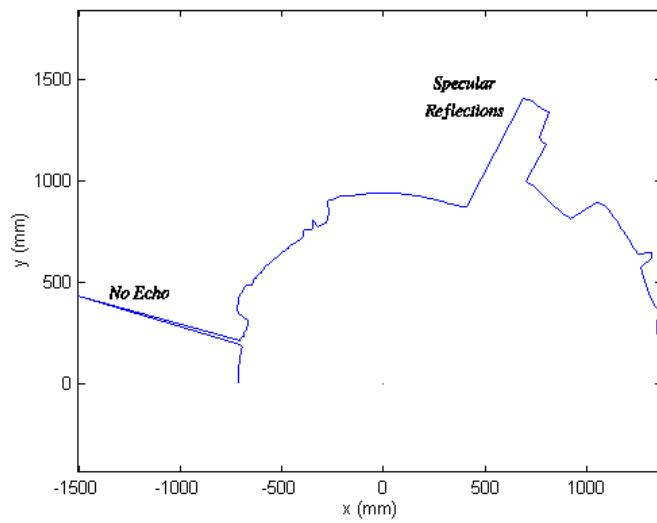


Figure 3.17: Raw Sonar Data

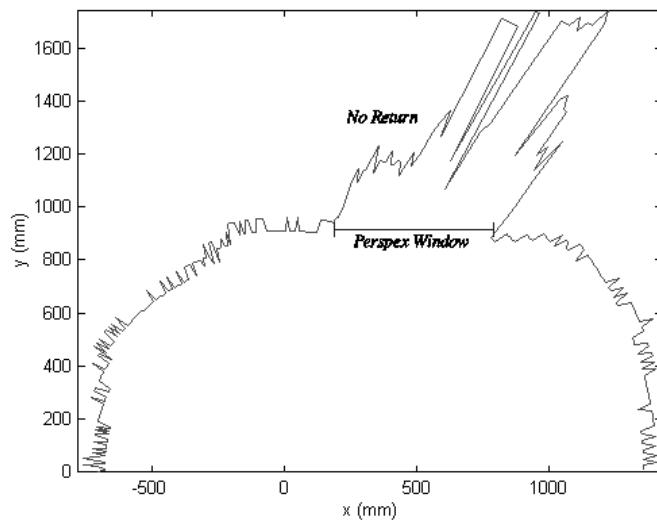


Figure 3.18: Raw Laser Data

the system will be inconsistent, due mainly to the inadequacy of the sonar sensor model. Thus, the information supplied by the two sensors is conflicting in many areas. For example where the sonar times out or there are specular reflections, the laser information is good. In areas where glass or perspex is present, the laser will return an incorrect measurement, and the sonar is far more accurate. Even in the case in the top right hand corner of the room where *both* sensors are at fault, the laser seeing through the perspex and the sonar receiving specular reflections, the faults will be detectable in the innovations of the filter because the *magnitude* of the fault is different.

Example 3.6.3 Consider the sensors positioned such that they are both facing due east measuring the perpendicular distance to the easternmost wall. Each sensor takes one hundred measurements of the range to the wall. For the filter described by Equation 3.21, the estimated range to the wall, using both sensors, is given by Figure 3.19. The corresponding innovation sequences for the laser and sonar are given by Figure 3.20.

It can be seen that even in this very simple case of perpendicular distance measurement, the filter is not consistent.

As can be seen in the laser innovations, there appears to be a large amount of quantisation noise (a characteristic not shared by all laser range sensors) which voids the basic assumption used in the derivation of the Kalman filter of measurements being corrupted by white noise only. This quantisation effect has not been modelled. The one hundred sonar samples were taken over a period of two hours. The temperature of the ambient air changed during this time causing the sonar range measurement to drift. This effect can be seen in the sonar innovations.

Because of these unmodelled effects, both innovation sequences are clearly non-white over the duration of the experiment. These inadequacies with the sensor models clearly lead to a system which is always at fault.

If only a single sensor had been used in this example, either the sonar or the laser, there would be no possible way for a filter to distinguish between a true measurement of the environment and a sensor fault, and if two identical sensors had been used the faults in each sensor would also be identical and thus be undetectable.

Therefore by using sensors based on different physical principles, all sensor faults have been rendered detectable, even in the case where both sensors fail simultaneously.

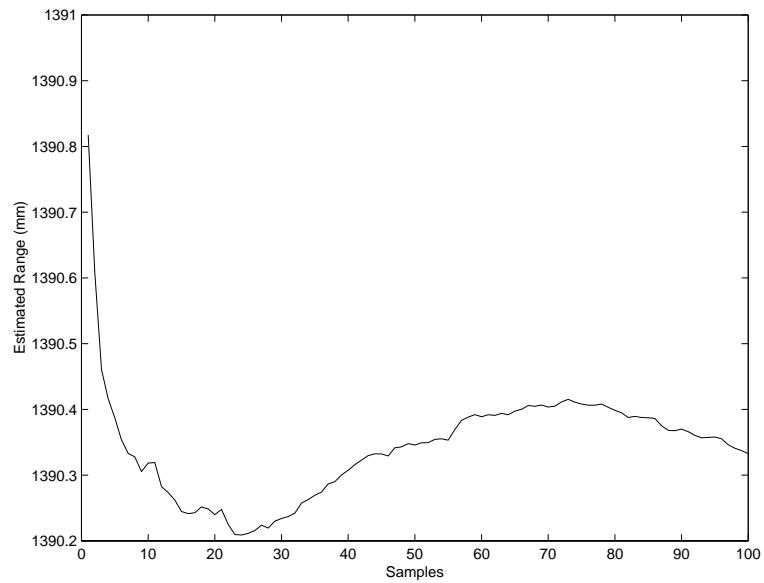
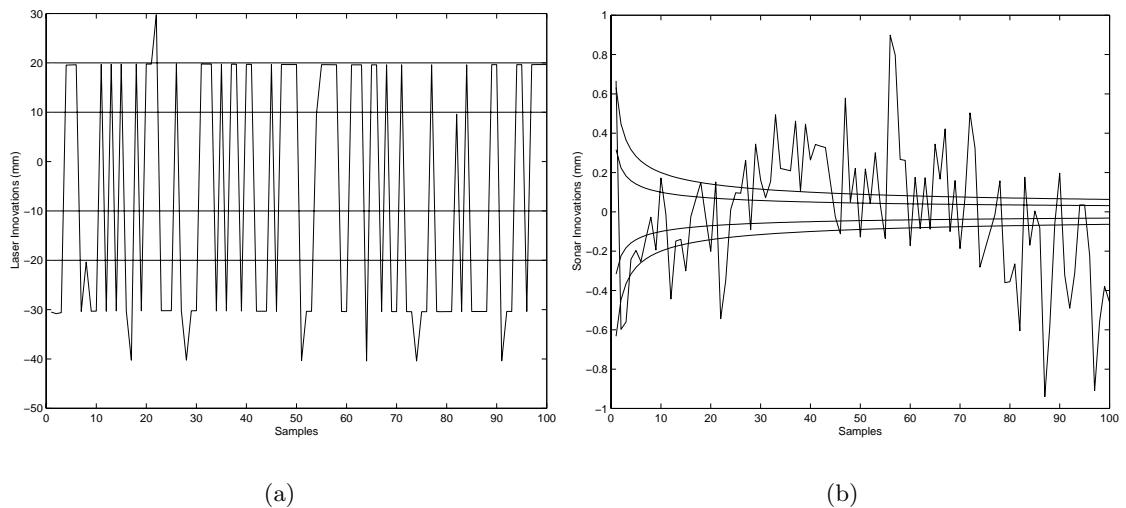


Figure 3.19: Estimated Range

Figure 3.20: (a) Laser Innovation Sequence with 1σ and 2σ bounds and (b) Sonar Innovation Sequence with 1σ and 2σ bounds

This example also showed that a sensor fault is as likely to be caused by the filter designer, by not providing an adequate model of the sensor's interactions with the environment, as by a hardware failure.

3.7 Nonlinear and Time Varying Systems

Within the Kalman filter framework, nonlinearities may manifest themselves in the process and observation models. In the extended Kalman filter, linear models are essentially replaced by linearised nonlinear models. If the linearised models are not close enough to the true models as to have the error resemble white noise, then these will be detectable as process or sensor faults as outlined in the previous sections.

Consider the example given in Section 2.10 of a mobile robot with nonlinear process and observation models. It was shown that a correctly functioning EKF is functionally equivalent to a time varying linear system. Further, this example showed that for typical vehicle and observation models used in navigation, the ‘family’ of transfer functions for the system are bounded. This means that the fault detectability metrics described in Section 3.4 will also consist of a family of bounded systems for nonlinear and time varying linear systems.

Completing the example started in Section 2.10, the sensor fault detectability plots for the range and bearing sensors are given by Figures 3.21, 3.22, 3.23 and 3.24. These plots were found by constructing the fault plot for every member of the family of related transfer functions.

It can be seen that the general shape of the fault plots is identical to the time invariant linear case. As predicted, the family of bounded transfer functions has given rise to a family of bounded fault plots. Interestingly, at very low frequencies and at very high frequencies, the fault plots for the EKF converge to a single value. It is only the mid frequencies whose values vary, changing the point at which a fault may be rendered detectable.

To guarantee fault detectability in nonlinear and time varying systems, one must be able to prove that the entire family of fault plots is non zero over the entire frequency spectrum. In practise this may be done simply and efficiently by taking the worst case transfer functions, i.e. the upper and lower bounds of the transfer, and constructing the equivalent fault plot for the system. If the worst case fault plots both remain non-zero, then one may safely

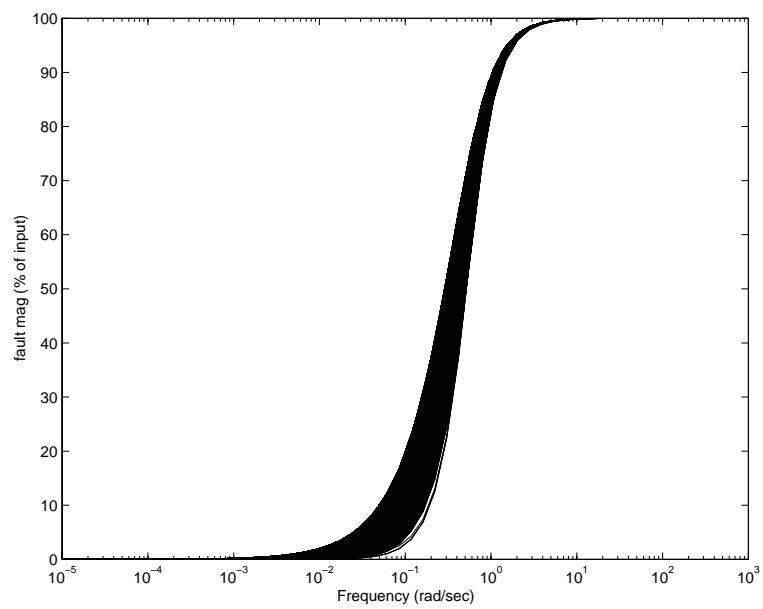


Figure 3.21: Fault Plot - Percentage of Range Measurement Fault Transmitted to Range Measurement Innovations

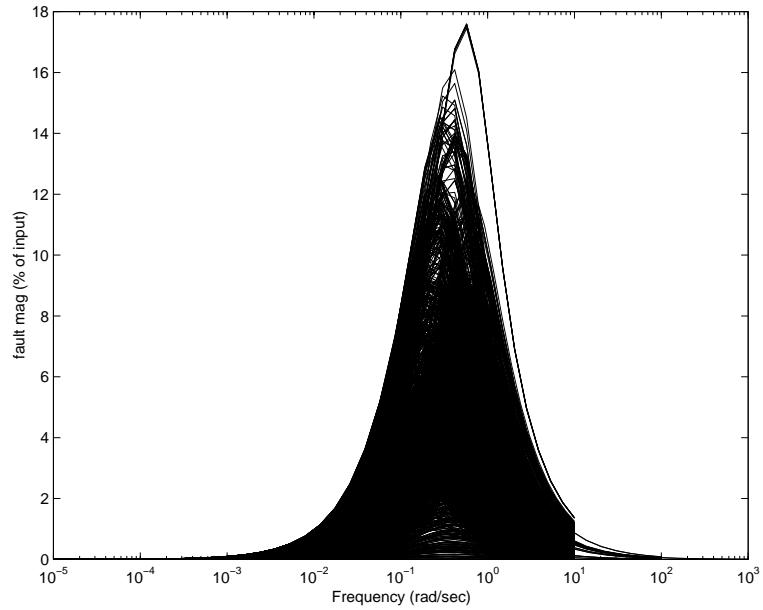


Figure 3.22: Fault Plot - Percentage of Range Measurement Fault Transmitted to Bearing Measurement Innovations

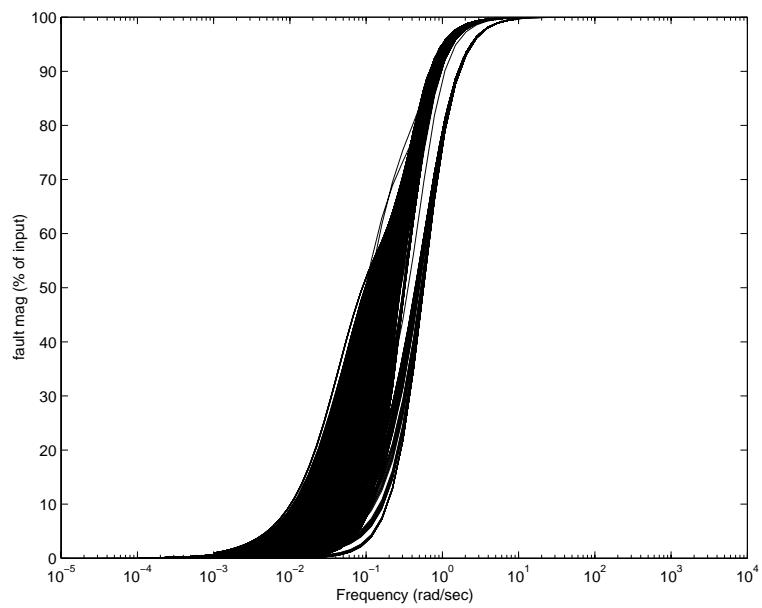


Figure 3.23: Fault Plot - Percentage of Bearing Measurement Fault Transmitted to Bearing Measurement Innovations

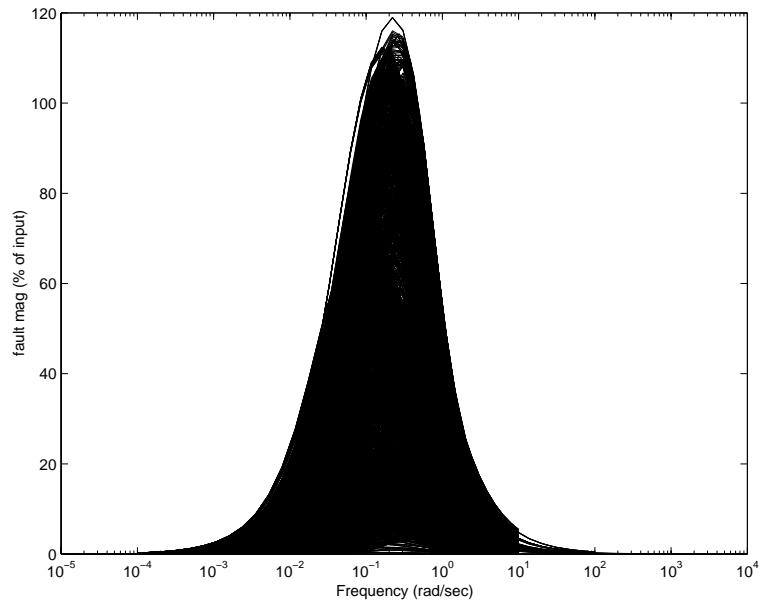


Figure 3.24: Fault Plot - Percentage of Bearing Measurement Fault Transmitted to Range Measurement Innovations

assume that the entire family of fault plots is non-zero, and thus all faults are detectable.

The technique of using upper and lower bounds of a transfer function for a worst case analysis is commonly used in \mathcal{H}_∞ control.

It is reasonable to assume that the general ideas presented to guarantee process and sensor fault detectability still apply to nonlinear and time varying systems provided the transfer functions of the system remain bounded. However, to provide a guarantee that all faults are detectable in these systems, the worst case analysis described above must be performed on a case by case basis.

3.8 Fault Detection

Chapter 2 described how particular model faults can be addressed and incorporated into the Kalman filter framework. These faults are detected by the Kalman filter itself, which estimates important unobserved parameters, such as the coloured noise properties of sensors, or the wheel slip of a vehicle. So, from the design point of view, simply by including important sources of error into the filter, the detection of such faults essentially takes care of itself.

But, if the models are wrong, or a sensor fails, the power stops, or if any one of a myriad of unforeseen events take place, how can these failures be detected? The answer is simply to test the innovations. The two main advantages of innovations testing are

- the detectability of faults in the innovations may be ascertained using the theory presented in the previous sections, and
- existing recursive time domain whiteness tests can be used.

A high integrity navigation system must have two important properties. Faults have to be guaranteed to be transmitted to the innovations, and the innovations must be tested for faults.

Appendix C presents the theory used to test the innovations for faults. Put simply, the innovations are zero mean white sequences with known variance. Innovations testing techniques exploit this fact to detect faults by determining on-line if the the innovations indeed have these statistics.

3.9 Conclusion

This chapter has examined the issue of detecting unmodelled faults. The analysis provided led to the proposal of methods which effectively guarantee that all faults will be detectable.

Section 3.2 examined the issue of fault detection with respect to high integrity navigation. It was argued that fault detection is a necessary part of any high integrity navigation system and furthermore, that some quantifiable measure of the detectability of faults must also be used. In Section 3.3, a survey of current methods of fault detection was given. It was concluded that innovations testing should be used for the detection of faults. To establish the detectability of faults, Section 3.4 derived a metric which allows the determination of the detectability of process and observation faults. This metric is based in part on the transfer function from the fault to the innovation sequence. The concept of a fault plot was introduced as a graphical method of quantifying the detectability information. From the insights gained in the development of this theory, Sections 3.5 and 3.6 were able to propose methods which effectively guarantee that all process and sensor faults are detectable. In Section 3.6 arguments were presented for the use of redundancy to guarantee fault detection, and a new term was introduced, that of frequency domain redundancy. The detectability of faults in Nonlinear and time variant systems were examined in Section 3.7. It was argued that bounded nonlinear and time variant systems can be treated similarly to linear systems in terms of fault detection and design. Finally in Section 3.8, a simple recursive method for checking filter consistency was given.

Chapter 4

High Integrity Design

4.1 Introduction

This chapter explores the issue of the design of high integrity navigation systems for AGVs. In particular, the theory presented in the previous chapters is utilised to provide a methodology for the design of high integrity navigation systems that achieve a given specification and are robust to failure.

In Section 4.2, a typical navigation system design specification is presented, and interpreted in terms of the Kalman filter. Section 4.3 examines the use of different sensor suites to achieve this specification. Sensor suite design is generalised in Section 4.4, and methods given to enable sensors to be chosen which minimise the cost of the sensor suite, whilst maintaining the specified performance level.

Section 4.5 introduces the concept of a high integrity navigation system, and utilising theory from the previous chapter, arguments are presented as to why sensor redundancy is necessary. This section also includes alternate Kalman filter structures which further improve the integrity of the system. A discussion of the information filter algorithm which is ideal for distributed systems is presented, and arguments are given as to why the theory presented in the previous chapters apply equally well to the information filter.

4.2 The Specification of Navigation Systems

To design a commercially viable navigation system, one must not only guarantee that all faults in the system will be detectable, but also that the system will meet certain design requirements.

The design requirements, however, impose some serious constraints that require a formal method of navigation system design to be developed. These are:

- Sensors must be chosen so as to minimise cost,
- Sensors must be chosen to achieve the design specification, and
- The system must be designed for high integrity.

The design specification will usually consist of:

- The highest frequency manoeuvre the vehicle is capable of, and
- The accuracy the navigation system needs to achieve in order to complete the desired task.

These specifications may be interpreted as being the highest frequency that the Kalman filter must track, and the associated steady state covariance.

Figure 4.1 graphically represents the frequency response design specification. It is desired that the filter provides good information at all frequencies up to the cut-off frequency for the state(s) of interest, which in this case occurs at approximately 20rads^{-1} . It was shown in Chapter 2 that the Kalman filter is always low pass, so it is known that signals above the cut-off frequency will be increasingly attenuated. This method of frequency domain ‘bounding’ is described in [21] as an aid to sensor suite selection.

For vehicle navigation systems, the accuracy the system must achieve is typically measured in terms of the vehicle’s position. In other words, the vehicle must be able to achieve a desired position within a desired tolerance. This is usually because the vehicle must perform an operation such as loading or unloading which requires the vehicle to be at a given position with known accuracy. In this case, the position standard deviation, which is a measure of the position accuracy, can be easily calculated from the associated covariance matrix.

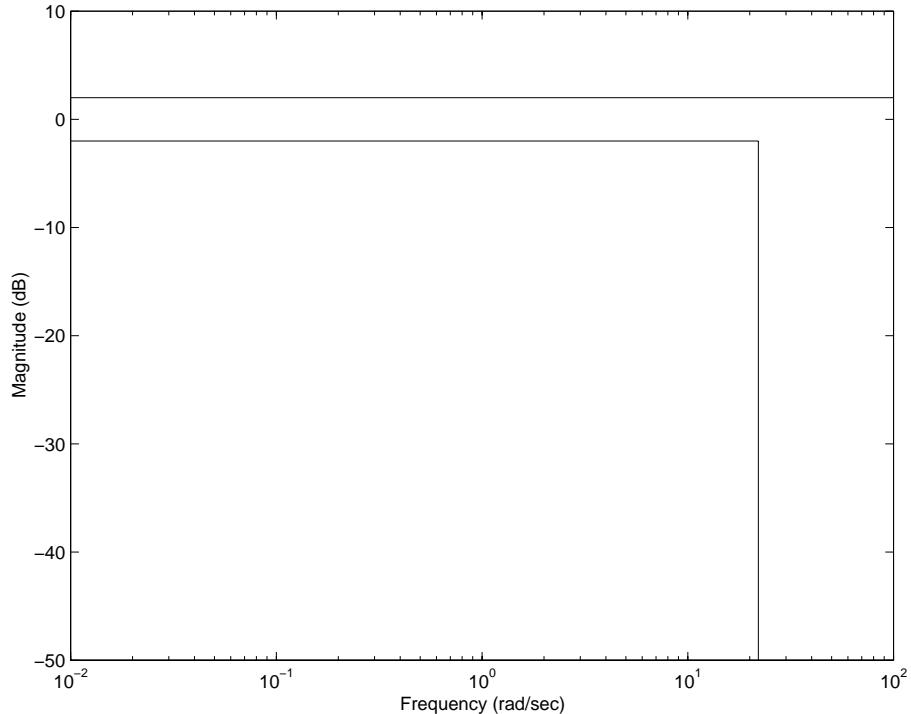


Figure 4.1: Bound Plot Showing Desired Frequency Response

4.3 Achieving the Specification

It is not unreasonable to assume that the target system (or vehicle) will essentially dictate the form of the process model. For this reason in the sections that follow, it is assumed that the designer has the following information

- The form of the process model is known
- The desired frequency response of the system is known
- The desired system accuracy is known

The last two points are, of course, the specification that the navigation system must meet.

4.3.1 Sensor Suites

As discussed in Section 2.4, many common navigation sensors output signals that are corrupted by non-white (or coloured) noise. For these sensors, which typically have poor low

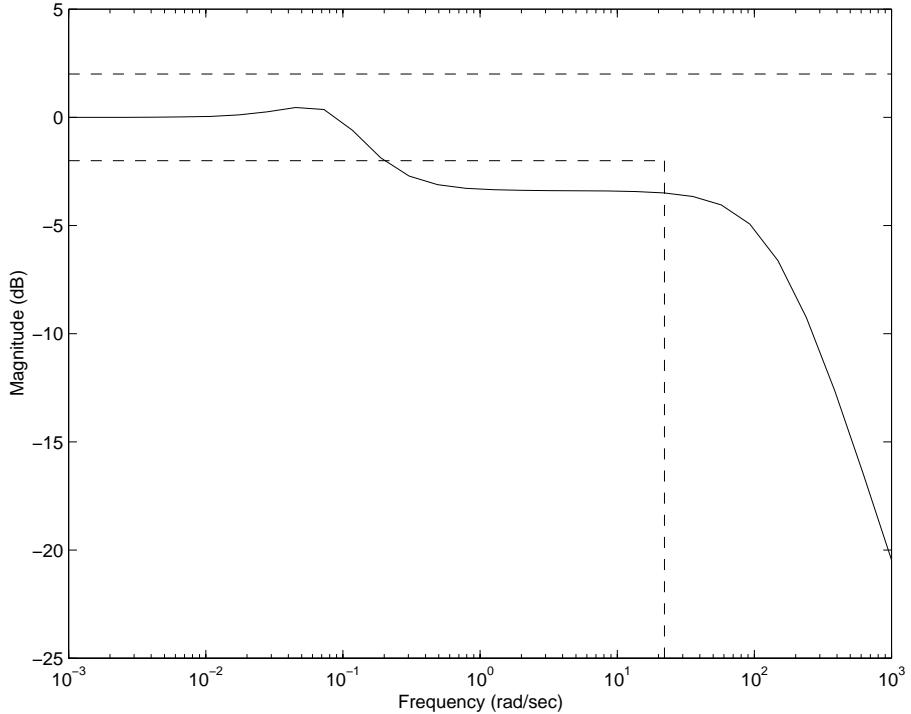


Figure 4.2: Frequency Response of Sensor with Time Correlated Measurement Noise

frequency performance, the transfer functions may become band-pass or some other arbitrary shape, and will therefore not satisfy the design requirements. In cases where a single sensor cannot meet the design specifications, it is therefore essential that multiple sensors, (or sensor suites) be used.

Consider the frequency response bound plot illustrated in Figure 4.1. For a sensor corrupted by exponentially time correlated noise, the frequency response of the system is shown in Figure 4.2. It can clearly be seen that this sensor *alone* will not satisfy the design requirements.

In [21] a method is presented for selecting sensor suites to satisfy the frequency response design requirement. In essence, sensors must be selected such that their frequency responses are complementary. For example, a sensor capable of tracking at high frequencies, such as an accelerometer (which typically has poor low frequency performance) should be coupled with a sensor that provides good low frequency information, such as a differential GPS.

Therefore, if the sensor with time correlated noise shown in Figure 4.2 is augmented with a relatively poor quality sensor with white noise only, it is reasonable to expect that the

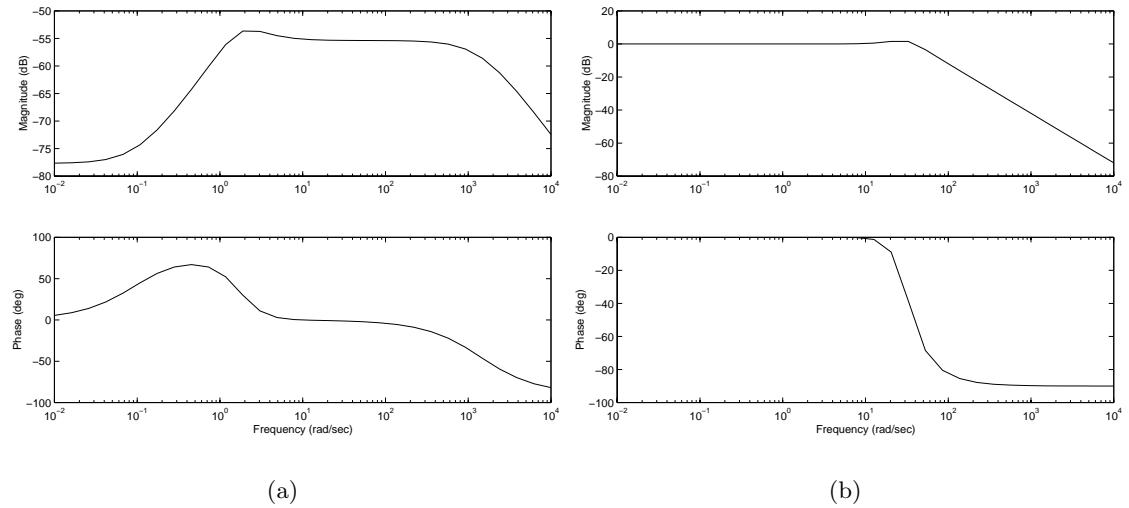


Figure 4.3: (a) Frequency Response of Sensor with Time Correlated Measurement Noise and (b) Frequency Response of Poor Quality Aiding Sensor

overall frequency response will improve. Continuing the example, Figure 4.3 shows the frequency response of the individual sensors to the state of interest. As the system is linear, these responses will sum at the output giving an overall response as shown in figure 4.4.

It can be seen that by augmenting the system with a relatively poor quality sensor, the systems' frequency response falls within the required bounds.

The analyses provided in Section 2.9 can be used to choose sensors whose frequency domain characteristics are complementary. For example, it was shown that type III sensors provide good high frequency information, but poor low frequency information. This sensor may be complemented with a type I sensor which is characterised by providing good low frequency information.

Figure 4.5 summarises the effects on the overall transfer functions and detectability plots of the different sensor combinations examined in Chapter 2. The graphs are only intended to give a general idea of the shape of the overall transfer functions and the primary detectability plots. The true shape of the graphs are dependent on the strength of the sensor noise covariance (and in the case of time correlated processes, the correlation time also).

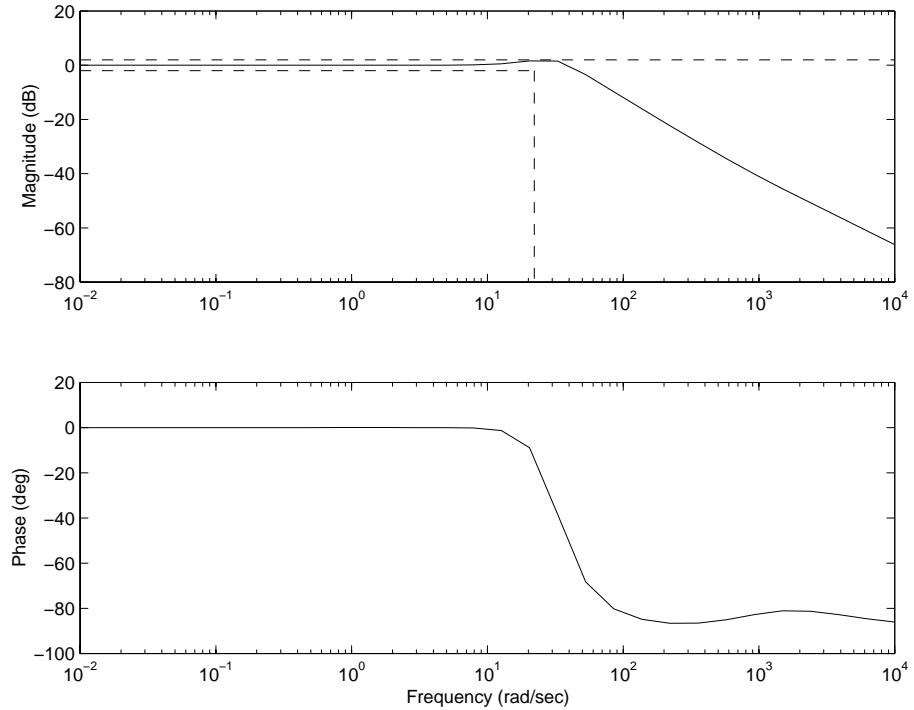


Figure 4.4: Overall Frequency Response of Sensor Suite

These frequency domain techniques may be used to find sensors whose characteristics are complementary. They do not, however, guarantee that the quality of the sensors is sufficient to meet the design requirements.

Also, as can be seen in Figure 4.5, simply designing a navigation system to meet the specifications does not in general guarantee that all faults will be detectable. This concept will be further expanded in Section 4.5.

4.4 Generalising Sensor Suite Design

The design of a sensor suite to achieve the design specifications as presented in [21] has several shortcomings. The main problem is that there is no mechanism for determining the accuracy of the system. Also, there is no way to specify the quality of sensors that should be used. Instead the method relies on trial and error to determine whether a given sensor suite fulfills the design requirements.

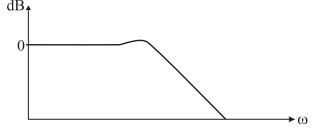
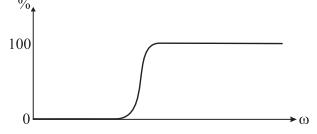
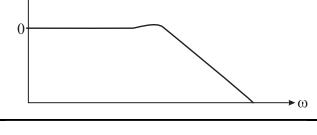
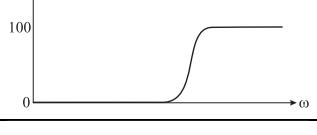
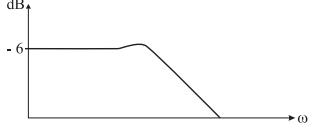
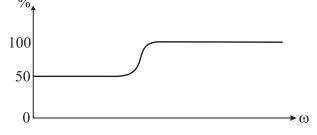
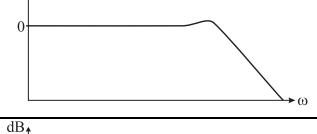
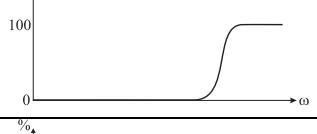
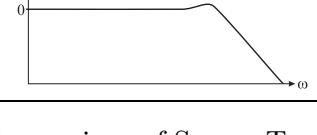
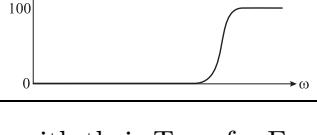
Sensor(s)	Transfer Function	Detectability Plot	Comments
Type I			Low frequency faults are undetectable
Type II			Coloured noise cannot be decorrelated without further sensing
Type III	—	—	Not observable
Dual Type I			All faults are detectable
Type I and Type II			Complementary sensors - Larger passband than a type I sensor alone
Type I and Type III			Complementary sensors - Larger passband than a type I sensor alone

Figure 4.5: A Comparison of Sensor Types with their Transfer Functions and Detectability Plots

This section examines a method by which the filter performance (the desired requirements) dictate the quality of the sensors used, rather than traditional design methods where the quality of the sensors used dictates the performance of the filter.

4.4.1 A Single Sensor Example

Example 4.4.1 Recalling the example presented in Section 2.2.2 it was shown that for the system described by

$$\begin{aligned} \begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \\ \mathbf{z}(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + v(t) \end{aligned} \quad (4.1)$$

with observation and process noise covariances given by

$$\mathbb{E}[w(t)w(t+\tau)] = Q\delta(\tau) \quad (4.2)$$

$$\mathbb{E}[v(t)v(t+\tau)] = R\delta(\tau) \quad (4.3)$$

the Kalman filter cutoff frequency was

$$\omega_n = Q^{\frac{1}{4}}R^{-\frac{1}{4}} \quad (4.4)$$

with associated steady state covariance

$$\mathbf{P} = \begin{bmatrix} \sqrt{2}Q^{\frac{1}{4}}R^{\frac{3}{4}} & \sqrt{QR} \\ \sqrt{QR} & \sqrt{2}Q^{\frac{3}{4}}R^{\frac{1}{4}} \end{bmatrix} \quad (4.5)$$

Assuming that the accuracy the system must meet is that the position standard deviation must not exceed A metres, and that the desired minimum frequency response is B rads $^{-1}$, then the problem equates to finding the solution(s) of the equations

$$\sqrt{\sqrt{2}Q^{\frac{1}{4}}R^{\frac{3}{4}}} < A \quad (4.6)$$

and

$$Q^{\frac{1}{4}}R^{-\frac{1}{4}} > B \quad (4.7)$$

So, for the arbitrary case where

$$\begin{aligned} Q &= 0.1m^2s^{-4}, \\ A &= 0.02m, \\ B &= 10rads^{-1} \end{aligned}$$

Equations 4.6 and 4.7 become

$$\begin{aligned} R &< 0.019m^2s^{-2} \\ R &< 0.00001m^2s^{-2} \end{aligned}$$

Clearly, to satisfy both inequalities, the position sensor must be chosen such that $R < 0.00001m^2s^{-2}$. Or, in other words, the position sensor's noise must have a standard deviation of less than $\sqrt{0.00001}ms^{-1}$.

This example illustrates that the design criteria of system accuracy and frequency response essentially dictate the *quality* of sensors that must be used to achieve the design specification. Typically the better the quality of sensor, the higher the price. This technique allows the minimum quality sensors to be chosen which achieve the desired specifications, thus minimising the potential cost.

4.4.2 The Riccati Equation

It was shown in Section 2.2.2 that for the continuous time Kalman filter, the steady state covariance matrix is determined by solving the algebraic Riccati equation

$$\dot{\mathbf{P}}(t) = \mathbf{F}\mathbf{P}(t)\mathbf{F}^T + \mathbf{G}\mathbf{Q}(t)\mathbf{G}^T - \mathbf{P}(t)\mathbf{H}^T\mathbf{R}^{-1}(t)\mathbf{H}\mathbf{P}(t) \quad (4.8)$$

letting $\dot{\mathbf{P}}(t)$ equal zero.

As was seen in Section 4.4.1, the solution of the Riccati equation is essential in determining the minimum values the noise covariance matrix must take in order to achieve the design

specification.

In general, the Riccati equation cannot be solved algebraically for systems of order greater than two[10]. There do exist, however, several numerical methods which can find solutions to Equation 4.8. One such technique based on eigenvector decomposition is summarised in Appendix D.

This basic property of the Riccati equation, means that *in general* to find the quality of sensors one must use to satisfy the design requirements, numerical rather than algebraic techniques must be used.

4.4.3 Numerical Techniques

To obtain the required noise characteristics of the sensors that satisfy the design criteria in situations where the state vector has order greater than two, a numerical solution to the Riccati equation must be found.

As the solutions are numerical, a search procedure must be used to calculate the values of the sensor noise covariance matrix. The following points summarise the general procedure.

- Solve Riccati equation and calculate steady state covariance and Kalman gains.
- Find the highest frequency pole - this will be the corner frequency.
- Form the steady state covariance matrix and evaluate the accuracy of the state(s) of interest.
- Store sensor noise covariance matrix if the design criteria are met.
- Repeat.

For the case where it is desired to find the sensor noise covariance matrix for a system with two sensors, Figure 4.6 shows the pseudocode used to determine this matrix. The variables R1 and R2 are the sensors' noise variance values. For each sensor minimum and maximum values are chosen that reflect the quality of available sensors. The highest frequency pole is found by solving the characteristic polynomial of the system which can be determined from Equation 2.13.

```

for(i = R1min; i<= R1max; i++){
    for(j = R2min; j<= R2max; j++){
        Covariance = SolveRiccati( System_Matrices );
        Kalman_Gain = Get_Gain( Covariance, System_Matrices );
        High_Pole = Get_HPole( Kalman_Gain, System_Matrices );
        if(( Covariance <= Desired_Accuracy )and
            ( High_Pole >= Desired_Cut_Off )){
            Store_Noise( i, j );
        }/* end if */
    }/* end for */
}/* end for */

```

Figure 4.6: Psuedocode to Find the set of Sensor Noise Covariance Matrices

This procedure will calculate a set of values for the noise covariance matrix for which the design criteria are met. It is then up to the designer to cull this set and choose values which will (typically) minimise the cost of the sensors used. The pseudocode shown outlines an iterative search procedure, searching between an assumed minimum and maximum for each variable. To decrease computational effort, a better search procedure such as a binary search should be used.

4.5 Designing for High Integrity

So far, this chapter has discussed methods for designing a navigation system such that certain design specifications are met. Although this provides a guarantee that the system will nominally perform to specification, there is no such guarantee that the system will be robust to failures, as was seen in Figure 4.5.

Chapter 3 discussed methods for design that guarantee that all faults will be detectable. Therefore, all that remains to the design of a high integrity navigation system is to apply these principles.

4.5.1 Redundancy

Following the arguments presented in Chapter 3, sensor redundancy *must* be employed in order to create a high integrity navigation system. The concept of frequency domain redun-

dancy specifies that the sensors employed should be based on different physical principles to reduce the possibility of sensors having common failure modes.

For a navigation system, the simplest way to provide frequency domain redundancy, and thus high integrity, is to repeat the whole system design, though using different sensing technologies, thereby obtaining two independent sensor suites. This design methodology allows the distribution of the navigation system over multiple computers, as will be seen in the following section.

4.5.2 High Integrity Filter Structures

For a high integrity navigation system, alternative filter structures exist for the way in which sensor information is integrated. These high integrity filter structures are designed to prevent the third class of failure identified in Section 3.2, those of software and hardware.

Figure 4.7 shows what is perhaps the most intuitive method for integrating multiple sensor suites. It follows that, because of the need for redundancy, at least two complementary sensor suites will be designed. And, as the sensors will be measuring the same (or comparable) information, the two suites may be directly integrated in a Kalman filter to provide a single consistent estimate of state. Additionally, the innovations of this filter will be sensitive to sensor failure *at all frequencies*, assuming that both sensor suites have been independently chosen to meet the design specifications. There is one important aspect the filter configuration shown in Figure 4.7 does not take into account, however. If the software or hardware (computer) on which the Kalman filter resides fails, then the entire navigation system will fail.

To overcome this problem, the filter must be distributed over multiple computers as shown in Figure 4.8. In this configuration three Kalman filters (implemented on three processors) are used, one for each sensor suite, and one to combine the estimates from both sensor suites. This has the advantage that if any single computer fails, there will always be an estimate of state available to the higher level control algorithms. However, if there is a sensor failure in one of the primary Kalman filters, sensor fault detection via the innovations can not in general be guaranteed (as there is no sharing of information between the sensor suites).

The advantage of this architecture, therefore, is that if a computer fails, the navigation system will be able to operate in a degraded mode until such time as a predetermined ‘safe’



Figure 4.7: Direct Kalman Filter with Two Sensor Suites

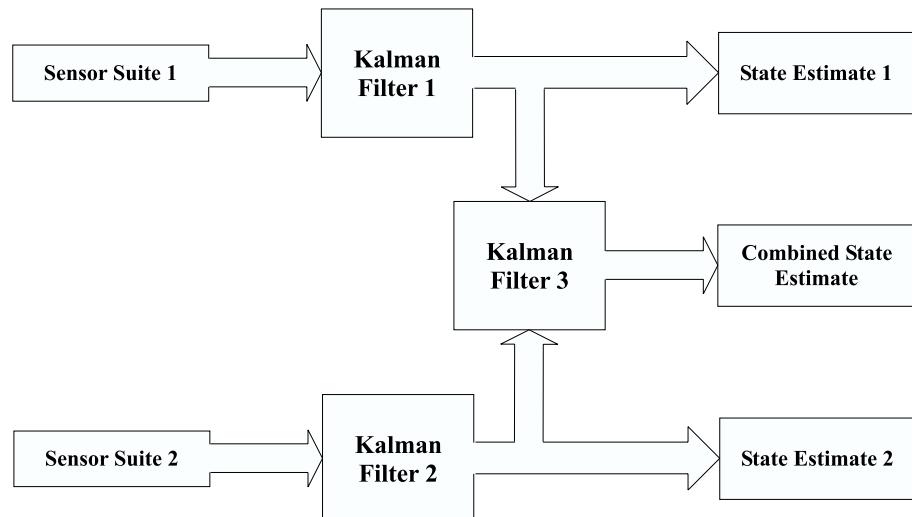


Figure 4.8: High Integrity Kalman Filter with Two Sensor Suites

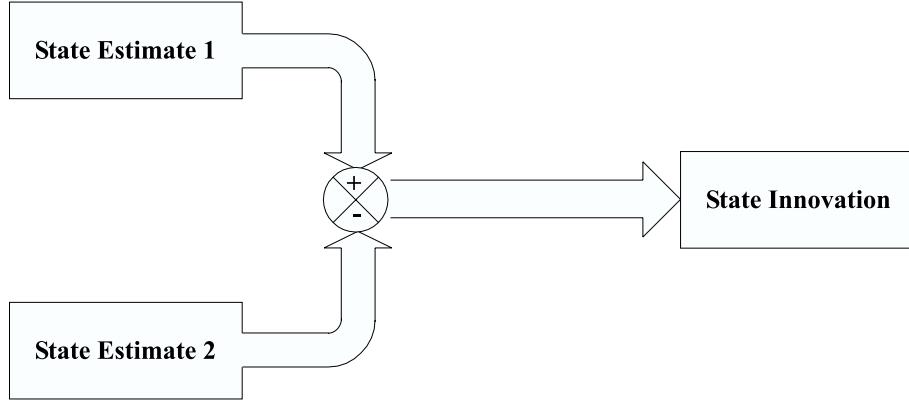


Figure 4.9: High Integrity Kalman Filter State Innovation Construction

configuration is reached (and repairs carried out).

To guarantee sensor fault detection for this configuration, the fault detection scheme for the filter shown in Figure 4.8 must necessarily be different from the standard innovation whiteness tests outlined in Section 3.8 because, as discussed previously, each individual filter will have frequency regions where faults are undetectable in the innovations. The information content provided by the sensors in the two filter structures discussed is identical, so it follows that there must be a method of determining sensor faults at all frequencies.

The key idea in this case is in understanding that faults not detectable in the innovation sequence of the individual filters will be passed to the state estimate of that filter. It can be shown[7] that under the assumption that the filter is consistent (no fault) that the estimation error is zero mean, uncorrelated in time with known covariance. Therefore if a state estimate innovation is constructed as shown in Figure 4.9, as a difference of the state estimated by both filters, then this error will also be zero mean with known variance. If a fault is undetectable in the innovations, then it will be detectable in the state estimate innovation measure.

Let $\hat{\mathbf{x}}_1(k|k)$ and $\hat{\mathbf{x}}_2(k|k)$ be the estimates from Kalman filter 1 and Kalman filter 2 respectively, with associated covariances $\mathbf{P}_1(k|k)$ and $\mathbf{P}_2(k|k)$.

Also, let the estimation errors of the Kalman filters be represented by

$$\begin{aligned}\tilde{\mathbf{x}}_1(k|k) &= \mathbf{x}(k) - \hat{\mathbf{x}}_1(k|k) \\ \tilde{\mathbf{x}}_2(k|k) &= \mathbf{x}(k) - \hat{\mathbf{x}}_2(k|k)\end{aligned}$$

The expected error from the true state $\mathbf{x}(k)$ for each filter can be derived in discrete time [6] as

$$\begin{aligned} \mathbb{E} [\tilde{\mathbf{x}}_1(k|k)] &= 0 \\ \mathbb{E} [\tilde{\mathbf{x}}_2(k|k)] &= 0 \end{aligned} \quad (4.9)$$

with covariance matrices defined as

$$\begin{aligned} \mathbb{E} [\tilde{\mathbf{x}}_1(k|k)\tilde{\mathbf{x}}_1(k|k)^T] &= \mathbf{P}_1(k|k) \\ \mathbb{E} [\tilde{\mathbf{x}}_2(k|k)\tilde{\mathbf{x}}_2(k|k)^T] &= \mathbf{P}_2(k|k) \end{aligned} \quad (4.10)$$

The expected value of the state innovation can then be constructed as

$$\mathbb{E} [\tilde{\mathbf{x}}_1(k|k) - \tilde{\mathbf{x}}_2(k|k)] = 0 \quad (4.11)$$

showing that the state innovation will have zero mean on the assumption that there are no faults present in the system.

The covariance matrix for the state innovation may be found as

$$\begin{aligned} \mathbf{P}_3(k|k) &= \mathbb{E} [(\tilde{\mathbf{x}}_2(k|k) - \tilde{\mathbf{x}}_1(k|k))(\tilde{\mathbf{x}}_2(k|k) - \tilde{\mathbf{x}}_1(k|k))^T] \\ &= \mathbf{P}_1(k|k) + \mathbf{P}_2(k|k) - \mathbf{P}_{12}(k|k) - \mathbf{P}_{21}(k|k) \end{aligned} \quad (4.12)$$

where the terms $\mathbf{P}_1(k|k)$ and $\mathbf{P}_2(k|k)$ are computed internally by the Kalman filters 1 and 2 and

$$\begin{aligned} \mathbf{P}_{12}(k|k) &= \mathbf{P}_{21}(k|k)^T \\ &= [\mathbf{I} - \mathbf{K}_1(k)\mathbf{H}_1(k)] [\mathbf{F}(k-1)\mathbf{P}_{12}(k-1|k-1)\mathbf{F}^T(k-1) + \mathbf{Q}(k-1)] \\ &\quad \times [\mathbf{I} - \mathbf{K}_2(k)\mathbf{H}_2(k)]^T \end{aligned} \quad (4.13)$$

as given in [6] (under the assumption of process correlations). The matrices $\mathbf{K}_i(k)$, $\mathbf{H}_i(k)$, $\mathbf{F}(k-1)$ and $\mathbf{Q}(k-1)$ retain their definitions from Chapter 2.

Now, if the processes models in Kalman filter 1 and 2 are uncorrelated, which is likely if frequency domain redundancy is employed, the cross correlation matrices introduced in

Equation 4.12 will equate to zero, giving the covariance for the state innovation as

$$\mathbf{P}_3(k|k) = \mathbf{P}_1(k|k) + \mathbf{P}_2(k|k) \quad (4.14)$$

which may be easily computed.

For the case where the processes in Kalman filter 1 and 2 *are* correlated, the track to track data fusion equations given in [6] must be used for Kalman filter 3, however the state innovations will still have known mean and covariance and can be tested using the whiteness tests outlined previously.

Therefore, to test the state innovation for faults, one must simply implement the same whiteness tests outlined in Section 3.8. In this structure, however, the whiteness tests must be applied to the filter innovations of Kalman filter 1 and 2, as well as to the state estimate innovations to ensure the detection of all possible faults. This is because only high frequency faults are detectable in filters 1 and 2, with the remaining low frequency faults being passed to the state innovation measure.

4.5.3 Other Architectures

The previous section argued that in terms of hardware integrity, the distribution of the filter over multiple processors is an advantage. This is not an uncommon goal and has been dealt with widely in the literature[41, 84, 89].

Perhaps the most useful algorithm for use in distributed systems is the information filter. Grime[39] provides an excellent overview, in which the information filter is shown to be directly analogous to the Kalman filter, with the additional benefit that it is relatively easy to implement in a distributed form.

The information filter, rather than estimating the state of the system $\mathbf{x}(k)$, estimates the information state $\mathbf{y}(k)$ defined as

$$\hat{\mathbf{y}}(k|k-1) \triangleq \mathbf{P}^{-1}(k|k-1)\hat{\mathbf{x}}(k|k-1) \quad (4.15)$$

The matrix $\mathbf{P}^{-1}(k|k-1)$ is known as the *information matrix*. The full derivation of the filter may be found in [39, 62].

In [62], it is shown that the filter innovations can be written in information space as

$$\begin{aligned}\nu'(k) &= \mathbf{H}(k)^T \mathbf{R}^{-1}(k) \nu(k) \\ &= \mathbf{H}(k)^T \mathbf{R}^{-1}(k) \mathbf{z}(k) - \mathbf{H}(k)^T \mathbf{R}^{-1}(k) \mathbf{H}(k) \hat{\mathbf{x}}(k|k-1)\end{aligned}\quad (4.16)$$

As there is a direct analogy of the innovations in information space, and the information filter satisfies the same criteria of minimum mean squared error estimation that the Kalman filter does, this means that the analyses provided for fault detection via the innovations are directly applicable to the information filter.

4.6 Conclusion

This chapter explored the issue of the design of high integrity navigation systems. Firstly, a method was presented which allowed sensors to be chosen to satisfy the dual goals of minimising overall cost and of meeting the design specifications. For a high integrity navigation system, it was concluded that the system must have redundant sensing, and if possible, be implemented in a distributed form.

In particular, Section 4.2 presented the design specification, interpreted in terms of the Kalman filter. It was shown that the design specification may be interpreted as the highest frequency the kalman filter must track, together with a maximum steady state covariance. Section 4.3 examined the use of sensor suites to achieve the design specification. In Section 4.4 a general method was presented for sensor suite design, subject to the constraints inferred by the design requirements. This section also presented a method to enable the automation of sensor suite design.

In Section 4.5 the concept of high integrity navigation systems was reintroduced. It was argued that sensor redundancy *must* be employed to achieve high reliability. Also, to improve reliability, alternate filter structures were proposed which distribute the filter across multiple computers. These distributed systems were proposed as a solution to computer hardware and software reliability problems.

Chapter 5

Experimental Results

5.1 Introduction

This chapter details the design, implementation and results of two very different experiments in high integrity navigation system design.

The first experiment, utilising a gyro and a laser sensor, is presented in Section 5.2. This experiment was designed to enable physical insight into the processes (namely fault propagation) studied in this thesis. Faults are injected into the system, and the results are compared with the results predicted by the frequency domain fault detectability metrics developed in Chapter 3.

The second experiment, outlined in Section 5.3, describes the initial results from an ongoing project into the design and implementation of a high integrity navigation system for an underground mining vehicle. The overall design of the navigation system is given together with extensive results obtained from one of the sensor suites. The preliminary navigation system described explores the issues of vehicle and sensor modelling.

5.2 The Gyro-Laser Experiment

In this section, an experiment designed to verify the fault detectability metrics developed in Chapter 3 is presented. The experimental platform consists of two sensors which may be considered complementary in the frequency domain. These sensors supply information

about the position and orientation of the platform, which are the states of most interest in navigation system design. By introducing faults into the system, and comparing these results to those predicted by the fault detectability metrics, an experimental validation of the theory may be ascertained.

The experimental results, given in Sections 5.2.3, 5.2.4, 5.2.5 and 5.2.6, show the effects of introducing bias to each of the sensors. Biases and slow drifts are arguably the most difficult types of fault to detect. Through the use of the frequency domain techniques developed in this thesis, the results explain why biases are difficult to detect, and why frequency domain redundancy is the key to detecting faults such as this.

The experimental results also show an example of a process fault.

5.2.1 Experimental Setup

The experiment uses two reasonably high quality sensors, which are:

- A GCS¹ laser
- An Andrews fibre optic gyroscope

These sensors are shown in Figures 5.1 and 5.2.

The GCS laser sensor operates by detecting the angle to a number of fixed beacons. Three beacons (actually retroreflective tape) were used in this experiment and were surveyed with a theodolite so their positions in absolute co-ordinates were accurately known. The strips were placed roughly in an equilateral triangle with the experimental rig situated roughly in the centre.

The Andrews gyro measures angular rate $\dot{\phi}$, but as with all inertial measurement devices is subject to a low frequency drift type fault (making it a type III sensor). This sensor, however, has extremely good high frequency performance, being able to detect rotation rates of up to 100 degrees per second.

The sensor suite utilised the gyro and laser for the following reasons:

¹ Guidance Control Systems Ltd. (UK)

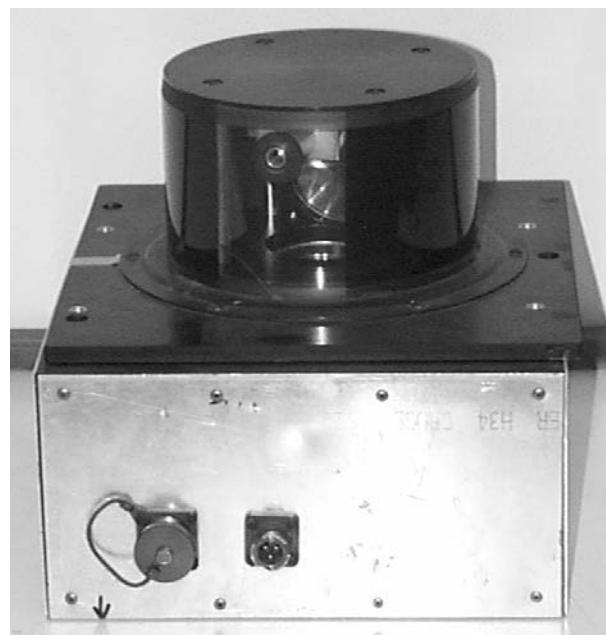


Figure 5.1: The GCS Laser

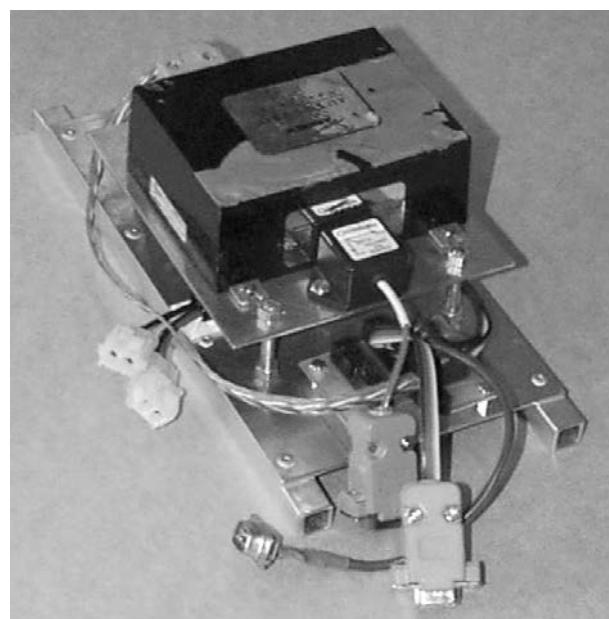


Figure 5.2: The Andrews Gyro

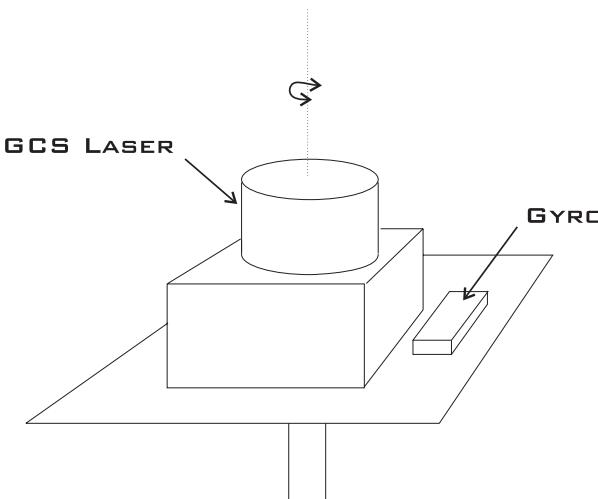


Figure 5.3: The Experimental Setup

- Over long time periods (the platform is stationary for 10-15 minute intervals over a 45 minute period), the gyro drift is significant. This sensor would therefore fail to estimate orientation correctly if used alone.
- The laser is a very good sensor at steady state, but sudden changes in orientation are almost impossible for it to track, due to its low bandwidth (a rotation rate of 1 Hz) and the sparseness of beacons. Again, this sensor would fail if used alone.

These facts may be interpreted in the frequency domains as follows: The laser is a type I sensor, and as such has good low frequency performance. The gyro, however, is a type III sensor, which has good high frequency performance. When combined, these sensors complement one another, as described in Chapter 2. Therefore, this combination should allow the whole spectrum of platform rotation manoeuvres to be able to be tracked.

The sensors were mounted on a plate such that the axis of rotation of the GCS laser was approximately located over the centre of the plate. The gyro was mounted at the periphery of the plate such that its sensitive axis was perpendicular to the surface of the plate. The plate itself was mounted on a stand such that it was free to rotate about its centre. A schematic of the experimental setup may be seen in Figure 5.3.

The data from the two sensors was logged and time stamped by a transputer network at a rate of 50 Hz, and post-processed to provide the results seen in this section.

The experiment itself was extremely simple. The platform was left for a period of time, usually 10 to 15 minutes, and then rotated quickly through approximately 90 degrees. This process was repeated several times over the course of a single run.

The first set of results (Section 5.2.3) for this experiment are for the nominal system, with no faults added. Section 5.2.4 gives the results of a run for which an additional bias was added to the gyro, and Section 5.2.5 shows results for the system in the presence of a laser bias. In Section 5.2.6, the effects of frequency domain redundancy are shown by simulating a second laser sensor.

5.2.2 Filter Design and Analysis

The orientation of the platform was modelled by a simple constant velocity model, allowing ϕ and $\dot{\phi}$ to be estimated. The gyro is known to drift, so a shaping filter must also be added. Also, to allow for small variations in position, x and y were also estimated.

The state vector for this system is therefore defined as

$$\mathbf{x}(t) = [x(t), y(t), \phi(t), \dot{\phi}(t), x_{sf}]^T \quad (5.1)$$

The continuous time linear process model may now be written as

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{\phi}(t) \\ \ddot{\phi}(t) \\ \dot{x}_{sf}(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \\ \phi(t) \\ \dot{\phi}(t) \\ x_{sf}(t) \end{bmatrix} + \begin{bmatrix} w_x(t) \\ w_y(t) \\ 0 \\ w_{\dot{\phi}}(t) \\ w_{x_{sf}}(t) \end{bmatrix} \quad (5.2)$$

Note that the shaping state x_{sf} , and the two position states x and y are all modelled as Brownian motion processes. The shaping state, however, is designed to reflect the coloured noise component of the gyro measurement, while the position state models are intended to reflect the uncertainty in the true state, and the rate (i.e. randomly) at which the true state is considered to vary.

The non-zero elements of the vector $[w_x, w_y, 0, w_{\dot{\phi}}, w_{x_{sf}}]^T$ are all assumed zero mean, uncorrelated Gaussian sequences with strengths σ_x^2 , σ_y^2 , $\sigma_{\dot{\phi}}^2$ and $\sigma_{x_{sf}}^2$ respectively.

σ_x^2	1^{-10} m^2
σ_y^2	1^{-10} m^2
σ_ϕ^2	$4.9^{-7} \text{ rad}^2 \text{s}^{-2}$
σ_{sf}^2	$1^{-10} \text{ rad}^2 \text{s}^{-2}$
σ_{gyro}^2	$8.1^{-7} \text{ rad}^2 \text{s}^{-2}$
σ_θ^2	1^{-8} rad^2

Table 5.1: The Noise Parameters

Linear observations are considered to be of the form

$$\mathbf{z}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k) \quad (5.3)$$

The gyro measurement (z_{gyro}) is a linear combination of the angular rate $\dot{\phi}$ and the shaping state x_{sf} as

$$z_{gyro}(t) = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \end{bmatrix} \mathbf{x}(t) + [v_{gyro}(t)] \quad (5.4)$$

where the gyro white noise component v_{gyro} is assumed a zero-mean, uncorrelated Gaussian sequence with strength σ_{gyro}^2 .

The laser provides a nonlinear observation which may be considered to be of the form

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}(t)) + \mathbf{v}(t) \quad (5.5)$$

The bearing to a beacon $\mathbf{B}_i = [X_i, Y_i]^T, i = 1, \dots, N$ is given by $\arctan(\frac{Y_i - y(t)}{X_i - x(t)})$. However the platform is oriented in the direction ϕ , so for this system, the measurement equation for each beacon detected by the laser is given by the model

$$\mathbf{z}_\theta^i(t) = \left[\arctan\left(\frac{Y_i - y(t)}{X_i - x(t)}\right) - \phi(t) \right] + [v_\theta^i(t)] \quad (5.6)$$

where the laser observation noise v_θ^i , is uncorrelated zero mean and Gaussian with strength σ_θ^2 , and is assumed to be identical for each beacon observed.

To filter the sampled data, the discrete Kalman filter was used. The continuous to discrete time conversion for this system is detailed in Appendix E.

Throughout this experiment, the noise strengths listed in the Table 5.1 were used.

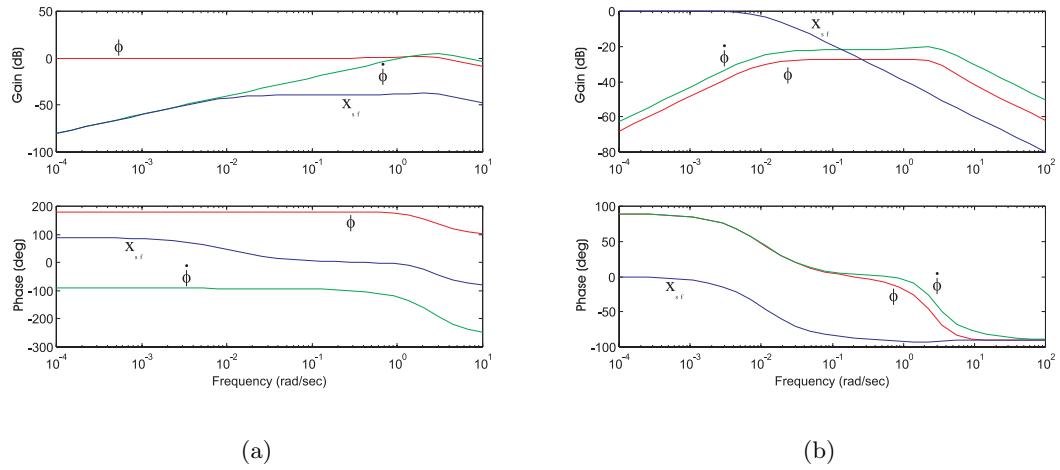


Figure 5.4: Bode Diagram of the Transfer Functions from (a) the Laser and (b) the Gyro to the states ϕ , $\dot{\phi}$ and x_{sf}

Figure 5.4 shows Bode diagrams of the theoretical transfer functions from the laser and gyro respectively. The transfer functions are only shown for the states of most interest, ϕ , $\dot{\phi}$ and x_{sf} . Note that the transfer functions are not directly comparable, as the laser provides an observation of ϕ , while the gyro observes $\dot{\phi}$. Figure 5.5 is the combined transfer function from the two sensors to the state ϕ . It can be seen that the cutoff frequency of the combined system is higher than that of either individual sensor.

The Fault Plots for this system can be seen in Figures 5.6 and 5.7. These figures show that only very low frequency faults in either the gyro or the laser will go undetected in the innovations. However, if the Bode plot for the gyro is examined (Figure 5.4(b)) at low frequencies, gyro information is transmitted to the shaping state. Therefore a fault in the gyro is either detectable, or transmitted to the shaping state.

As the gyro drift is explicitly modelled and incorporated in the Kalman filter, a low frequency fault in the gyro will not affect the estimated states of interest. It is instead estimated by the shaping state.

A low frequency laser fault on the other hand *is* transmitted to the states of interest and is undetectable in the innovations.

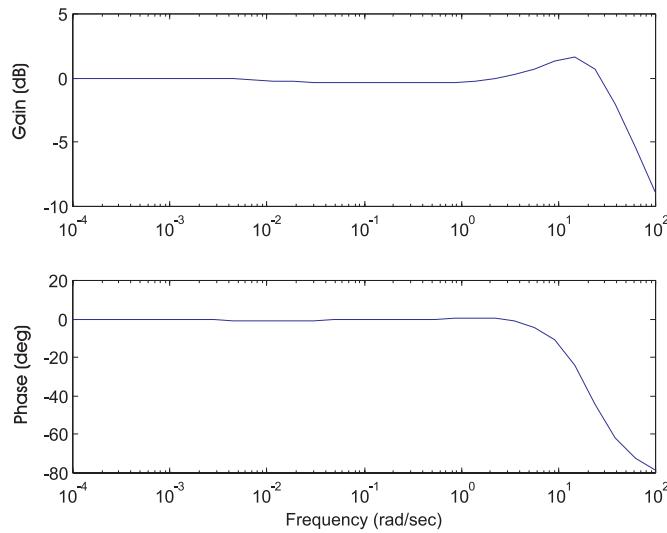


Figure 5.5: Bode Diagram of the Combined Transfer Functions to the State ϕ

This may be seen from the laser Bode and fault detectability plots and is due to the fact that in the Kalman filter design, the laser is assumed to provide good low frequency information.

Following the arguments presented in Chapter 3, all process faults are guaranteed to be detectable in this system except those that *only* affect the shaping state.

This is due to the fact that only the shaping state is unobserved.

5.2.3 Results: The Nominal System

Figure 5.8 shows the orientation of the platform as estimated by the filter. This result agrees well with the experiment performed: the platform was stationary for relatively long periods of time, then rotated quickly by approximately 90 degrees. This result can be directly compared to that shown in Figure 5.9. This figure shows the effect of simply integrating the gyro measurements to obtain the platform's orientation. It can be seen that the integrated gyro drifts by approximately 90 degrees over the course of the experiment. It may also be seen that the small spikes at the end of some of the rotations actually occur (they are due to human error) and are estimated correctly.

Figure 5.10 shows the estimated angular rate of the platform. The angular rate can be seen

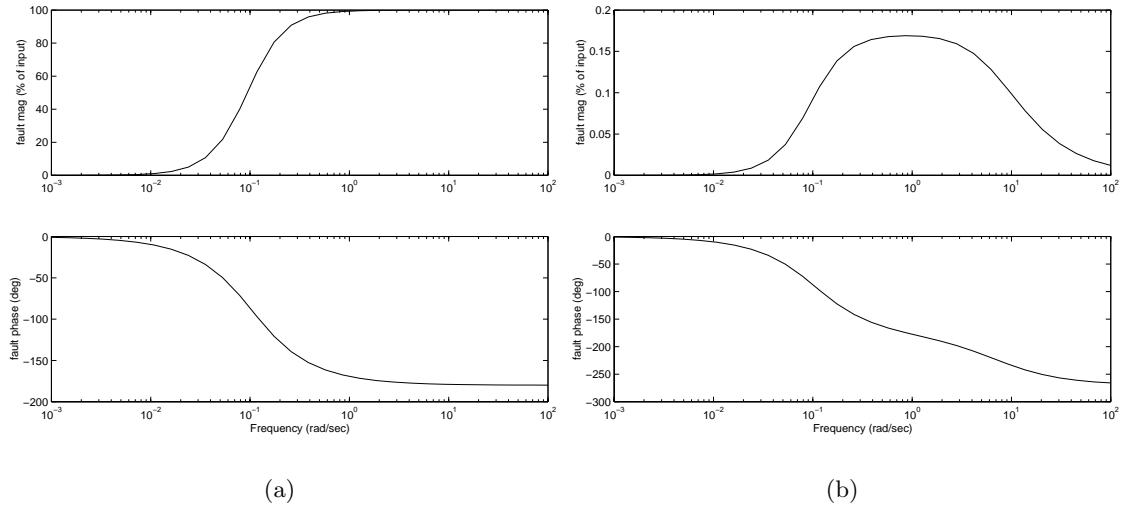


Figure 5.6: Fault Plots - Percentage of Fault in Laser Sensor Transmitted to (a) the Laser Innovations (b) the Gyro Innovations

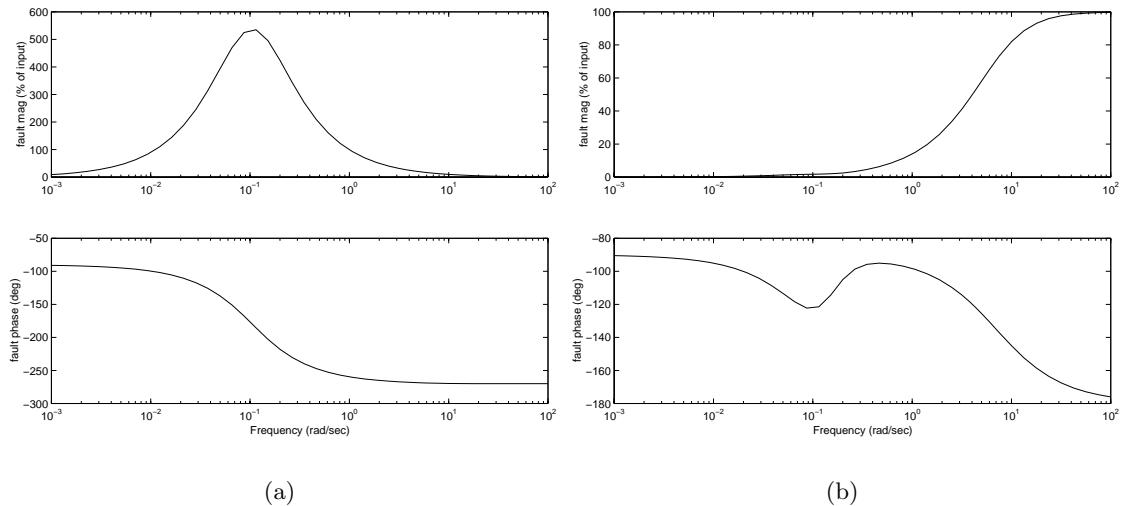


Figure 5.7: Fault Plots - Percentage of Fault in Gyro Transmitted to (a) the Laser Innovations (b) the Gyro Innovations

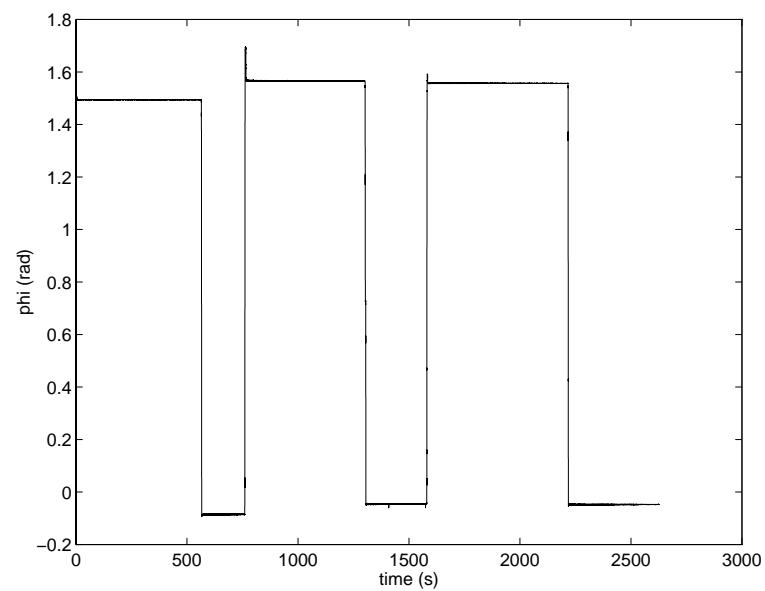
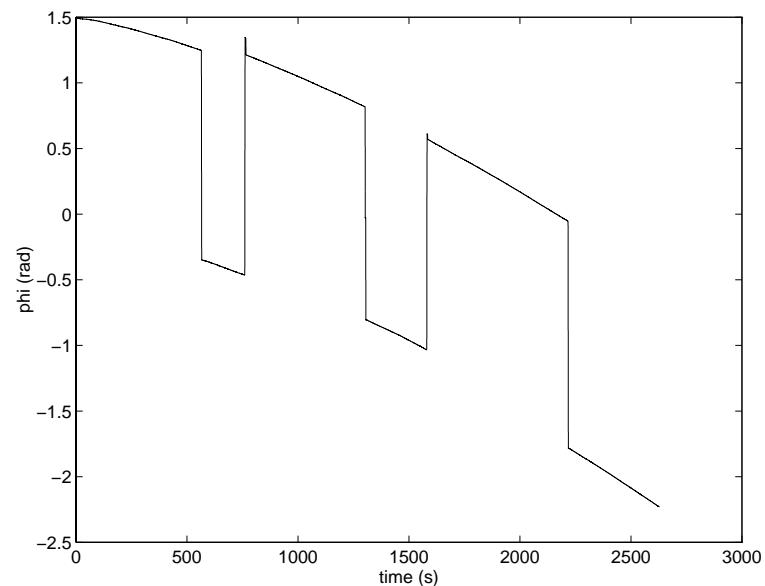
Figure 5.8: Estimated Orientation ϕ 

Figure 5.9: Integrated Gyro Output

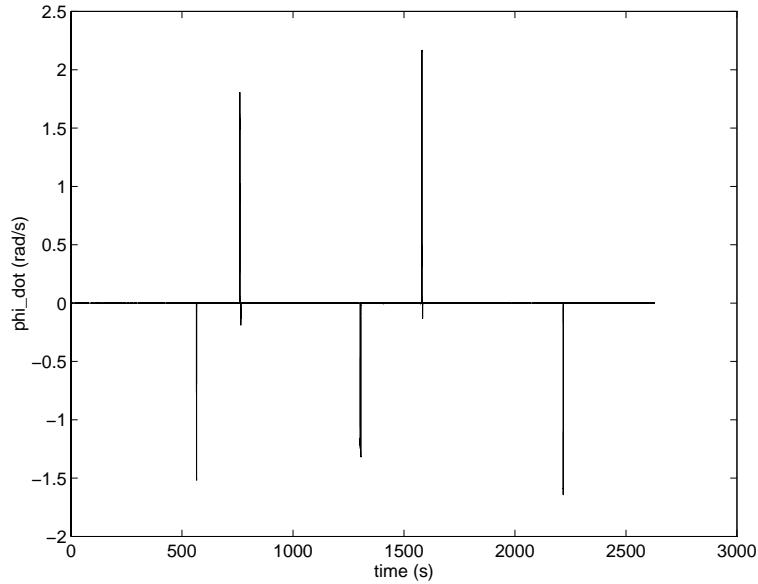


Figure 5.10: Estimated Angular Rate $\dot{\phi}$

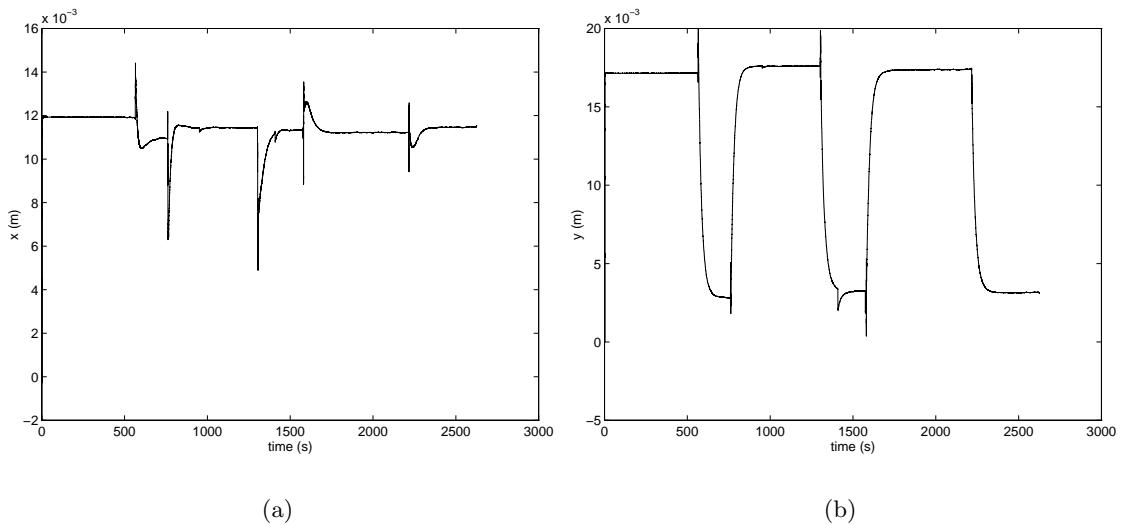
to be approximately zero when the platform is stationary, and large when the platform is rotated, a result which also concurs with the experiment performed.

In principle, the platform rotates about the geometric centre of the laser, and therefore the laser does not move in the xy plane. However, Figure 5.11 shows that indeed the laser moved in the order of a few millimetres during rotation.

This highlights the need for accurate models - if the xy position were not estimated, this small error would feed into the other states, perhaps causing filter divergence.

The spikes in these estimates, which correspond to the times at which the platform was rotated, are an artifact of bearing only tracking. The laser is the only sensor which supplies information to the position states, allowing them to be estimated. The fact that the laser supplies bearing only, causes the filter to have a greater weight for the orientation state than for the position states. This means that the position states take longer to converge to steady state than the orientation. In a system that is much more tightly coupled, for example by using a sensor that supplies range and bearing, one could reasonably expect the estimates of position to converge at a much faster rate.

Figure 5.12 shows the estimated shaping state x_{sf} . It can be seen that the shaping state ‘absorbs’ the gyro drift. Again, there are spikes corresponding to the changes in orientation

Figure 5.11: Estimated Position (a) x and (b) y

of the platform. In this case, *this is due to a process fault* (as can be seen and easily detected in the gyro innovation sequence shown in Figure 5.13). The process model used is an extremely poor model of the dynamic system. There is no way for the process model to predict the onset of movement in the platform. The shaping state tries to compensate for some of this error (thus the spikes), but at steady state, only the gyro drift is estimated.

Following the theory presented in Chapter 3, as both the orientation and the angular rate are being observed, any process faults occurring in either of those states is guaranteed to be detectable in the innovations.

This can be seen in Figure 5.13 which shows the gyro innovation. The innovation is zero mean and white when the platform is stationary, indicating correct filter performance. However at the points where the platform is being rotated, the innovation sequence jumps, indicating a fault has occurred. The laser innovation sequence exhibits very similar behaviour.

Figures 5.14 and 5.15 show the innovations of the gyro and laser respectively during periods when the platform is stationary. Both innovation sequences appear to be unbiased and zero mean, indicating that when the platform is stationary, the filter is performing correctly.

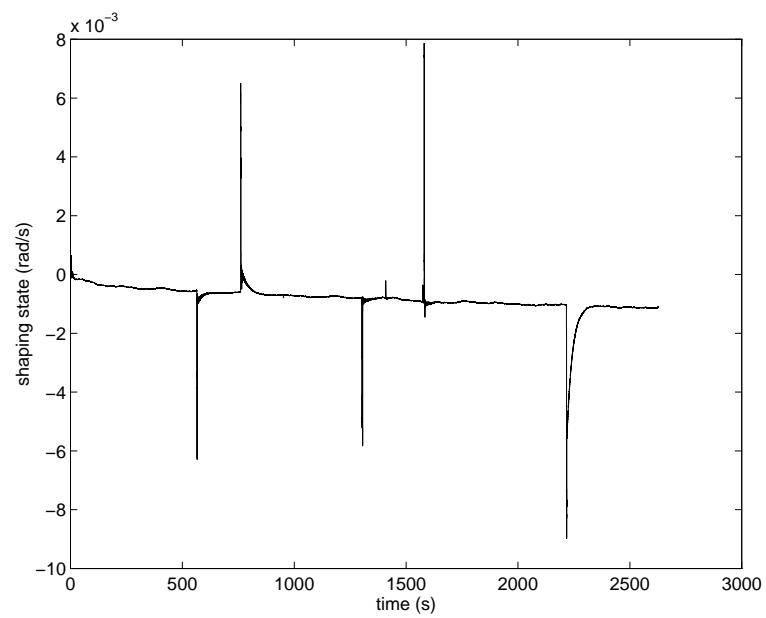
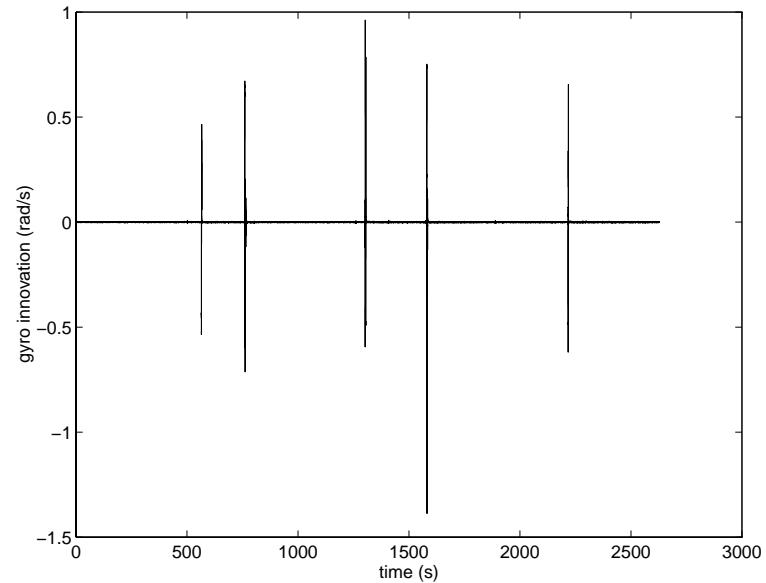
Figure 5.12: Estimated Shaping State x_{sf} 

Figure 5.13: Gyro Innovation Sequence

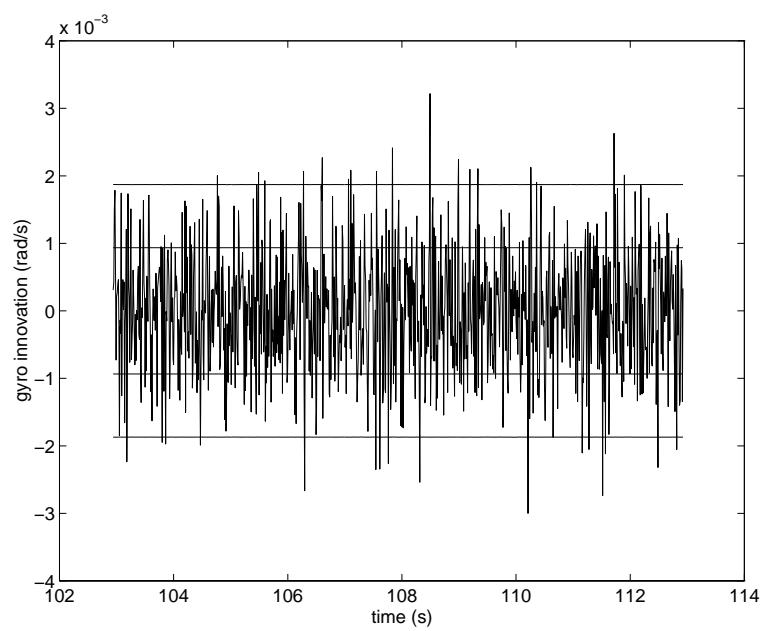


Figure 5.14: Gyro Innovation Sequence with 1σ and 2σ Bounds

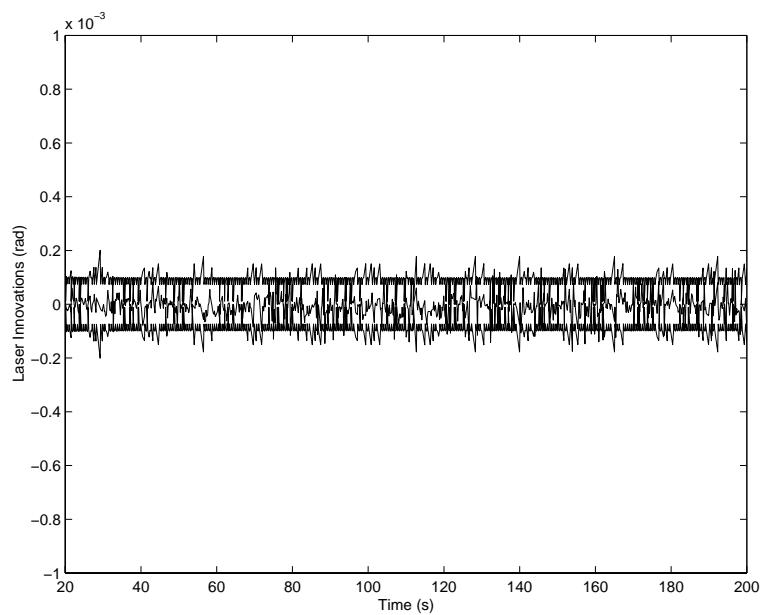


Figure 5.15: Laser Innovation Sequence with 2σ Bound

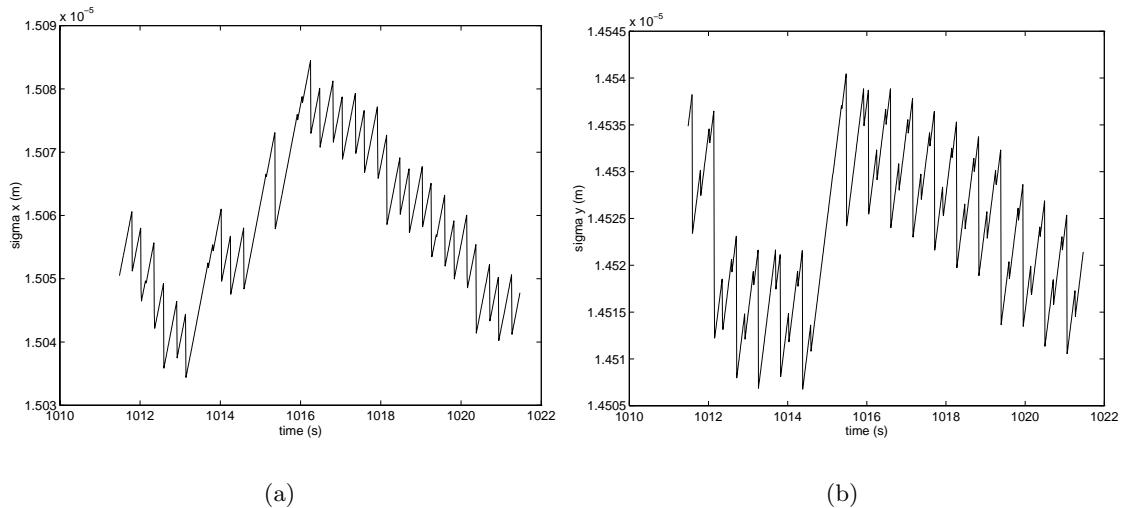


Figure 5.16: Position Standard Deviation (a) x and (b) y

The standard deviations of each of the parameters gives an indication of the accuracy of the system. Figures 5.16, 5.17 and 5.18 show the standard deviations of the states of the system. It can be seen that the average standard deviation in orientation is approximately 0.0001 radians (0.0057 degrees), and the position standard deviation is in the order of 0.015 millimetres. These standard deviations are well within the tolerances demanded by most applications.

In summary, these results show that a constant velocity process model is not sufficient to accurately model the *dynamic* behaviour of the experimental rig. In steady state, the constant velocity process model is a perfectly valid model of system behaviour. However, when the platform was rotated significant error was induced, as evidenced by the innovations. The constant velocity model does not accurately track the system when the velocity is not constant, as is the case when the platform is rotated.

By sensing the orientation states, however, this fault is detectable in the innovations, as predicted by the theory presented in Chapter 3.

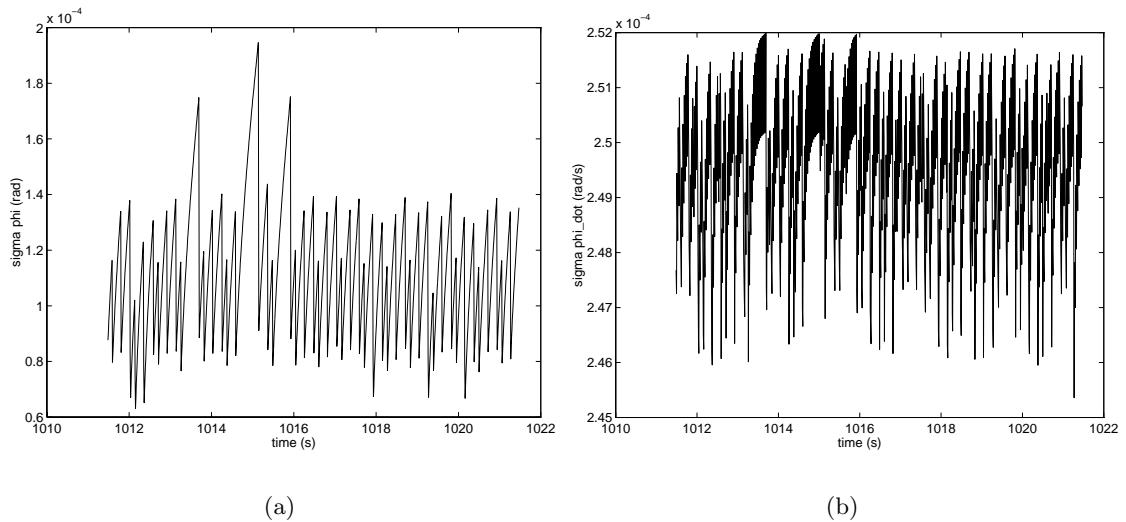


Figure 5.17: (a) Orientation Standard Deviation and (b) Angular Rate Standard Deviation

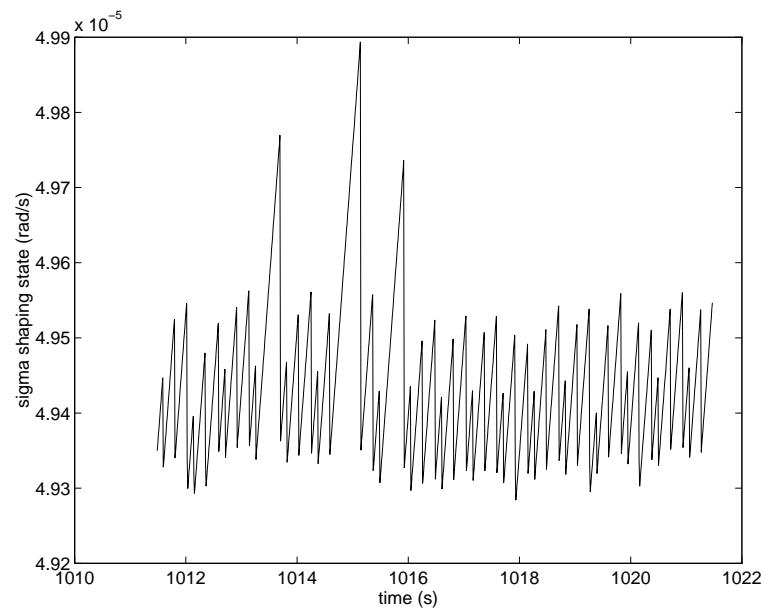


Figure 5.18: Shaping State Standard Deviation

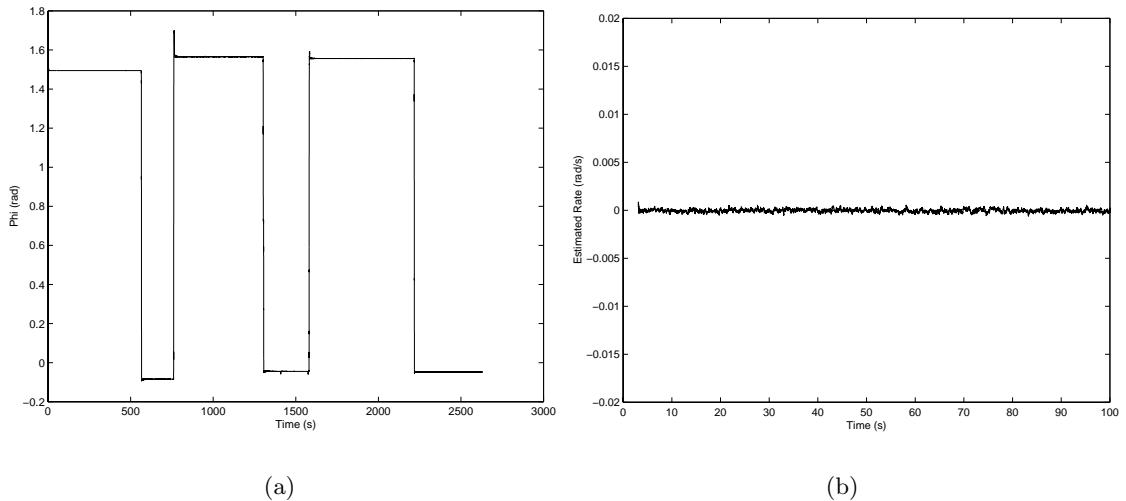


Figure 5.19: (a) Estimated Orientation ϕ and (b) Estimated Angular Rate $\dot{\phi}$

5.2.4 Results: in the Presence of Gyro Bias

One of the results of the previous section was the estimation of the gyro drift. In this experiment, an artificial bias was added (in software) to the gyro of 0.001rad/s . Integrated, this corresponds to an angular drift of 149 degrees over the course of the experiment.

Figure 5.19 shows the estimated orientation and angular rate of the system. These results agree with those presented in the previous section. The position states x and y (not shown) also closely match the previous results.

Therefore, the gyro bias (a very low frequency fault) *does not* affect the state estimates, as predicted.

In Figure 5.20, the raw gyro output is shown. The bias can be clearly seen. Figure 5.21 shows the estimated shaping state. This figure is nearly identical to the one presented previously, however the initial offset has moved from zero to approximately 0.001 reflecting the magnitude of the added bias.

As the gyro drift is explicitly modelled, the filter very quickly identifies and removes the gyro bias.

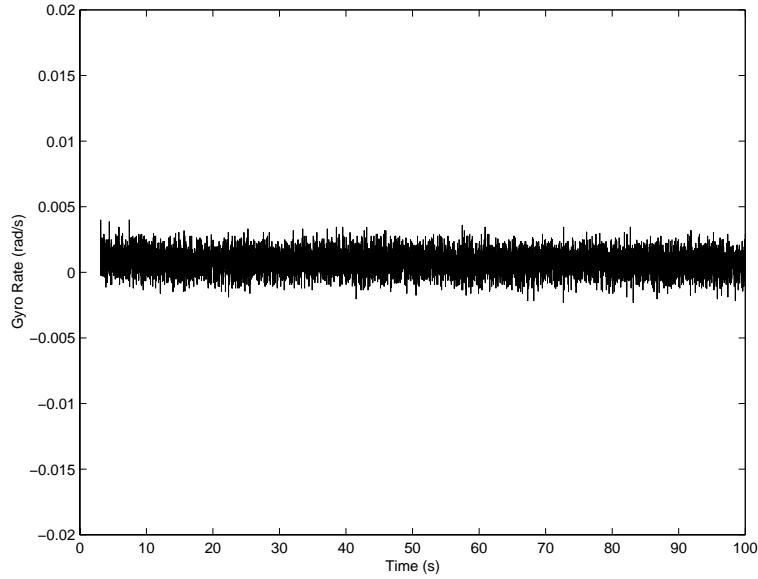


Figure 5.20: Gyro Observation of Angular Rate $\dot{\phi}$

In Section 2.4 it was shown that the sensor error model for Brownian motion is identical to the model for unknown bias. Therefore, the gyro shaping state absorbs both drift *and* bias.

The gyro innovations (see Figure 5.22) remain zero mean and white. Note that the innovations 2σ bound lies at around $0.003\text{rad}/\text{s}$. A constant bias of $0.001\text{rad}/\text{s}$ is therefore not insignificant.

This experiment serves to show that very low frequency faults in the gyro, such as bias and drift, are modellable and estimable.

The bias in the gyro has no effect on the estimate of orientation. This is because the filter uses only the high frequency content of the gyro signal when the platform is in rotation. The low frequency component of the gyro signal is fed into the shaping state. This effect was predicted by the Bode diagrams given in Figure 5.4.

5.2.5 Results: in the Presence of Laser Bias

For this experiment a bias was added to the GCS laser of 0.1 radians (approximately 5.7 degrees). This bias simulates a misalignment of the sensor with respect to the platform.

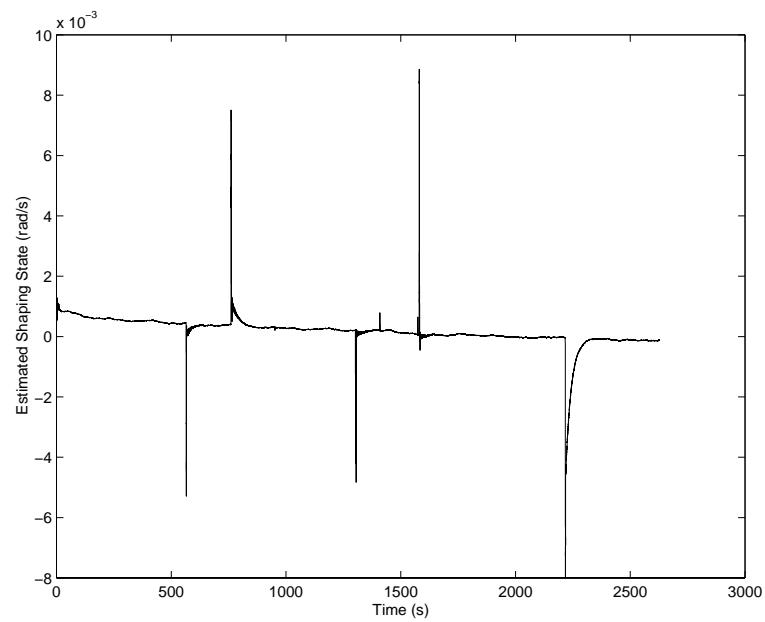
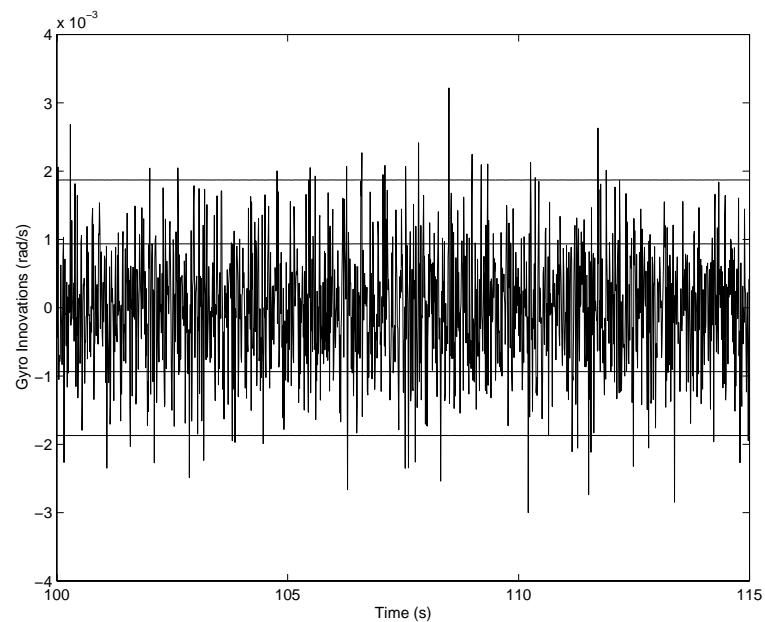
Figure 5.21: Estimated Shaping State x_{sf} 

Figure 5.22: Gyro Innovations

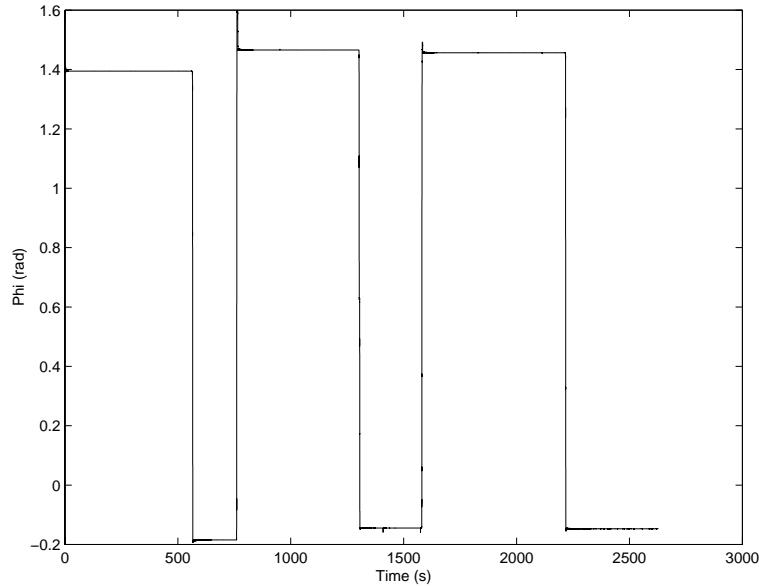


Figure 5.23: Estimated Orientation ϕ

This is not an uncommon problem in the field, as it is extremely difficult to align a sensor such as the GCS laser to a tolerance that is smaller than the accuracy of the sensor.

Figure 5.23 shows the estimated orientation of the platform.

As predicted by the Bode diagrams, the bias is passed straight through to the estimate, resulting in the estimated orientation being consistently 0.1 radians offset from the unbiased case.

Figure 5.24 shows the innovation sequence for the gyro and laser respectively. Again, they appear unbiased and zero mean.

As predicted by the laser fault plots, the laser bias is undetectable in the innovations.

This is an example of the worst kind of sensor fault. There is no way to internally ascertain that the system is at fault, as the filter continues to produce consistent estimates with innovations that are white with zero mean. The estimates, however, *do not reflect the true state*. The only way to detect the bias in a system such as this is to provide redundancy (following the arguments from Chapter 3).

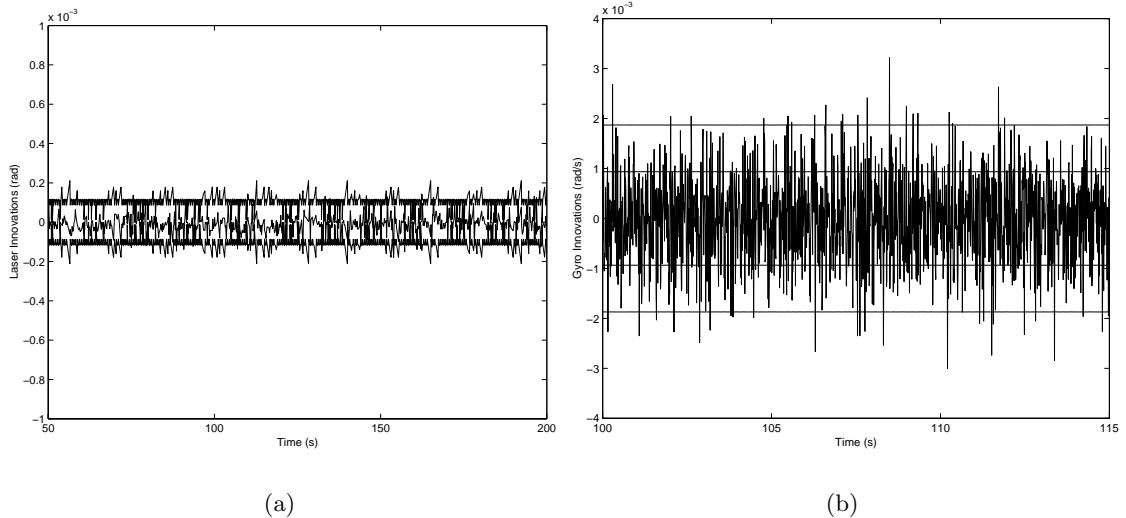


Figure 5.24: (a) Laser Innovations and (b) Gyro Innovations

This example shows why biases and slow drifts in low frequency sensors are so hard to detect. In cases such as this, these faults are proveably *undetectable*.

It should be noted that in this case the sensor did not fail in hardware. The simulated fault (misalignment) is one that is caused by human error. With redundancy, even faults such as these are detectable.

5.2.6 Results: Laser Bias with Redundancy

In this experiment, a second (identical) laser sensor was simulated in software. The second laser sensor had a bias of 0.01 radians. Practically, the second laser was simulated by simply adding the bias to the measurements made by the first laser sensor, and using this value as the measurement of the simulated sensor.

By simulating the laser sensor, this experiment shows the effect of frequency domain redundancy. In Chapter 3 it was argued that if frequency domain redundancy is employed, then all sensor faults will be rendered detectable.

Figure 5.25 shows the percentage of fault in the simulated sensor transmitted to the simulated sensor's innovation sequence. This shows that a fault in this sensor is detectable at

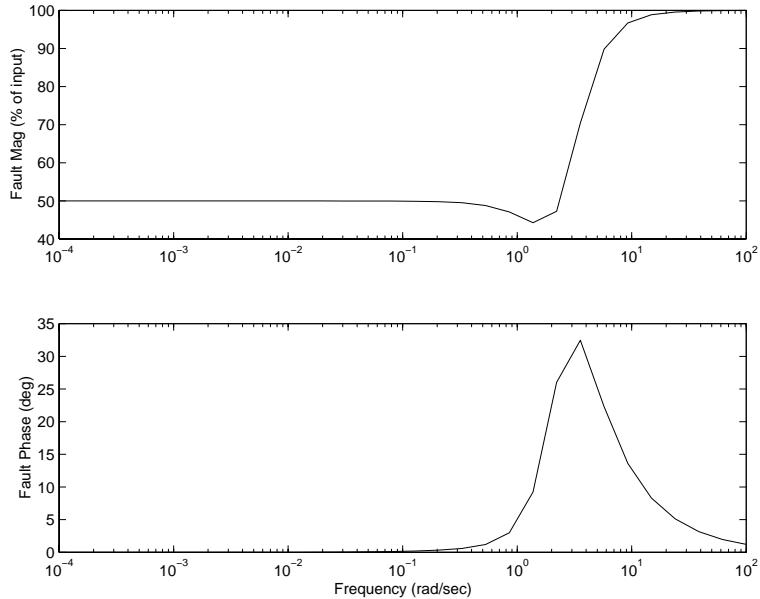


Figure 5.25: Fault Plot - Percentage of Simulated Laser Fault Transmitted to Simulated Laser Innovations

all frequencies, as expected. It can be seen that at low frequencies only 50% of the fault will be transmitted to the innovations. This is due to the fact that both laser sensors have the same observation noise covariance.

Figure 5.26 gives the percentage of a fault in the simulated sensor transmitted to the real laser sensors innovations. Again, at low frequencies 50% of the fault is transmitted.

In Figure 5.27, the fault plot for a fault in the simulated sensor to the gyro innovations is shown. At low frequencies, none of the fault is transmitted.

The estimated orientation of the platform for the run is shown in Figure 5.28(a). The filter only remains nondivergent for the first 35 seconds before failing catastrophically. It can be seen that the filter has trouble resolving the orientation of the platform when one sensor is at fault, as indicated by the noise on the estimated orientation.

As predicted by the fault plots, the bias in the simulated laser is not transmitted to the gyro innovations as can be seen in Figure 5.28(b). The bias *is* transmitted to the laser innovations, as can be seen in Figure 5.29.

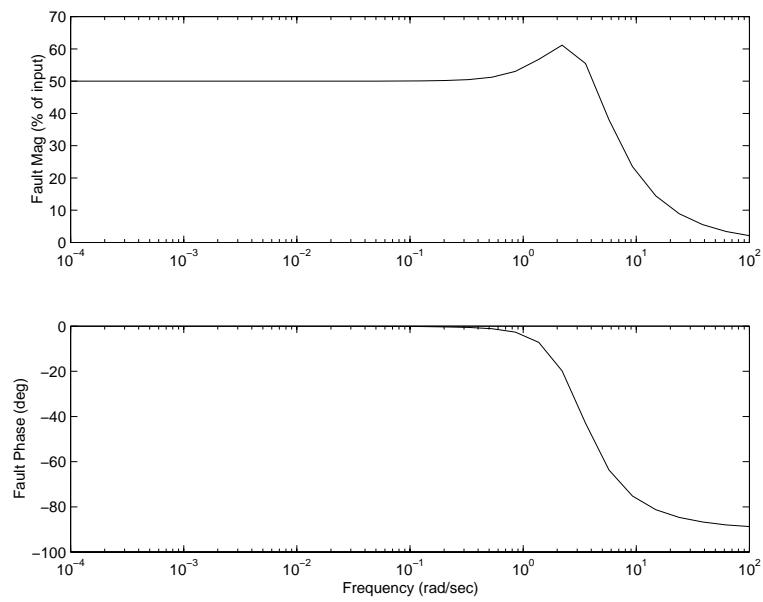


Figure 5.26: Fault Plot - Percentage of Simulated Laser Fault Transmitted to Real Laser Innovations

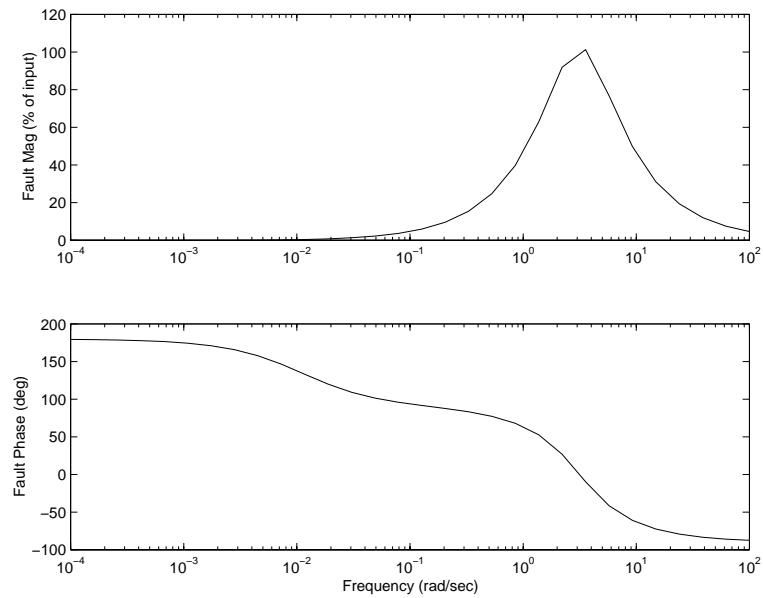


Figure 5.27: Fault Plot - Percentage of Simulated Laser Fault Transmitted to Gyro Innovations

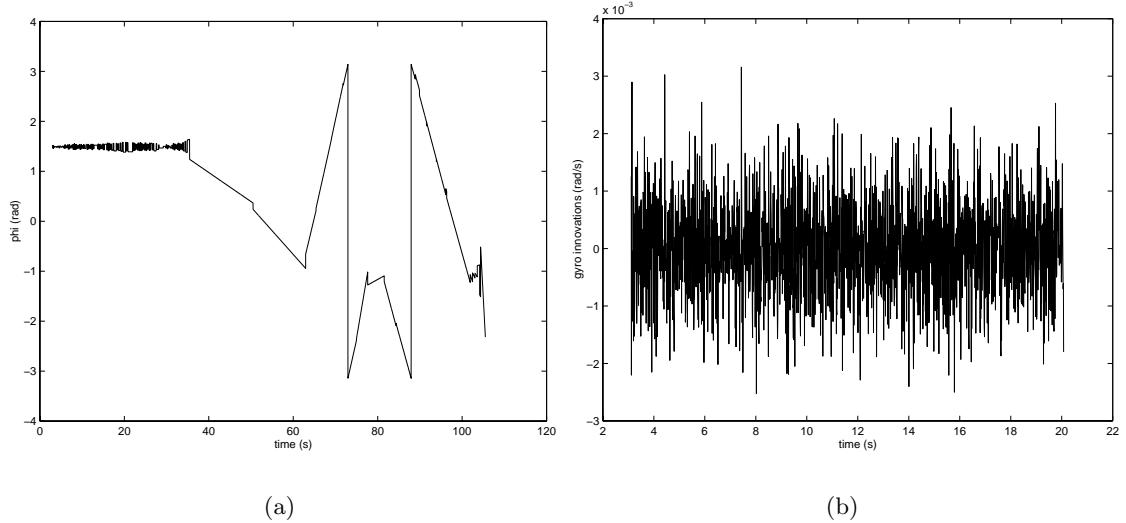


Figure 5.28: (a) Estimated Orientation ϕ and (b) Gyro Innovations

The mean of the innovation sequence shown in Figure 5.29(a) is -0.0048 radians, and the mean of the sequence shown in Figure 5.29(b) is 0.0052 , verifying the 50% transmittance of the 0.01 radian bias predicted by the fault plots.

5.2.7 Summary

These experiments have verified the fault detectability theory developed in this thesis.

In particular, these experiments have shown that for a particularly hard to detect class of faults, that Bode diagrams accurately predict the transfer of information to the state vector, and that fault detectability plots accurately predict the transfer of information to the filter innovations.

Further, this experiment has verified the theory developed in Chapter 3 to guarantee that process and sensor faults are detectable.

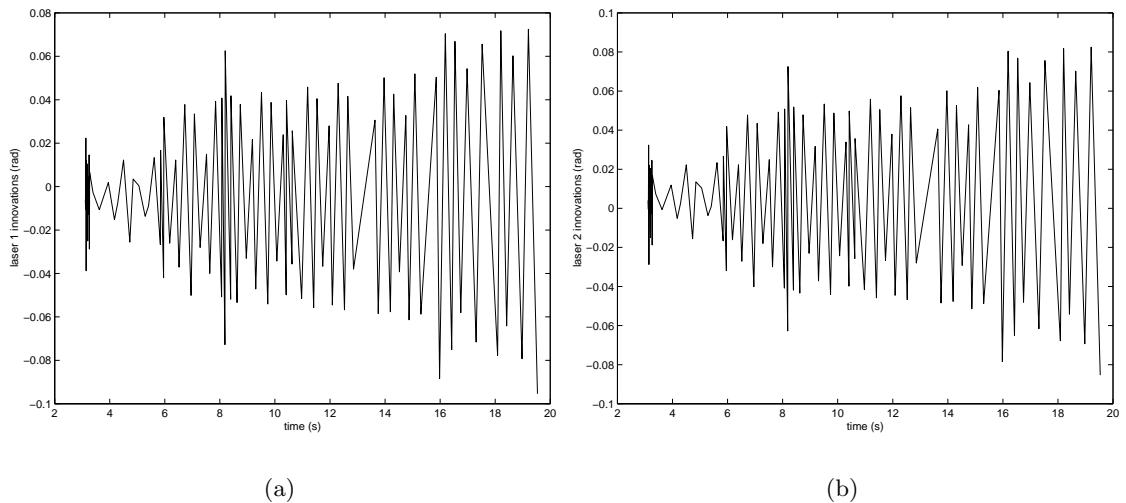


Figure 5.29: (a) Simulated Laser Innovations and (b) Real Laser Innovations

5.3 A High Integrity Navigation System for an LHD

During the course of this thesis, work was carried out at the Mt. Isa² lead-zinc mine in collaboration with the CRC³ for Mining Technology and Equipment (CMTE) and Mt. Isa Mines (MIM) to determine the feasibility of automating underground mining vehicles, including haul trucks and Load Haul and Dump (LHD) trucks (see Figure 5.30).

There is a strong case for automation of these vehicles for reasons of both safety and productivity. The safety issue has been addressed with several working implementations of teleoperated systems, including the well publicised Canadian systems [54, 55]. Teleoperation, however, offers little to increase productivity, as these systems tend to run at speeds slower than conventional manned systems, resulting in significantly lower productivity levels with the additional overhead of the infrastructure required to teleoperate. Many mines are now looking to increase productivity whilst maintaining safety and therefore wish to fully automate, using as little in-mine infrastructure as possible.

Several autonomous LHD systems have been trialed, including [5] which uses a retroreflective stripe on the tunnel roof (back) detected by cameras to guide the vehicle. This approach, however, requires relatively major levels of infrastructure, which must be changed whenever

²Located in Queensland, Australia

³Co-operative Research Centre

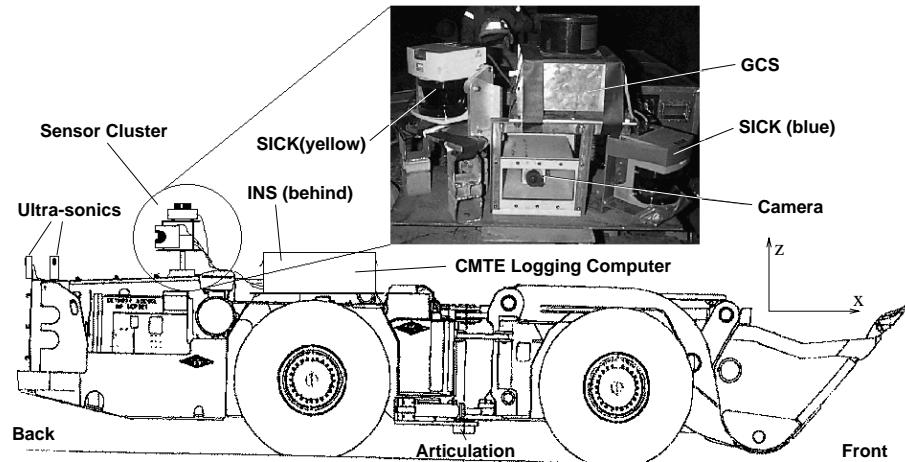


Figure 5.30: The LHD Showing Sensor Layout

a new path is required for the vehicle. In [90], a system is proposed that uses ultrasonic sensors to follow the walls of a tunnel. The wall-following approach suffers from lack of an Earth-fixed coordinate system, which is necessary for any planning operations, particularly traffic management. It is also difficult with wall following to accurately differentiate a side tunnel from a corner.

A much more robust approach is presented in [61], where information from various sensors is fused via an EKF, to provide an estimate of position and orientation. This approach, however, suffers from under-utilisation of data. The vehicle model presented uses heading provided from a gyro, coupled with encoder-derived velocity information. Information about articulation angle of the vehicle and vehicle dynamics are simply ignored.

For many mines, however, the problem of underground vehicle automation is still an open problem. The problem of navigation without infrastructure is far from solved. The problem of high integrity navigation has never previously been addressed. This section details the initial results of an ongoing project into the design of a high integrity navigation system for underground mining vehicles.

Section 5.3.1 describes the experimental setup in detail. In Section 5.3.2, the proposed high integrity navigation structure is outlined. This structure essentially uses two distinct sensor suites to estimate the vehicle parameters. The first, feature-based, sensor suite is described in Section 5.3.3.

The second sensor suite is geometrically based. The detailed derivation and implementation of the geometric-based navigation system begins in Section 5.3.4. The navigation system derivation explores the issue of vehicle modelling, using the ideas presented in Section 2.7.1. The key theoretical issues addressed in the design are the explicit modelling of slip parameters as the vehicle manoeuvres over rough terrain, and the use of inertial measurements to estimate these parameters.

5.3.1 Experimental Setup

The set of sensors tested underground on an LHD mining vehicle were:

- A strap-down Inertial Measurement Unit (IMU).
- Encoders/Odometers.
- A bearing only laser scanner (GCS).
- Two time-of-flight range and bearing laser scanners (SICK).
- Two sets of ultrasonic sensors (MASA/Polaroid).
- Vision - structured lighting.

These sensors and their layout on the vehicle may be seen in Figure 5.30.

The data from the different sensors needed to be recorded so that it could be post-processed. To achieve this, a distributed network of computers was located on an LHD, synchronised by a time stamp sent from the master computer to the slaves at a rate of $50Hz$. This enabled the data from each sensor to be individually time stamped. This is important as some of the sensors produce data asynchronously. The overall system is shown in Figure 5.31⁴.

Figure 5.32 shows the area used during the trials. The tunnel is approximately 150 metres long. The sensor data was collected while the LHD was running back and forward between the black circular marks. The small square in the figure shows the area of the tunnel for which results will be presented.

⁴Figures 5.30 and 5.31 were created by Jonathan Roberts of CMTE

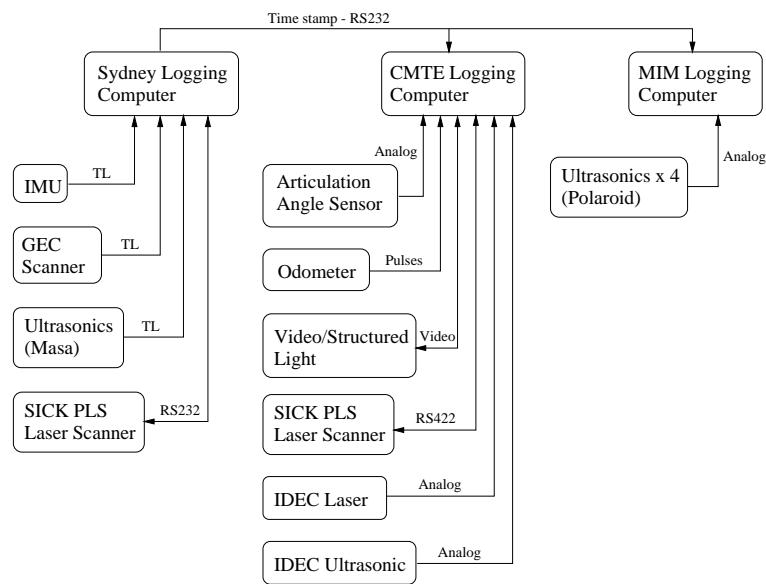


Figure 5.31: The Data Logging System

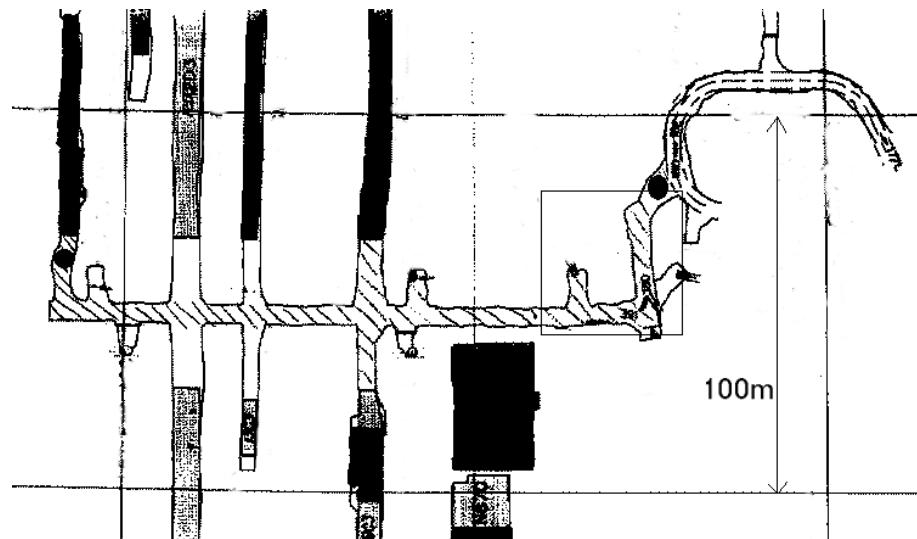


Figure 5.32: The Trials Area

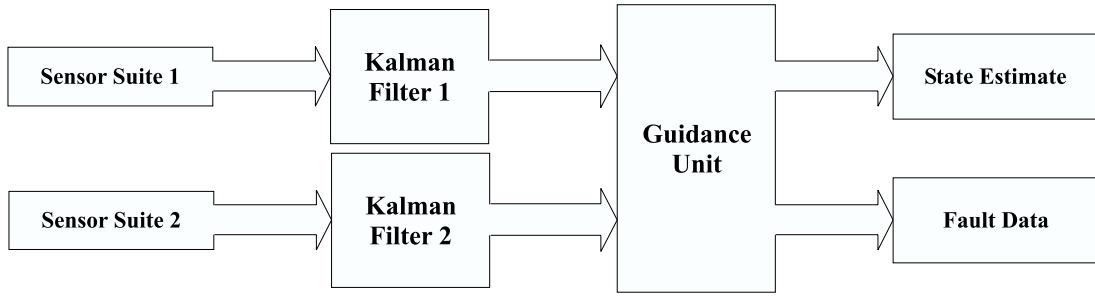


Figure 5.33: The Proposed Navigation System Architecture

5.3.2 Navigation Architecture

The overall structure of the proposed high integrity navigation system is shown in Figure 5.33.

The navigation system essentially consists of two physically independent navigation loops, indicated by the two separate sensor suites. Each loop consists of both high frequency and low frequency sensors in feedback configuration.

Provided that both navigation systems meet the design specifications, all faults in the system will be detectable, according to the arguments given in Chapter 4.

The guidance unit is responsible for combining the state estimates produced by each navigation loop, and for detecting faults as they occur. If enough information exists for the guidance unit to identify which loop is at fault, then the system will operate in a degraded mode, using only the estimates from the error-free navigation system, until such time as a safe configuration can be reached, or repairs carried out.

For a more detailed discussion of the proposed high integrity navigation architecture, refer to [27].

From the results of the underground sensor trials[86], two main navigation solutions presented themselves which are constrained by the wish for low infrastructure, and the desire to use physically different sensing technology. These systems will be used for the independent navigation loops shown in Figure 5.33.

The first navigation solution is feature based. An IMU or odometry is used for dead reckoning. This information is fused with information of natural landmarks, gained from

sensors such as the SICK laser or ultrasonics. Section 5.3.4 provides a general overview of the work being conducted towards this goal.

The second navigation system is geometric based, using encoder based dead reckoning fused with periodic updates to artificial landmarks using a sensor such as the GCS laser. This solution requires infrastructure, though far less than existing systems. This system was developed during the course of this thesis and is described in detail in Section 5.3.4.

5.3.3 The Feature-Based Navigation System

Research has been carried out at the University of Sydney, in conjunction with CMTE into an infrastructure free navigation system for underground mining vehicles.

The approach uses data from a SICK laser combined with dead reckoning information to estimate vehicle pose. In particular, the SICK range and bearing data is matched with an existing map using the Iterative Closest Point (ICP)[9, 97] algorithm. The approach used differs from the conventional ICP algorithm, in that an EKF is used to fuse feature data with the dead reckoning information. Figure 5.34⁵ shows a laser scan being matched to the known map of the environment, and the path estimated by this feature-based navigation system.

A more detailed discussion of this navigation system may be found in [60], which shows that the method is robust to occlusions and outliers.

5.3.4 The Geometric-Based Navigation System

The next sections detail the design and implementation of the geometric-based navigation system to be used as one loop of the high integrity navigation architecture. This navigation system was developed during the course of this thesis and primarily examines the issues of vehicle and sensor modelling outlined in Chapter 2.

5.3.5 The Vehicle Model

The LHD is an articulated vehicle, which means that the vehicle has a front and rear body which can rotate relative to each other. The front and rear wheel sets are fixed to remain

⁵Figure 5.34 was created by Raj Madhavan of the University of Sydney

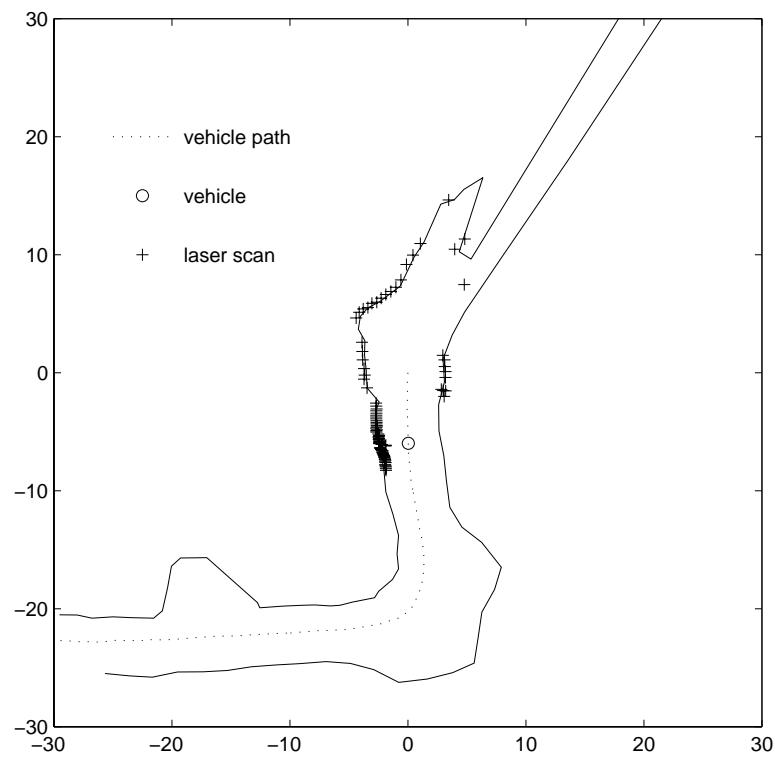


Figure 5.34: Estimated LHD Path - Using ICP Approach

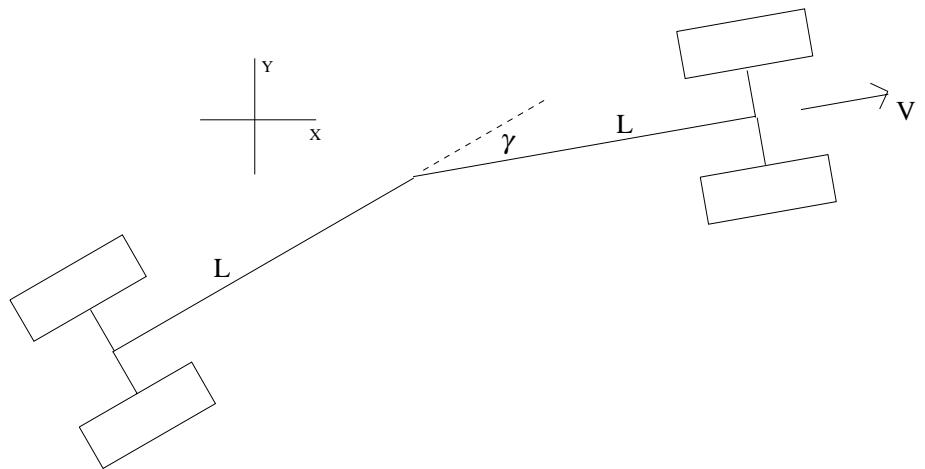


Figure 5.35: LHD Kinematic Geometry

parallel with the body of the vehicle, and steering is achieved by driving the articulation joint located midway between the front and rear axles. Figure 5.35 shows the geometry of a typical LHD.

5.3.6 The No-Slip Model

The kinematic model of the vehicle derived from rigid body and rolling motion constraints is given by;

$$\begin{aligned}\dot{x}(t) &= V \cos(\phi) \\ \dot{y}(t) &= V \sin(\phi) \\ \dot{\phi}(t) &= \frac{V \tan(\frac{\gamma}{2})}{L}\end{aligned}\tag{5.7}$$

where x and y denote the position of the vehicle relative to some fixed global frame of reference. The angle ϕ is the orientation of the vehicle with respect to the x axis, while V represents the linear velocity of an imaginary front wheel located midway between the real front wheels. The angle γ is defined as the articulation angle of the vehicle, and L is the half-length of the vehicle, the distance between the front or rear set of wheels and the articulation joint. The derivation is omitted here, but the result concurs with [90].

However, for the particular LHD used in the experimental work, (an Atlas-Copco Wagner ST-7) the drivetrain delivers equal power to both the front and rear wheel sets through a single transmission. This guarantees that wheel slip must occur when the rate of change of the articulation angle is non-zero. This means that the constraint of zero velocity in the direction of the axles (rolling motion constraint) is not valid, and in fact this model greatly overestimates the rate of change of orientation $\dot{\phi}$. A kinematic model that explicitly models this slip is therefore developed as follows.

5.3.7 Accounting for Slip

To take into account that the vehicle will slip during motion, two slip variables α and β are introduced. These variables are chosen to represent the slip angles of the vehicle, that is, the angle between the kinematically indicated velocity (perpendicular to the axles) and the

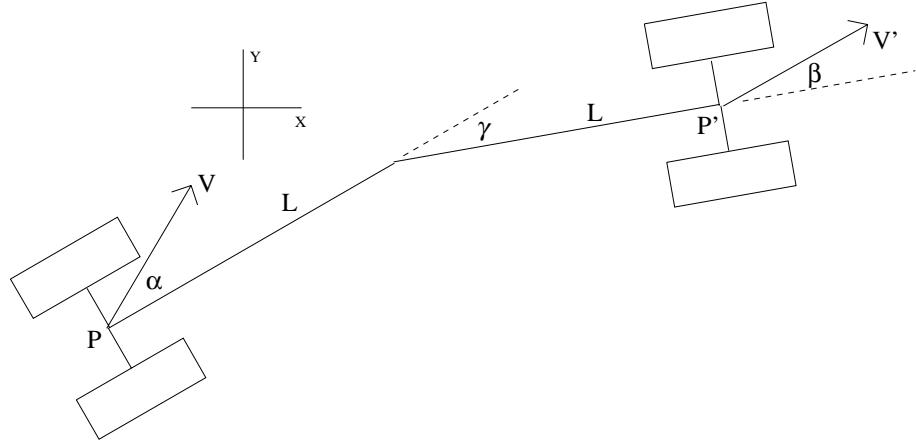


Figure 5.36: LHD Kinematic Geometry - Including Slip Angles

true velocity. The variation between true and kinematic velocities is by definition entirely dependent on the slip.

Consider the velocity of the point P in the rear body, and the velocity at a symmetric point P' in the front body as shown in Figure 5.36. Generally, the prime will indicate quantities related to the front of the vehicle, whilst no prime indicates quantities related to the rear.

By obtaining the velocity of P' in the direction perpendicular to V' and equating this to zero, we obtain the following relation;

$$V \sin(\alpha - \beta - \gamma) + L \cos(\beta + \gamma) \dot{\phi} + L \cos(\beta)(\dot{\phi} + \dot{\gamma}) = 0 \quad (5.8)$$

which may be solved for $\dot{\phi}$ to yield

$$\dot{\phi} = \frac{(\omega R \sin(\beta - \alpha + \gamma) - \dot{\gamma} L \cos(\beta))}{L(\cos(\beta) + \cos(\beta + \gamma))} \quad (5.9)$$

If the velocity V of P is now set to equal ωR , the wheel's angular velocity multiplied by the nominal wheel radius, and resolved into x and y , the following kinematic equations for

the motion of point P are obtained;

$$\begin{aligned}\dot{x} &= \omega R \cos(\alpha + \phi) \\ \dot{y} &= \omega R \sin(\alpha + \phi) \\ \dot{\phi} &= \frac{(\omega R \sin(\beta - \alpha + \gamma) - \dot{\gamma} L \cos(\beta))}{L(\cos(\beta) + \cos(\beta + \gamma))}\end{aligned}\quad (5.10)$$

It can be seen that the vehicle moves in the direction given by the slip angle, ie the vehicle heads in the direction $\alpha + \phi$, whilst the rate of change of orientation is dependant on the slip angles, the articulation angle, and the rate of change of articulation angle.

5.3.8 The Importance of Slip

Comparing Equation 5.10 to Equation 5.7, it can be seen that the models are significantly different, particularly for the term $\dot{\phi}$. The simple model described by Equation 5.7 greatly overestimates the turning rate of the vehicle, causing the navigation system to continuously fight to correct the modelling error (a typical process fault). Equation 5.10 is far more accurate if the slip angles are known. The Extended Kalman Filter (EKF) is very powerful in this respect, as states that are not directly observable, such as the slip parameters, may be estimated and used to improve the model of the vehicle.

As the vehicle drives both the front and rear wheel sets at the same angular velocities, insight into why the vehicle does indeed slip may be gained by examining the different expressions for V and V' . The equation governing V' is;

$$V' = \frac{V(\cos(\alpha - \gamma) + \cos(\alpha)) - L\dot{\gamma} \sin(\gamma)}{\cos(\beta) + \cos(\beta + \gamma)} \quad (5.11)$$

When the slip angles α and β are zero (ie no slip is present), we obtain the following expression for the velocity of the front wheel.

$$V' = V - L\dot{\gamma} \tan\left(\frac{\gamma}{2}\right) \quad (5.12)$$

It is clear from this that the only conditions under which the front and rear velocities are equal (under no-slip assumptions) is when the articulation angle γ is constant or zero.

Under normal operating conditions, wheel slip *must* be present.

Experimental data shows that the rate of change of articulation angle $\dot{\gamma}$ can be quite high, and may cause the vehicles tyres to scrub the ground. Human operators tend to use this effect to their advantage.

5.3.9 The Error Model

When the model addressed in Equation 5.10 is considered, it can be seen that the primary sources of error are due to the time varying parameters ω , γ , $\dot{\gamma}$, α and β , as errors in these parameters propagate directly through to the states. The variables ω , γ and $\dot{\gamma}$ represent well known control inputs and thus do not need to be estimated. The slip parameters α and β , however, are not directly measured, and therefore cannot be treated in the same way as the control inputs. It is interesting to note that the effective wheel radius of the vehicle can also be considered to be time varying, taking into account factors such as loading and tyre wear. The wear on a typical LHD tyre can be as much as 15 – 20cm in radius and will therefore introduce a bias to the system over the operating lifetime of the tyres if a constant wheel radius is assumed. The concept of estimating wheel radius was introduced in [28], and is extremely beneficial to this application. Therefore, the states to be estimated should not only include position and orientation, but also the slip angles and wheel radius.

It is salient at this point to describe how the error in each of the parameters may be modelled, as it is important to know how the errors propagate through time. The errors in control input are modelled as simple additive noise $\delta\omega(t)$, $\delta\gamma(t)$ and $\delta\dot{\gamma}(t)$ about their respective means $\bar{\omega}(t)$, $\bar{\gamma}(t)$ and $\dot{\bar{\gamma}}(t)$ at time t such that

$$\begin{aligned}\omega(t) &= \bar{\omega}(t) + \delta\omega(t) \\ \gamma(t) &= \bar{\gamma}(t) + \delta\gamma(t) \\ \dot{\gamma}(t) &= \dot{\bar{\gamma}}(t) + \delta\dot{\gamma}(t)\end{aligned}\tag{5.13}$$

The errors in R , α and β , however, are extremely difficult to model accurately as they tend to involve a combination of other parameters, and are caused fundamentally by the vehicle dynamics. For example, the error in slip angle will change with respect to vehicle speed, mass, tyre-terrain interaction and articulation angle in a highly non-linear way. A compromise which has been found to work well is to model the errors in these parameters

as random walks (or Brownian motion), such that the error in each of the parameters is the integral of white noise as follows,

$$\begin{aligned}\dot{R}(t) &= \delta R(t) \\ \dot{\alpha}(t) &= \delta \alpha(t) \\ \dot{\beta}(t) &= \delta \beta(t)\end{aligned}\tag{5.14}$$

The noise sources $\delta\omega(t)$, $\delta\gamma(t)$, $\delta\dot{\gamma}(t)$, $\delta R(t)$, $\delta\alpha(t)$ and $\delta\beta(t)$ are assumed zero-mean, uncorrelated Gaussian sequences with variance σ_ω^2 , σ_γ^2 , $\sigma_{\dot{\gamma}}^2$, σ_R^2 , σ_α^2 and σ_β^2 respectively. Although in practice these parameters may not evolve in a strictly Brownian manner, the Brownian model reflects the growth in uncertainty in their true value, and the rate at which the true value is considered to vary.

The continuous time vehicle model shown in Equation 5.10 may now be rewritten at time t , adding the additional states to be estimated, as:

$$\begin{aligned}\dot{x}(t) &= \omega(t)R(t) \cos(\alpha(t) + \phi(t)) \\ \dot{y}(t) &= \omega(t)R(t) \sin(\alpha(t) + \phi(t)) \\ \dot{\phi}(t) &= \frac{\omega(t)R(t) \sin(\beta(t) - \alpha(t) + \gamma(t))}{L(\cos(\beta(t)) + \cos(\beta(t) + \gamma(t)))} - \frac{\dot{\gamma}(t)L \cos(\beta(t))}{L(\cos(\beta(t)) + \cos(\beta(t) + \gamma(t)))} \\ \dot{R}(t) &= \delta R(t) \\ \dot{\alpha}(t) &= \delta \alpha(t) \\ \dot{\beta}(t) &= \delta \beta(t)\end{aligned}\tag{5.15}$$

5.3.10 Observation Model

The navigation system described here uses the GCS laser described in the previous experiment. Also used are the three gyroscopes contained in the IMU unit.

The information from the three gyros is first converted to a single rotation about the vehicle's z axis, then fused in the filter to provide a better estimate of state.

As seen in Section 5.2, the errors in the gyro are adequately modelled as Brownian motion, the integral of white noise. The state vector must therefore be augmented with a shaping

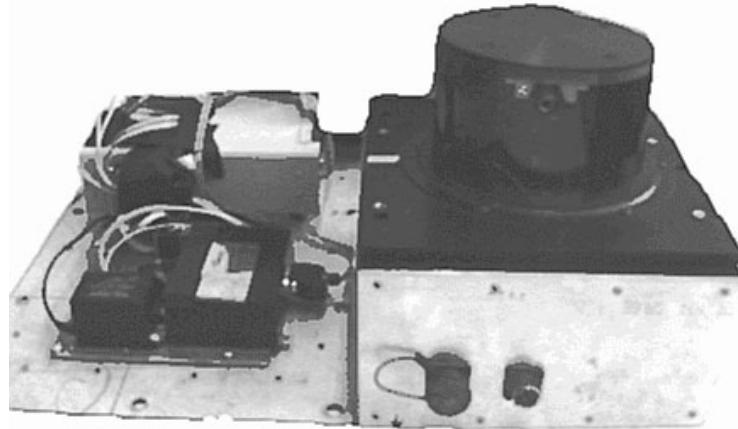


Figure 5.37: The IMU and Laser used in the field trials

state as follows.

$$\mathbf{x}(t) = [x(t), y(t), \phi(t), R(t), \alpha(t), \beta(t), x_{sf}(t)]^T \quad (5.16)$$

The observation equations are the identical to those presented in Section 5.2, and are

$$\mathbf{z}_\theta^i(t) = \left[\arctan\left(\frac{Y_i - y(t)}{X_i - x(t)}\right) - \phi(t) \right] + [v_\theta^i(t)] \quad (5.17)$$

and

$$\mathbf{z}_{gyro}(t) = [\phi(t) + x_{sf}(t)] + [v_{gyro}(t)] \quad (5.18)$$

for the laser and gyro respectively.

5.3.11 Implementing the Navigation System

To implement a Kalman Filter based on a continuous time model, the model must be discretised, as the control signals ω and γ and the sensor observations are sampled at regular discrete intervals.

The discrete-time Kalman filter derivation for this system is found in Appendix F.

5.3.12 Results

This section presents the results of implementing the geometric-based navigation system for an LHD described previously. The results were obtained using data from the field trials on an LHD in an underground mine at Mt. Isa in Queensland, Australia. The sensors and their layout is shown in Figure 5.30. The sensor positions were measured accurately so that coordinate transformations could be reliably achieved. The experimental setup is described in Section 5.3.1, and more fully in [86].

The section of tunnel, populated with retroreflective strips used as beacons for the laser scanner, is shown in Figure 5.32. The strips were surveyed to provide their positions in the tunnel. As described previously, the data from all sensors on board the vehicle was logged and time-stamped during the trials, and post-processed to provide the results seen here.

Figure 5.38 shows the tracked path of the LHD during a greater than right angle turning manoeuvre. The maximum standard deviation, or error bound in position, during this manoeuvre (as estimated by the EKF) was approximately eight centimetres, as can be seen in Figure 5.39. The spikes in the standard deviation correspond to periods during the run when no beacons could be seen or accurately identified. Figures 5.41, 5.42 and 5.43 show the estimated wheel radius and the slip angles α and β respectively. These may be compared to Figure 5.40 which shows the variation in articulation angle over time. As can be seen from these graphs, the estimated slip is highly correlated with the articulation angle, verifying the intuitive assumption that the vehicle will slip more while cornering. This is of particular importance in underground mining where a majority of the tunnel corners are tight right angled bends. In this example, the forward slip angle β approaches six degrees during the corner, while the rear slip angle α exceeds twenty degrees. From the kinematic model of the vehicle shown in figure 5.36, it can be seen that the sign of the slip angles corresponds to the vehicle turning the corner faster than the no-slip model would allow. The magnitude of the slip angles indicates that the rear of the vehicle slips more than the front. Figure 5.41 shows that the estimated wheel radius quickly converges from the initial value of $0.9m$ to a value of approximately $0.75m$, which was in fact the correct wheel radius at the time of the trials. It is also interesting to note that the estimated wheel radius R decreases by approximately $0.1m$ during the turning manoeuvre. The reason for this effect is clear when it is considered that the vehicle's forward velocity is given by $V = \omega R$. Any change in the forward velocity of the vehicle (due to forward slip) is estimated as a change in wheel radius

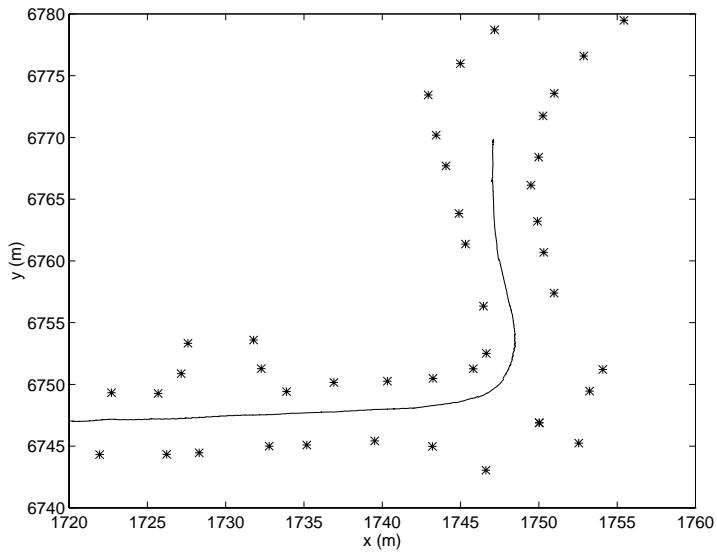


Figure 5.38: Estimated LHD Path - stars represent retroreflective markers

by the EKF. The slip angles account for the across axis, or transverse, slip.

Further experimental results (not shown) based on the data recorded during other runs show that the filter reliably estimates vehicle position over the entire operating range of vehicle speeds. The only noticeable difference between the slow and high speed runs is the estimated position standard deviation, which is larger at higher speeds.

Figure 5.44 shows the gyro innovation sequence for the course of the experimental run. It can be seen that the system is not consistent. This is not surprising considering the large number of nonlinearities that have not been modelled. The vehicle pitches and rolls to a significant degree, but the vehicle and sensor models assume the vehicle moves in two dimensions only. Therefore if this navigation system were used in conjunction with the fault detection techniques described in this thesis, the system would permanently be at fault.

What has been achieved, though, is a vehicle model that does not cause the filter to diverge (even though it is not consistent).

5.3.13 Summary

This section has described the early results of an ongoing project to design a high integrity navigation system for underground mining vehicles, subject to the constraint of minimum

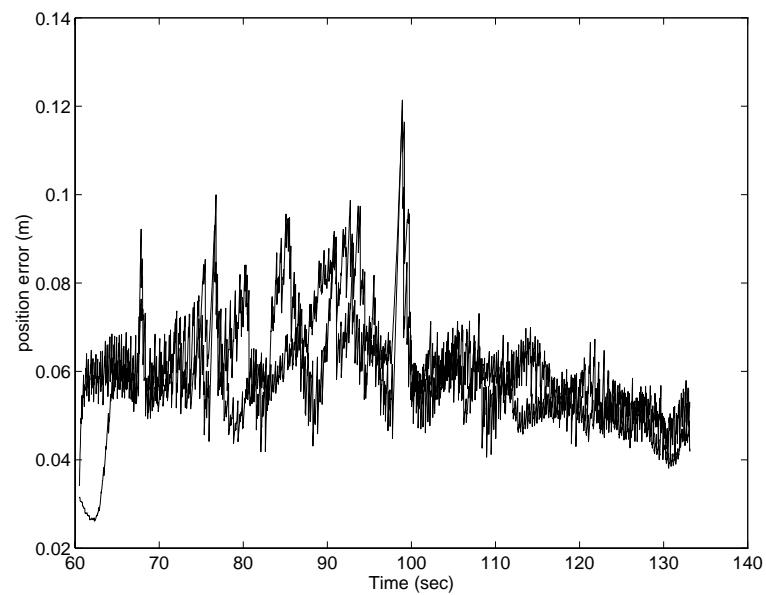


Figure 5.39: Position standard deviation during run (x and y)

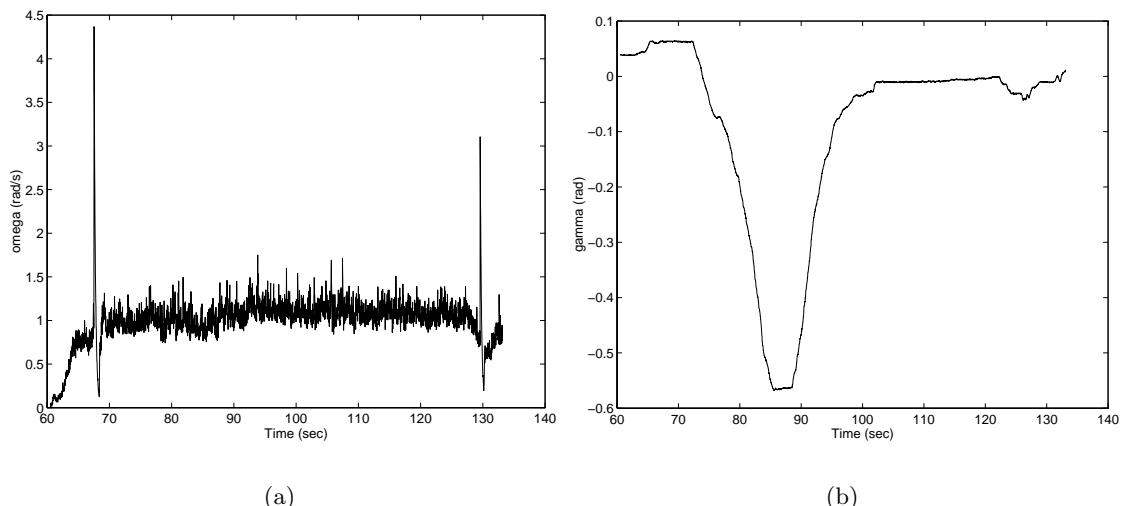


Figure 5.40: Vehicle Velocity and Steer Angle

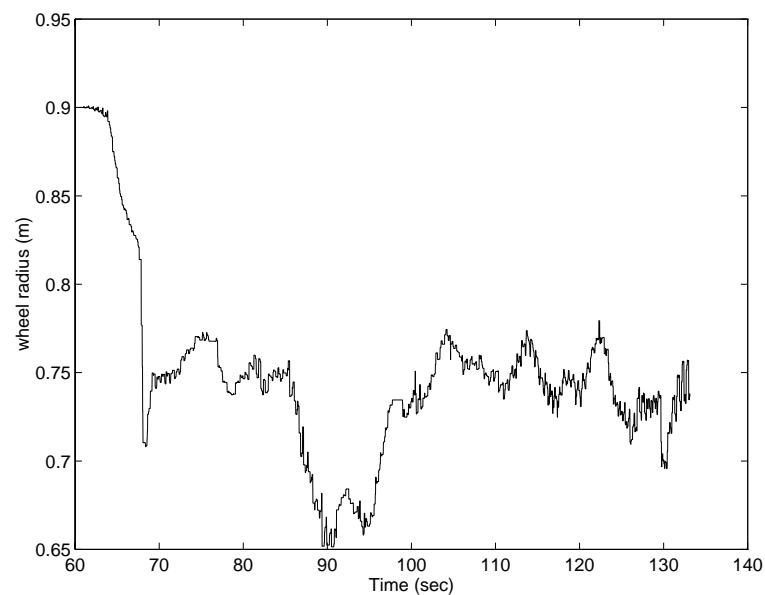


Figure 5.41: Estimated Wheel Radius

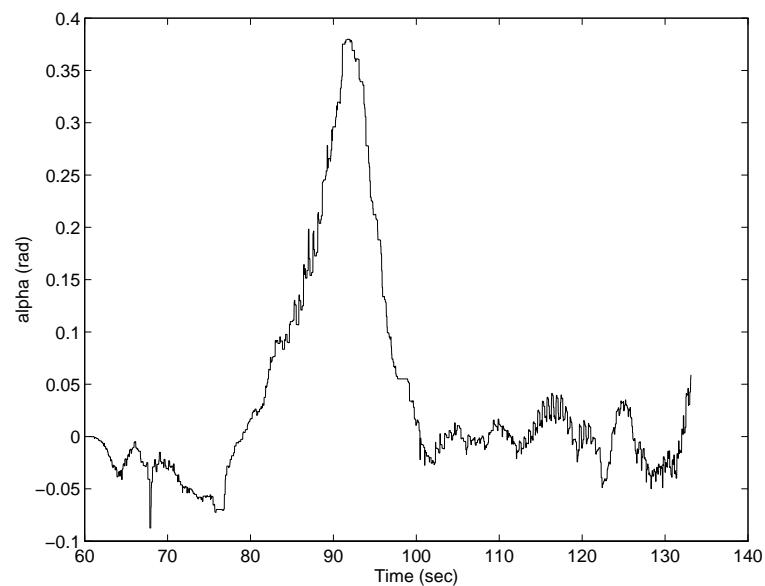


Figure 5.42: Estimated Slip Angle Alpha

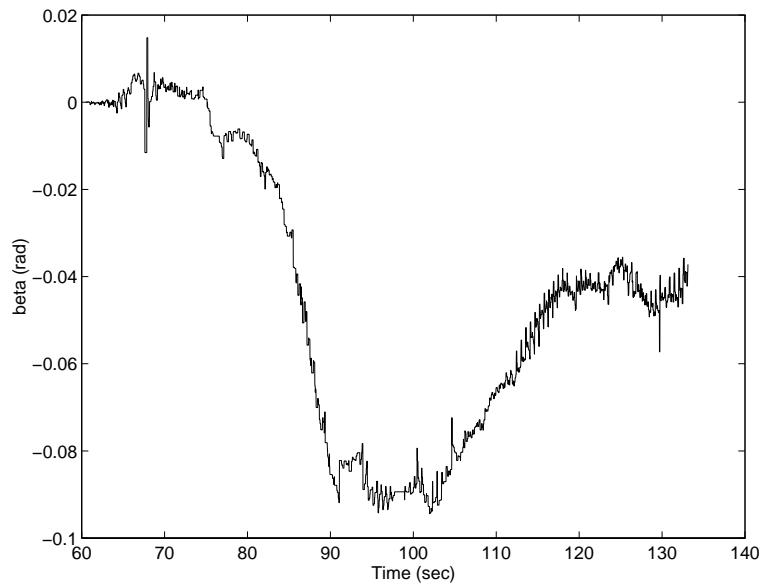
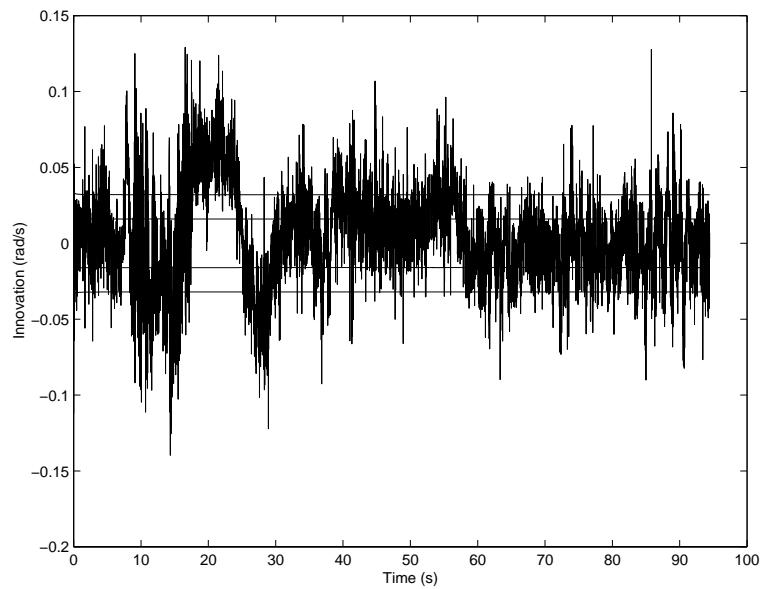


Figure 5.43: Estimated Slip Angle Beta

Figure 5.44: Gyro Innovation Sequence with 1σ and 2σ bounds

infrastructure. The proposed high integrity navigation architecture was described, and was shown to be proveably robust to failures. Early results from the two navigation loops were also presented.

The results from the initial geometric-based navigation system justify the effort of using models which explicitly account for common navigation errors. However, these results also showed that even better models must be used to ensure filter consistency.

5.4 Conclusion

The first half of this chapter (Section 5.2) described a simple experiment designed to verify and aid in the understanding of the theory presented in Chapter 3. The experiment utilised a gyro and a laser sensor to estimate the orientation of the experimental platform, and demonstrated examples of typical sensor and process faults. It was shown that in some cases faults may be undetectable, remotivating the need for redundancy.

The second half of this chapter (Section 5.3), described the early results of ongoing research into the design of a high integrity navigation system for underground mining vehicles. Two navigation systems were proposed: a feature-based solution, and a geometric-based solution. The geometric navigation system developed in this thesis explored the issues of vehicle and sensor modelling on an LHD underground mining vehicle. It was shown that conventional kinematic vehicle models are not adequate to describe the motion of the vehicle. A new vehicle model was derived that explicitly incorporates slip angles and tyre deformation parameters. The vehicle's position in the test tunnel was estimated using information supplied by a laser and gyro sensors.

Chapter 6

Summary and Conclusions

6.1 Introduction

This thesis aimed to provide the tools and insights needed to design a high integrity navigation system. High integrity navigation is essential to the commercial success of large outdoor AGV systems.

This chapter summarises the work developed in this thesis to achieve this goal, highlighting the major contributions made during the execution of this work.

In Section 6.2, a chapter by chapter summary of this thesis is presented. Section 6.3 outlines the contributions of this thesis, and finally, Section 6.4 suggests possible future research in this area.

6.2 Summary

In **Chapter 1**, the primary motivation for the work carried out in thesis was given. It was argued that autonomous navigation systems with the potential to do harm *must* be designed with high integrity. By designing a system with high integrity the designer is acknowledging the fact that no matter how reliable the system, it will eventually fail. This chapter also explored the concept of AGV navigation, introducing the Kalman filter as the preferred localisation algorithm.

Chapter 2 motivated the Kalman filter as the tool used in thesis to perform the task of localisation. In particular, Section 2.2 presented the continuous time Kalman filter together with its analysis in the frequency domain. The frequency domain was used as a tool for examining filter behaviour, allowing insights which are not immediately obvious in the time domain. Alternative filter structures were examined which explicitly take into account the frequency response of different sensors.

Section 2.3 examined the nature of failures within the Kalman filter framework by examining the assumptions that were made during the filter's derivation. It was concluded that process faults, observation faults and nonlinearities may cause the filter to fail. In Section 2.4, methods were presented that allow some common observation faults to be modelled and incorporated into the Kalman filter framework. Section 2.5 presented the EKF, which is widely used for the estimation of nonlinear systems. The EKF is essentially the equivalent to a time varying linear Kalman filter. In Section 2.6, the discrete Kalman filter was presented. The use of the discrete Kalman filter is motivated by the fact that most navigation systems execute on digital computers and incorporate sampled data. In Section 2.7, the issue of modelling process faults was examined. A simple example was given which showed the potential sources of failure and their effect.

A short survey of common navigation sensors was given in Section 2.8. This survey showed that sensors commonly fall into three categories, according to their noise properties. This motivated a study of sensor error and filter transfer which was given in Section 2.9. Section 2.10 examined the frequency interpretation of the EKF with the aid of a simple example.

The problem of detecting unmodelled faults was examined in **Chapter 3**. In Section 3.2, motivation was given for the need for fault detection. It was argued that for high integrity a guarantee must be given that all faults are detectable. In Section 3.3 a survey was given of existing fault detection techniques, highlighting the fact that no existing fault detection method provides a guarantee that all possible faults will be detected. Innovations testing techniques are used as the fault detection technique in this thesis for the reasons that they provide the only internal measure of filter consistency without compromising the Kalman filter's optimality.

Metrics for determining the detectability of process faults and sensor faults were developed in Section 3.4. These metrics made possible a theoretical framework which allowed techniques to be developed which effectively guarantee fault detectability, which were given in

Sections 3.5 and 3.6. For sensor fault detectability it was shown that frequency domain redundancy must be employed for all sensor faults to be detectable, and for process faults, it was shown that all affected states must be observed to ensure process fault detectability.

In Section 3.7 the detectability metrics developed previously were applied to a nonlinear system. It was shown that for bounded nonlinear and time varying linear systems, that the techniques developed for ensuring fault detectability apply equally well to these systems.

A recursive technique for detecting faults via innovations testing was presented in Section 3.8.

Chapter 4 expanded the concept of high integrity design. The first part of this chapter simply deals with the issue of designing a navigation system such that it meets some prespecified design requirements. Specifically, Section 4.2 introduced the typical design specifications that an AGV must meet for useful operation. These requirements were summarised as the maximum frequency response of the system and the minimum accuracy required to achieve a given task.

It was shown in Section 4.3 that to meet the design requirements, the choice of sensors is essentially the only variable, and that in many cases, sensor suites must be used to achieve the desired performance characteristics. Section 4.4 generalised the method of designing sensor suites for navigation systems. Rather than design by trial and error, the design specifications essentially dictate the quality of sensors that must be used. This design problem was identified as being closely linked to the algebraic Riccati equation, and a method was proposed to automate sensor suite selection based upon the numerical solution of the Riccati equation.

The second part of **Chapter 4** dealt with the issue of designing a high integrity navigation system. It was argued that for all sensor faults to be guaranteed detectable, one must simply provide frequency domain redundancy. In essence this means that there must be at least two complementary navigation systems. Further, it was argued that the navigation system should be implemented in a distributed fashion so that a computer hardware failure will not cause the whole system to fail. The information filter and a distributed Kalman filter were introduced as potential methods for achieving this goal.

In **Chapter 5** the results from two separate experiments were given. The first experiment detailed the design and operation of a simple experimental rig used to verify the theory

developed in this thesis regarding the propagation of faults in navigation systems. Results were shown which verified the theory.

The second experiment detailed in **Chapter 5** described initial results of an ongoing research project to design a high integrity navigation system for an underground mining vehicle. The proposed navigation architecture was described, highlighting the need for two physically independent navigation loops. Initial results from the first, feature-based navigation system were described briefly. The second, geometric-based, navigation system was developed in greater detail. The modelling used in this navigation system derived from ideas presented in this thesis, notably through the inclusion of parameters to allow the estimation of slip. Results were given that show that the extra modelling effort resulted in a navigation system that remained nondivergent.

6.3 Contributions

The major contribution of this thesis was the development of the theoretical tools needed to provide the guarantees necessary for a navigation system to be considered of high integrity. This section outlines the individual contributions of this thesis, which together have achieved this goal.

6.3.1 Sensor Error Models

This thesis presented an analysis of sensor error models and their effect on the transfer function of the Kalman filter. It was shown that navigation sensors may be grouped in three broad categories according to their noise characteristics. The effect on the Kalman filter transfer function of each of these three sensor types, and of combinations of these sensors was shown. It was argued that this type of analysis should influence the type of sensors that are chosen for use in a high integrity navigation system.

6.3.2 Navigation Failure

The issue of navigation failure was examined in detail. A navigation failure was identified as any fault that will cause the Kalman filter to produce inconsistent estimates of state.

Specifically, it was shown that a navigation failure is caused by any effect that invalidates the assumptions made in the Kalman filter derivation, whether caused by an incorrect mathematical model, or hardware failure.

6.3.3 Fault Detectability Metrics

An important contribution of this thesis was the development of metrics which describe the detectability of process and sensor faults in the filter innovations. These metrics make no assumptions as to the source or nature of the fault. A fault is characterised by its frequency content only.

The fault detectability metrics enabled the development of techniques which effectively guarantee that all process and sensor faults will be detectable.

6.3.4 Sensor Suite Selection

This thesis developed methods for the selection of the minimum quality sensors that guarantee that the navigation system design requirements are met. It was shown that this result is dependent on the solution of the algebraic Riccati equation, and that this equation is not in general solvable. Iterative numerical methods were developed to overcome this problem.

6.3.5 High Integrity Navigation System Design

An important contribution of this thesis was to quantify the essential elements that are necessary for an autonomous navigation system to have high integrity.

It was shown that a high integrity navigation system must have two important attributes. Firstly, it must meet the design requirements. This element was solved through the design of appropriate sensor suites.

The second requirement of a high integrity navigation system was identified to be fault detection, provided a guarantee of the detectability of all faults exists. It was shown that in order to provide this guarantee, sensor redundancy must exist, preferably using sensors based on different physical principles.

This thesis also showed that to take into account the possibility of software and hardware failure, the high integrity navigation system should be implemented in a distributed form.

6.3.6 Experimental Investigations

The final contributions of this thesis were experimental investigations into high integrity navigation.

The first experiment was designed as a testbed for the theory presented in the body of this thesis on the detectability of faults in navigation systems. Experimental results were shown to agree extremely well with the theory.

The second experiment showcased the initial results of an ongoing project to design a high integrity navigation system for an underground mining vehicle.

6.4 Future Work

Several areas related to high integrity navigation system design may be pursued in future research. Some of these areas are identified in this section.

6.4.1 Fault Detection

This thesis derived the theory needed to provide a guarantee that all faults are detectable by the navigation system. A recursive algorithm for the detection of these faults was suggested.

Important future work in this area should include an analysis as to the performance of the detection algorithm. Further, research should be conducted to identify and exploit other fault detection algorithms. Estimation techniques suggest themselves as ideal candidates for an alternate method of fault detection due to the stochastic nature of the filter innovations.

6.4.2 Identification

As the primary motivation of this thesis was the design of high integrity navigation systems, it was only desired to be able to detect that a fault has occurred.

Therefore, the issue of fault identification was not addressed. It should be possible, using techniques similar to those presented in this thesis, to develop a measure of the ability to identify faults. There is also scope for the development of algorithms to identify the source of a fault on-line.

Techniques such as this would greatly aid in the diagnosis of faults and therefore decrease the time to repair.

6.4.3 High Integrity LHD Navigation

There are several issues with the high integrity LHD navigation system presented in this thesis that need to be addressed.

The LHD vehicle model (although better than pre-existing articulated vehicle models) is simplistic. This simplicity was shown to affect the navigation system performance, producing estimates that although nondivergent, were inconsistent. To improve performance, the vehicle model should be extended to three dimensions to take into account the significant pitch and roll experienced by the vehicle in its operational environment. Further use of inertial sensors, including accelerometers and perhaps tilt sensors, could be used to assist in the estimation of the three dimensional vehicle parameters.

For the LHD navigation system to be considered to have high integrity, the geometric navigation solution outlined in Chapter 5 must be fully developed. In fact, research is currently being undertaken at the University of Sydney into this area. The two navigation solutions must then be combined with a fault detection scheme for fault detection to be guaranteed. Further research is also required to develop on-line implementations of both LHD navigation systems.

Appendix A

Power Spectral Density Estimation

To compute an accurate PSD, the sensor data must be collected with the sensor at a precisely known state so that the correlated noise may be estimated. With sensors such as GPS, this is usually achieved by taking the data sets with the sensor stationary.

Usually, 50 or more sequences are taken with length at least twice that of the longest time constant expected. The sample interval is Δt . The assumption must be made that the noise process is stationary and ergodic in order that ensemble averages may be replaced by time averages in the equations that follow.

After taking the data sets, to compute the PSD for the sensor the fast Fourier transform of the data is taken. Let $X_i(s)$ be the Fourier transform of the data

$$X_i(n\omega_s) = \sum_{k=0}^{M-1} x(t_0 + k\Delta t)e^{-jk\Delta tn\omega_s} \quad n = 0, \dots, M-1, \quad i = 1, \dots, N \quad (\text{A.1})$$

where $\omega_s = \frac{2\pi}{N\Delta t}$, N is the number of samples and M is the number of data points per sample. An estimate of the PSD may then be calculated as

$$\Psi_i(n\omega_s) = \frac{1}{N}|X_i(n\omega_s)|^2 \quad n = 0, \dots, M-1, \quad i = 1, \dots, N \quad (\text{A.2})$$

The corresponding autocorrelation $R_i(s)$ is found by taking the inverse Fourier transform

of the estimated PSD

$$R_i(n\Delta t) = \sum_{k=0}^{M-1} \Psi_i(k\omega_s) e^{jk\Delta tn\omega_s} \quad n = 0, \dots, M-1, \quad i = 1, \dots, N \quad (\text{A.3})$$

The autocorrelation is then multiplied by a hamming window to prevent frequency leakage

$$R'_i(n\Delta t) = \omega_n R_i(n\Delta t) \quad n = 0, \dots, M-1, \quad i = 1, \dots, N \quad (\text{A.4})$$

as the data sets are finite in length.

The smoothed PSD estimate is now found simply by taking the Fourier transform of the windowed autocorrelation, averaged over the number of samples.

$$\Psi(n\omega_s) = \frac{1}{N} \sum_{i=1}^N \sum_{k=0}^{M-1} R'_i(k\Delta t) e^{-jk\Delta tn\omega_s} \quad (\text{A.5})$$

Appendix B

Continuous to Discrete Time Conversion

The following equations show how continuous time processes may be discretised. Examples are given in Appendices E and F showing how discretisation may be performed with real systems.

Given the initial condition,

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (\text{B.1})$$

the solution to the linear differential equation given in equation 2.1 is

$$\mathbf{x}(t) = \mathbf{F}_d(t, t_0)\mathbf{x}_0 + \int_{t_0}^t \mathbf{F}_d(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau + \int_{t_0}^t \mathbf{F}_d(t, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau \quad (\text{B.2})$$

where \mathbf{F}_d is the discrete time state transition matrix and is given by[63]

$$\mathbf{F}_d(t, t_0) = \mathbf{F}_d(t - t_0) = e^{\mathbf{F}_c(t - t_0)} \quad (\text{B.3})$$

where \mathbf{F}_c is the continuous time process model.

Defining a sample interval ΔT such that

$$\Delta T \triangleq t - t_0, \quad (\text{B.4})$$

the state transition matrix becomes

$$\mathbf{F}_d(\Delta T) = e^{\mathbf{F}_c(\Delta T)} \quad (\text{B.5})$$

and the state at time k is given by

$$\begin{aligned} \mathbf{x}(k) = & \mathbf{F}_d(\Delta T)\mathbf{x}(k-1) + \int_{k-1}^k \mathbf{F}_d(\Delta T)\mathbf{B}(\tau)\mathbf{u}((k-1)\Delta T + \tau)d\tau \\ & + \int_{k-1}^k \mathbf{F}_d(\Delta T)\mathbf{G}(\tau)\mathbf{w}((k-1)\Delta T + \tau)d\tau \end{aligned} \quad (\text{B.6})$$

which is a complete discrete time representation of Equation 2.1.

Appendix C

Innovation Whiteness Testing

The innovation mean can be found as

$$\begin{aligned}
 \mathbb{E} [\nu(k) | \mathbf{Z}^{k-1}] &= \mathbb{E} [\mathbf{z}(k) - \hat{\mathbf{z}}(k|k-1) | \mathbf{Z}^{k-1}] \\
 &= \mathbb{E} [\mathbf{z}(k) | \mathbf{Z}^{k-1}] - \hat{\mathbf{z}}(k|k-1) \\
 &= 0
 \end{aligned} \tag{C.1}$$

with variance

$$\begin{aligned}
 \mathbf{S}(k) &= \mathbb{E} [\nu(k)\nu^T(k)] \\
 &= \mathbb{E} [(\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1))(\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1))^T] \\
 &= \mathbb{E} [(\mathbf{H}(k)[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1)] + \mathbf{v}(k))(\mathbf{H}(k)[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1)] + \mathbf{v}(k))^T] \\
 &= \mathbf{R}(k) + \mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}^T(k)
 \end{aligned} \tag{C.2}$$

So, to summarise, the innovation sequence is white with zero mean and variance $\mathbf{S}(k)$. To check whether the actual innovations indeed have these properties, two on-line tests must be performed. These are

- Test that the innovations are unbiased and
- Test that the innovations are uncorrelated.

which follow from the proof that when the filter is consistent, the innovations are white with zero mean.

The normalised innovations squared $q(k)$ are defined as

$$q(k) = \nu^T(k) \mathbf{S}^{-1}(k) \nu(k) \quad (\text{C.3})$$

which under the assumption that the filter is consistent can be shown to be a χ^2 random distribution with m degrees of freedom, where $m = \dim(\mathbf{z}(k))$, the dimension of the measurement sequence[7].

From this definition, it can be stated that

$$\mathbb{E}[q(k)] = m \quad (\text{C.4})$$

In principle, to compute the expected value (or mean) of $q(k)$, one must compute the sample mean averaged over N independent runs as

$$\bar{q}(k) = \frac{1}{N} \sum_{i=1}^N q^i(k) \quad (\text{C.5})$$

where $q^i(k)$ are the N runs of the filter.

However, in practice it is desirable that the fault detection algorithm be implemented online, so N samples will of course not be available. Because the innovation sequence is zero mean and white, it is therefore ergodic, and the sample mean computed in Equation C.5 may be replaced by a simple time average of a single sequence. The test in Equation C.5 may therefore be implemented as

$$\bar{q} = \frac{1}{N} \sum_{i=1}^N q(i) \quad (\text{C.6})$$

using the normalised innovations squared for a single run. Then $N\bar{q}(k)$ is a χ^2 distribution with Nm degrees of freedom. Therefore, to test the unbiasedness of the innovations, a confidence interval can be constructed using χ^2 tables, within which the average \bar{q} should lie assuming the hypothesis that the innovations are unbiased and white.

To test the whiteness (or uncorrelatedness) of the innovation sequence, one must prove that

$$\mathbb{E} [\nu^T(i)\nu(j)] = \mathbf{S}(i)\delta_{ij} \quad (\text{C.7})$$

which is achieved by computing the time averaged biased autocorrelation of the sequence as

$$\mathbf{r}(\tau) = \frac{1}{N} \sum_{i=0}^{N-\tau-1} \nu^T(i)\nu(i+\tau) \quad (\text{C.8})$$

assuming, again, that the innovation sequence is white and therefore ergodic. This test statistic $\mathbf{r}(\tau)$ should be approximately zero for an uncorrelated sequence. For large sample size, though, it is approximately normally distributed with standard deviation $\frac{1}{\sqrt{N}}$. So for 95% confidence bounds one must simply test that the autocorrelation $\mathbf{r}(\tau)$ lies within two standard deviations, $\pm \frac{2}{\sqrt{N}}$, of the mean, which in this case is zero.

Appendix D

A Numerical Solution of the Riccati Equation

In general, the steady state covariance matrix is found as the solution to the algebraic Riccati equation,

$$\mathbf{F}\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T + \mathbf{G}\mathbf{Q}(t)\mathbf{G}^T - \mathbf{P}(t)\mathbf{H}^T\mathbf{R}^{-1}(t)\mathbf{H}\mathbf{P}(t) = 0 \quad (\text{D.1})$$

In general, however, it is not possible to explicitly compute an algebraic solution (such as in the example of Section 4.4.1) for the steady state covariance matrix for state vectors of order greater than two[10]. As typical navigation systems use state vectors many times this size, a method must be developed to overcome this limitation.

The Riccati equation may be solved numerically by computing the Jordan form reduction of the associated Hamiltonian matrix \mathcal{H} where

$$\mathcal{H} = \begin{bmatrix} \mathbf{F}^T & -\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \\ -\mathbf{G}\mathbf{Q}\mathbf{G}^T & -\mathbf{F} \end{bmatrix} \quad (\text{D.2})$$

The Jordan form of \mathcal{H} is found by calculating the matrix of eigenvectors T that transform the Hamiltonian matrix into a diagonal form as

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{F}^T & -\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \\ -\mathbf{G}\mathbf{Q}\mathbf{G}^T & -\mathbf{F} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} -\Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \quad (\text{D.3})$$

where Λ is a diagonal matrix consisting of the eigenvalues in the left-half complex plane.

It can be shown [10, 13, 31] that the Riccati equation solution \mathbf{P} is found by solving the system of linear equations

$$\mathbf{P} = T_{21}T_{11}^{-1} \quad (\text{D.4})$$

Although this method of eigenvector decomposition works well in most practical situations, severe numerical difficulties can arise when the Hamiltonian matrix \mathcal{H} has multiple repeated eigenvalues (or even very close eigenvalues). To overcome this problem, the Hamiltonian may be decomposed using Schur vectors and a method nearly identical to the one presented above is used to solve for \mathbf{P} . Most matrix manipulation software such as Matlab contain code which calculates both steady state covariance and Kalman gain matrices by using either eigenvector or Schur decomposition methods.

Appendix E

The Gyro-Laser Experiment Discretisation Equations

E.1 Discretisation

From Equation B.6, the discrete time state transition matrix is found as

$$\begin{aligned}\mathbf{F}_d &= \exp(\mathbf{F}_c \Delta T) \\ &= \mathbf{I} + \Delta T \mathbf{F}_c + \frac{\Delta T^2 \mathbf{F}_c^2}{2} + \dots \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta T & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{E.1})\end{aligned}$$

As there are no control inputs in this system, the second term in Equation B.6 equates to zero.

The third term, the discrete time equivalent noise process is found as

$$\begin{aligned}
 \mathbf{w}(k) &= \int_{k-1}^k \exp(\mathbf{F}_c(\Delta T - \tau)) \begin{bmatrix} w_x((k-1)\Delta T + \tau) \\ w_y((k-1)\Delta T + \tau) \\ 0 \\ w_{\dot{\phi}}((k-1)\Delta T + \tau) \\ w_{sf}((k-1)\Delta T + \tau) \end{bmatrix} d\tau \\
 &= \int_{k-1}^k \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta T - \tau & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_x((k-1)\Delta T + \tau) \\ w_y((k-1)\Delta T + \tau) \\ 0 \\ w_{\dot{\phi}}((k-1)\Delta T + \tau) \\ w_{sf}((k-1)\Delta T + \tau) \end{bmatrix} d\tau \\
 &= \int_{k-1}^k \begin{bmatrix} w_x((k-1)\Delta T + \tau) \\ w_y((k-1)\Delta T + \tau) \\ (\Delta T - \tau)w_{\dot{\phi}}((k-1)\Delta T + \tau) \\ w_{\dot{\phi}}((k-1)\Delta T + \tau) \\ w_{sf}((k-1)\Delta T + \tau) \end{bmatrix} d\tau \tag{E.2}
 \end{aligned}$$

The discrete time process noise covariance matrix is given by

$$\begin{aligned}
 \mathbf{Q} &= \mathbb{E}[\mathbf{w}(i)\mathbf{w}^T(j)] \\
 &= \mathbb{E} \left\{ \int_{k-1}^k \begin{bmatrix} w_x(i\Delta T + \tau) \\ w_y(i\Delta T + \tau) \\ (\Delta T - \tau)w_{\dot{\phi}}(i\Delta T + \tau) \\ w_{\dot{\phi}}(i\Delta T + \tau) \\ w_{sf}(i\Delta T + \tau) \end{bmatrix} \begin{bmatrix} w_x(j\Delta T + \tau) \\ w_y(j\Delta T + \tau) \\ (\Delta T - \tau)w_{\dot{\phi}}(j\Delta T + \tau) \\ w_{\dot{\phi}}(j\Delta T + \tau) \\ w_{sf}(j\Delta T + \tau) \end{bmatrix}^T d\tau \right\} \tag{E.3}
 \end{aligned}$$

If the process noise is assumed uncorrelated,

$$\mathbb{E}[w(t)w(\tau)] \equiv \sigma^2 \delta(t - \tau) \tag{E.4}$$

therefore

$$\begin{aligned}
 \mathbf{Q}(k) &= \int_{k-1}^k \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 & 0 \\ 0 & 0 & (\Delta T - \tau)^2 \sigma_{\dot{\phi}}^2 & (\Delta T - \tau) \sigma_{\dot{\phi}}^2 & 0 \\ 0 & 0 & (\Delta T - \tau) \sigma_{\dot{\phi}}^2 & \sigma_{\dot{\phi}}^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{sf}^2 \end{bmatrix} d\tau \\
 &= \begin{bmatrix} \Delta T \sigma_x^2 & 0 & 0 & 0 & 0 \\ 0 & \Delta T \sigma_y^2 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} \Delta T^3 \sigma_{\dot{\phi}}^2 & \frac{1}{2} \Delta T^2 \sigma_{\dot{\phi}}^2 & 0 \\ 0 & 0 & \frac{1}{2} \Delta T^2 \sigma_{\dot{\phi}}^2 & \Delta T \sigma_{\dot{\phi}}^2 & 0 \\ 0 & 0 & 0 & 0 & \Delta T \sigma_{sf}^2 \end{bmatrix} \quad (E.5)
 \end{aligned}$$

Appendix F

LHD Discretisation Equations

F.1 Discrete Vehicle Model

The control signals ω and γ and the sensor observations are sampled at regular discrete intervals. The sample interval ΔT is synchronous and small enough to capture manoeuvres occurring in the operating frequency range. The control inputs are assumed approximately constant over this interval. Replacing the continuous time index t with a discrete time index k , Equation 5.15 becomes

$$\begin{aligned}
 x(k+1) &= x(k) + \Delta T \omega(k) R(k) \cos(\alpha(k) + \phi(k)) \\
 y(k+1) &= y(k) + \Delta T \omega(k) R(k) \sin(\alpha(k) + \phi(k)) \\
 \phi(k+1) &= \phi(k) + \Delta T \frac{(\omega(k) R(k) \sin(\beta(k) - \alpha(k) + \gamma(k)) - \dot{\gamma}(k) L \cos(\beta(k)))}{L(\cos(\beta(k)) + \cos \beta(k) + \gamma(k)))} \\
 R(k+1) &= R(k) + \Delta T \delta R(k) \\
 \alpha(k+1) &= \alpha(k) + \Delta T \delta \alpha(k) \\
 \beta(k+1) &= \beta(k) + \Delta T \delta \beta(k) \\
 x_{sf}(k+1) &= x_{sf}(k) + \Delta T \delta \mathbf{w}_{sf}(k)
 \end{aligned} \tag{F.1}$$

The discrete time state vector at time $k + 1$ can now be defined as

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k), k) + \mathbf{v}(k) \\ &= [x(k+1), y(k+1), \phi(k+1), R(k+1), \alpha(k+1), \beta(k+1), x_{sf}(k+1)]^T\end{aligned}\quad (\text{F.2})$$

where

$$\mathbf{u}(k) = [\omega(k), \gamma(k), \dot{\gamma}(k)]^T$$

is the control vector at time k , $\mathbf{v}(k)$ is additive process noise representing the uncertainty in slip angles and wheel radius, and $\mathbf{f}(\cdot)$ is the nonlinear state transition function mapping the previous state and current control inputs to the current state, here represented by Equation F.1.

F.2 State Prediction

Assume the system has an estimate available at time k which is equal to the conditional mean

$$\hat{\mathbf{x}}(k|k) = \mathbb{E}[\mathbf{x}(k)|\mathbf{Z}^k] \quad (\text{F.3})$$

The prediction $\hat{\mathbf{x}}(k+1|k)$ for the state at time $k+1$ based on information up to time k is given by expanding Equation F.2 as a Taylor series about the estimate $\hat{\mathbf{x}}(k|k)$, eliminating second and higher order terms, and taking expectations conditioned on the first k observations, giving

$$\hat{\mathbf{x}}(k+1|k) = \mathbb{E}[x(k+1)|\mathbf{Z}^k] = \mathbf{f}(\mathbf{x}(k|k), \mathbf{u}(k), k) \quad (\text{F.4})$$

The vector state prediction function $\mathbf{f}(\cdot)$ is defined by Equation F.1 assuming zero process and control noise. The prediction of state is therefore obtained by simply substituting the previous state and current control inputs into the state transition equation with no noise.

Define the noise source vector as

$$\delta\mathbf{w}(k) = [\delta\omega(k), \delta\gamma(k), \delta\dot{\gamma}(k), \delta R(k), \delta\alpha(k), \delta\beta(k), \delta\mathbf{w}_{sf}(k)]^T$$

The error between the true state and the estimated state is given by

$$\delta \mathbf{x}(k|k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k) \quad (\text{F.5})$$

the prediction of covariance is then obtained as

$$\mathbf{P}(k+1|k) = E[\delta \mathbf{x}(k|k)\delta \mathbf{x}(k|k)^T | \mathbf{z}^k] \quad (\text{F.6})$$

$$= \nabla \mathbf{f}_x(k) \mathbf{P}(k|k) \nabla \mathbf{f}_x^T(k) + \nabla \mathbf{f}_w(k) \Sigma(k) \nabla \mathbf{f}_w^T(k) \quad (\text{F.7})$$

where $\nabla \mathbf{f}_x(k)$ represents the gradient or Jacobian of $\mathbf{f}(\cdot)$ evaluated at time k with respect to the states, $\nabla \mathbf{f}_w(k)$ is the Jacobian of $\mathbf{f}(\cdot)$ with respect to the noise sources, and $\Sigma(k)$ is the noise strength matrix given by

$$\Sigma(k) = \begin{bmatrix} \sigma_\omega^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_\gamma^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\dot{\gamma}}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_R^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_\alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_\beta^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{sf}^2 \end{bmatrix} \quad (\text{F.8})$$

$\nabla \mathbf{f}_x(k)$ and $\nabla \mathbf{f}_w(k)$ are given by the following,

$$\nabla \mathbf{f}_x(k) = \begin{bmatrix} 1 & 0 & f_{13} & f_{14} & f_{15} & 0 & 0 \\ 0 & 1 & f_{23} & f_{24} & f_{25} & 0 & 0 \\ 0 & 0 & 1 & f_{34} & f_{35} & f_{36} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{F.9})$$

where,

$$\begin{aligned}
f_{13} &= -\Delta T \bar{\omega}(k) \hat{R}(k|k) \sin(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
f_{14} &= \Delta T \bar{\omega}(k) \cos(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
f_{15} &= -\Delta T \bar{\omega}(k) \hat{R}(k|k) \sin(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
f_{23} &= \Delta T \bar{\omega}(k) \hat{R}(k|k) \cos(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
f_{24} &= \Delta T \bar{\omega}(k) \sin(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
f_{25} &= \Delta T \bar{\omega}(k) \hat{R}(k|k) \cos(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
f_{34} &= \frac{\Delta T \bar{\omega}(k) \sin(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))} \\
f_{35} &= -\frac{\Delta T \bar{\omega}(k) \hat{R}(k|k) \sin(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))} \\
f_{36} &= \Delta T \left[\begin{array}{c} \frac{\dot{\gamma}(k)L \sin(\hat{\beta}(k|k)) + \bar{\omega}(k)\hat{R}(k|k) \cos(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))} \\ -\frac{(\sin(\hat{\beta}(k|k)) + \sin(\hat{\beta}(k|k) + \bar{\gamma}(k)))(\bar{\gamma}(k)L \cos(\hat{\beta}(k|k)) - \bar{\omega}(k)\hat{R}(k|k) \sin(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k)))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))^2} \end{array} \right]
\end{aligned}$$

$$\nabla \mathbf{f}_w(k) = \begin{bmatrix} g_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ g_{21} & 0 & 0 & 0 & 0 & 0 & 0 \\ g_{31} & g_{32} & g_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta T & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta T & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta T \end{bmatrix}$$

where

$$\begin{aligned}
g_{11} &= \Delta T \hat{R}(k|k) \cos(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
g_{21} &= \Delta T \hat{R}(k|k) \sin(\hat{\alpha}(k|k) + \hat{\phi}(k|k)) \\
g_{31} &= \frac{\Delta T \hat{R}(k|k) \sin(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))} \\
g_{32} &= \Delta T \left[\begin{array}{c} \frac{\bar{\omega}(k)\hat{R}(k|k) \cos(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))} \\ -\frac{\sin(\hat{\beta}(k|k) + \bar{\gamma}(k))(\bar{\gamma}(k)L \cos(\hat{\beta}(k|k)) - \bar{\omega}(k)\hat{R}(k|k) \sin(\hat{\beta}(k|k) - \hat{\alpha}(k|k) + \bar{\gamma}(k)))}{L(\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k)))^2} \end{array} \right] \\
g_{33} &= -\frac{\Delta T \cos(\hat{\beta}(k|k))}{\cos(\hat{\beta}(k|k)) + \cos(\hat{\beta}(k|k) + \bar{\gamma}(k))}
\end{aligned}$$

F.3 Observation Prediction

The predicted observation is found by taking expectations conditioned on all previous observations, truncating at first order to give

$$\begin{aligned}\hat{\mathbf{z}}(k+1|k) &= \mathbb{E}[\mathbf{z}(k+1)|\mathbf{Z}^k] \\ &= \mathbf{h}(\hat{\mathbf{x}}(k+1|k))\end{aligned}\quad (\text{F.10})$$

If there is a predicted state for the vehicle, $\hat{\mathbf{x}}(k+1|k)$, we can therefore predict the observations that will be made at that state. From Equations 5.17 and 5.18 and from Equation F.10, we have the predicted observations as

$$\mathbf{z}(k+1|k) = \begin{bmatrix} \hat{\mathbf{z}}_\theta^i(k+1|k) \\ \hat{\mathbf{z}}_{gyro}(k+1|k) \end{bmatrix} = \begin{bmatrix} \arctan\left(\frac{Y_i - \hat{y}(k+1|k)}{X_i - \hat{x}(k+1|k)}\right) - \hat{\phi}(k+1|k) \\ \hat{\phi}(k+1|k) + \hat{x}_{sf}(k+1|k) \end{bmatrix} \quad (\text{F.11})$$

Now, the innovation or observation prediction error covariance $\mathbf{S}(k)$, used in the calculation of the Kalman gains must be computed. This is found by squaring the estimated observation error and taking expectations of the first k measurements to give the following,

$$\mathbf{S}(k+1) = \nabla \mathbf{h}_x(k+1) \mathbf{P}(k+1|k) \nabla \mathbf{h}_x^T(k+1) + \mathbf{R}(k+1) \quad (\text{F.12})$$

In this case, the Jacobian $\nabla \mathbf{h}_x(k+1)$ is given by

$$\nabla \mathbf{h}_x(k+1) = \begin{bmatrix} -\frac{\hat{y}(k+1|k) - Y_i}{d^2} & \frac{\hat{x}(k+1|k) - X_i}{d^2} & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{F.13})$$

where $d = \sqrt{(X_i - \hat{x}(k+1|k))^2 + (Y_i - \hat{y}(k+1|k))^2}$ is the predicted distance between a beacon and the vehicle.

The observation variance term $\mathbf{R}(k)$ is given by

$$\mathbf{R}(k) = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_{gyro}^2 \end{bmatrix} \quad (\text{F.14})$$

F.4 Update Equations

The equation for the estimate of covariance is identical to that given in Section 2.6, however the state update equation for the EKF is given by

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)[\mathbf{z}(k+1) - \mathbf{h}(\hat{\mathbf{x}}(k+1|k))] \quad (\text{F.15})$$

where

$$\mathbf{K}(k+1) = \mathbf{P}(k+1|k)\nabla\mathbf{h}_x^T(k+1)\mathbf{S}(k+1)^{-1} \quad (\text{F.16})$$

Bibliography

- [1] A. Agogino, K. Goebel, and S. Alag. Intelligent Sensor Validation and Sensor Fusion for Reliability and Safety Enhancement in Vehicle Control. Technical Report UCB-ITS-PRR-95-40, University of California, 1995.
- [2] B. Anderson and J. Moore. *Optimal Filtering*. Prentice Hall, New Jersey, 1979.
- [3] B. Anderson and J. Moore. *Optimal Control*. Prentice Hall, New Jersey, 1989.
- [4] Anonymous. Fault Tree Handbook. Technical Report NUREG-0492, Nuclear Regulatory Commission, 1981.
- [5] G. Baiden. Multiple LHD Teleoperation and Guidance at Inco Limited. In *Proceedings of the International Mining Congress*, 1993.
- [6] Y. Bar-Shalom and T.E. Fortman. *Tracking and Data Association*. Academic Press, 1988.
- [7] Y. Bar-Shalom and X. Li. *Estimation and Tracking: Principles, Techniques and Software*. Artech House, Boston, 1993.
- [8] T. Basar and P. Bernhard. \mathcal{H}_∞ *Optimal Control and Related Minimax Design Problems : a Dynamic Game Approach*. Birkhauser, Boston, 1995.
- [9] P. Besl and N. McKay. A Method for Registration of 3-D Shapes. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 239–256, 1992.
- [10] S. Bittanti, A. Laub, and J. Willems. *The Riccati Equation*. Springer-Verlag, Berlin, 1991.
- [11] M. Blanke, R. Izadi-Zamanabadi, S. Bogh, and C. Lunau. Fault-Tolerant Control Systems - A Holistic View. *Control Engineering Practise*, 5:693–702, 1997.
- [12] R. Brown and P. Hwang. *Introduction to Random Signals and Applied Kalman Filtering 2nd Ed.* John Wiley and Sons, 1992.
- [13] R. Bucy and P. Joseph. *Filtering for Stochastic Processes with Applications to Guidance*. John Wiley and Sons, 1968.
- [14] R. Byrne. Global Positioning System Receiver Evaluation Results. Technical Report SAND93-0827, Sandia National Laboratories, 1993.

- [15] J. Cavallaro and I. Walker. A Survey of NASA and Military Standards on Fault Tolerance and Reliability Applied to Robotics. In *Proceedings AIAA-NASA Conference on Intelligent Robots in Field, Factory, Service, and Space*, pages 282–286, 1994.
- [16] J. Center. Practical Nonlinear Filtering of Discrete Observations by Generalised Least Squares Approximation of the Conditional Probability Distribution. In *Proceedings of the 2nd. Symposium on Nonlinear Estimation*, pages 88–99, 1971.
- [17] R. Chatila and S. Lacroix. Adaptive Navigation for Autonomous Mobile Robots. *International Symposium on Robotics Research*, 1995.
- [18] J. Chen, R. Patton, and G. Liu. Optimal Residual Design for Fault Diagnosis using Multi-Objective Optimisation and Genetic Algorithms. *International Journal of Systems Science*, 27:567–576, 1996.
- [19] H. Chin and K. Danai. A Method of Fault Signature Extraction for Improved Diagnosis. *Journal of Dynamic Systems, Measurement and Control*, 113:634–638, 1991.
- [20] D. Clarke. Sensor, Actuator and Loop Validation. *IEEE Control Systems*, pages 39–45, 1995.
- [21] S. Cooper. *A Frequency Response Method for Sensor Suite Selection with an Application to High-Speed Vehicle Navigation*. PhD thesis, University of Oxford, 1996.
- [22] S. Cooper and H. Durrant-Whyte. A Frequency Response Method for Multi-Sensor High-Speed Navigation Systems. In *IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, 1994.
- [23] I. Cox. Blanche - An Experiment in Guidance and Navigation of an Autonomous Robot Vehicle. *IEEE Transactions on Robotics and Automation*, 7:193–204, 1991.
- [24] X. Ding, L. Guo, and P. Frank. A Frequency Domain Approach to Fault Detection of Uncertain Dynamic Systems. In *Proc. 32nd Conf. Decision and Control*, 1993.
- [25] R. Douglas, J. Speyer, D. Mingori, R. Chen, D. Malladi, and W. Chung. Fault Detection and Identification with Application to Advanced Vehicle Control Systems. Technical Report UCB-ITS-PRR-95-26, University of California, 1995.
- [26] H. Durrant-Whyte. Where Am I? *Industrial Robot*, 21(2), 1994.
- [27] H. Durrant-Whyte, E. Nebot, S. Scheding, S. Sukkarieh, and S. Clark. The Design of Ultra-High Integrity Navigation Systems for Large Autonomous Vehicles. In *International Symposium on Experimental Robotics*, 1997.
- [28] H. F. Durrant-Whyte. An Autonomous Guided Vehicle for Cargo Handling Applications. *International Journal of Robotics Research*, 15, 1996.
- [29] H. Evers, G. Kasties, and R. Teske. Differential GPS in a Real Time Land Vehicle Environment - Satellite Based Van Carrier Location Equipment. In *IEEE Location and Navigation Symposium*, pages 361–365, 1994.

- [30] L. Feng, J. Borenstein, and H. Everett. Where Am I? : Sensors and Methods for Autonomous Mobile Robot Positioning. Technical report, University of Michigan, 1994.
- [31] G. Franklin, J. Powell, and M. Workman. *Digital Control of Dynamic Systems 2nd Ed.* Addison-Wesley, New York, 1990.
- [32] A. Gelb. *Applied Optimal Estimation*. The MIT Press, 1974.
- [33] J. Gertler. Survey of Model-Based Failure Detection and Isolation in Complex Plants. *IEEE Control Systems Magazine*, 8:3–11, 1988.
- [34] J. Gertler. Fault Detection and Isolation Using Parity Relations. *Control Engineering Practise*, 5:653–661, 1997.
- [35] R. Gianoglio. A Fault Tolerant Four Wheel Steering System. In *Proceedings of the American Control Conference*, 1991.
- [36] J. Gilmore and R. McKern. A Redundant Strapdown Inertial System Mechanisation - SIRU. In *Proceedings of AIAA Guidance, Control and Flight Mechanics Conference*, pages 17–19, 1970.
- [37] N. Gordon, D. Salmond, and A. Smith. A Novel Approach to Nonlinear-NonGaussian Bayesian State Estimation. *IEE Proceedings-F*, 140:107–113, 1993.
- [38] M. Grewal and A. Andrews. *Kalman Filtering: Theory and Practise*. Prentice Hall, 1993.
- [39] S. Grime and H. Durrant-Whyte. Data Fusion in Decentralised Sensor Networks. *Control Engineering Practise*, 2:849–863, 1994.
- [40] J. Handschin. Monte Carlo Techniques for Prediction and Filtering of Nonlinear Stochastic Processes. *Automatica*, 6:555–563, 1970.
- [41] H. Hashemipour, S. Roy, and A. Laub. Decentralised Structures for Parallel Kalman Filtering. *IEEE Transactions on Automatic Control*, 33:88–93, 1988.
- [42] P. Hwang and R. Brown. GPS Navigation: Combining Pseudorange with Continuous Carrier Phase Using a Kalman Filter. *Journal of the Institute of Navigation*, 37:181–196, 1990.
- [43] R. Isermann. Process Fault Detection Based on Modelling and Estimation Methods - A Survey. *Automatica*, 20:387–404, 1984.
- [44] R. Isermann. Supervision, Fault-Detection and Fault-Diagnosis Methods - An Introduction. *Control Engineering Practise*, 5:639–652, 1997.
- [45] R. Isermann and P. Balle. Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes. *Control Engineering Practise*, 5:709–719, 1997.
- [46] R. Isermann and B. Freyermuth. Process Fault Diagnosis Based on Process Model Knowledge - Part I: Principles for Fault Diagnosis with Parameter Estimation. *Journal of Dynamic Systems, Measurement and Control*, 113:620–626, 1991.

- [47] R. Isermann and B. Freyermuth. Process Fault Diagnosis Based on Process Model Knowledge - Part II: Case Study Experiments. *Journal of Dynamic Systems, Measurement and Control*, 113:627–633, 1991.
- [48] S. Julier. Process Models for the High-Speed Navigation of Conventional Road Vehicles. Technical report, University of Oxford, 1994.
- [49] S. Julier. *Process Models for the Navigation of High Speed Land Vehicles*. PhD thesis, University of Oxford, 1996.
- [50] J. Keller, L. Summerer, M. Boutayeb, and M. Darouach. Generalised Likelihood Ratio Approach for Fault Detection in Linear Dynamic Stochastic Systems with Unknown Inputs. *International Journal of Systems Science*, 27(12):1231–1241, 1996.
- [51] A. Kelly. A 3D State Space Formulation of a Navigation Kalman Filter for Autonomous Vehicles. *CMU Robotics Institute Technical Report*, 1994.
- [52] R. King. Underground Mining and Construction Robotics. In *Proceedings of SPIE Vol. 1613*, pages 348–358, 1991.
- [53] S. Kweon and T. Kanade. High-Resolution Terrain Map from Multiple Sensor Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:278–292, 1992.
- [54] F. Labonte and P. Cohen. Perceptual Aspects of Mining Equipment Teleoperation. In *Proceedings of the 6th Canadian Symposium on Mining Automation*, 1994.
- [55] F. Labonte, J. Giraud, and V. Polotski. Telerobotics Issues in the Operation of a LHD Vehicle. In *Proceedings of the Third Canadian Conference on Computer Applications in the Mineral Industry*, 1995.
- [56] D. Lane and P. Maybeck. Multiple Model Adaptive Estimation Applied to the Lambda URV for Failure Detection and Identification. In *Proc. 33rd Conf. Decision and Control*, 1994.
- [57] J.J. Leonard. *Directed Sonar Sensing for Mobile Robot Navigation*. PhD thesis, University of Oxford, 1991.
- [58] S. Leonhardt and M. Ayoubi. Methods of Fault Diagnosis. *Control Engineering Practice*, 5:683–692, 1997.
- [59] E. MacLeod and M. Chiarella. Navigation and Control Breakthrough for Automation Technology. In *Proceedings SPIE Conference on Mobile Robots VIII*, volume 2058, pages 57–68, 1993.
- [60] R. Madhavan, M. Dissanayake, and H. Durrant-Whyte. Autonomous Underground Navigation of an LHD Using a Combined ICP and EKF Approach. In *Submitted to IEEE International Conference of Robotics and Automation*, 1998.
- [61] H. Makela, H. Lehtinen, K. Rintanen, and K. Koskinen. Navigation System for LHD Machines. *Intelligent Autonomous Vehicles*, pages 314–319, 1995.

- [62] J. Manyika and H. Durrant-Whyte. *Data Fusion and Sensor Management: a Decentralised Information-Theoretic Approach*. Ellis Horwood, 1994.
- [63] P. Maybeck. *Stochastic Models, Estimation, and Control*, volume 1. Academic Press, New York, 1979.
- [64] P. McKerrow. *Introduction to Robotics*. Addison-Wesley, 1991.
- [65] T. Menke and P. Maybeck. Multiple Model Adaptive Estimation Applied to the Vista F-16 Flight Control System with Actuator and Sensor Failures. In *Proceedings IEEE National Aerospace and Electronics Conference*, 1992.
- [66] M. Nayebi and M. Aref. Limitations of the Gaussian Model for Radar Signals. In *Proceedings IEEE International Symposium on Information Theory*, page 76, 1997.
- [67] E. Nebot, H. Durrant-Whyte, and S. Scheding. Frequency Domain Modelling of Aided GPS with Application to High-Speed Vehicle Navigation Systems. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 1997.
- [68] E. Nebot, M. Karim, and J. Romagnoli. Implementation of a Failure Detection-Identification Algorithm for Dynamical Systems. *J. Int. Soc. Mini and Microcomputers*, 5:59–65, 1986.
- [69] E. Nebot, S. Sukkarieh, and H. Durrant-Whyte. Inertial Navigation Aided with GPS Information. In *Mechatronics and Machine Vision in Practise*, 1997.
- [70] V. Nelson. Safety in Numbers. *BYTE Magazine*, 16(8):175–184, 1991.
- [71] G.P.R. Netzler. Ndc news. Technical Report 6, Netzler & Dahlgren Company AB, SE-430 40 SARO, Sweden, 1994.
- [72] P. Olin and G. Rizzoni. Design of Robust Fault Detection Filters. In *Proceedings of the American Control Conference*, 1991.
- [73] A. Orrick, M. McDermott, D. Barnett, E. Nelson, and G. Williams. Failure Detection in an Autonomous Underwater Vehicle. *IEEE Symposium on Autonomous Underwater Vehicle Technology*, 6:377–382, 1994.
- [74] A. Papoulis. *Probability, Random Variables, and Stochastic Processes 3rd Ed.* McGraw-Hill, 1991.
- [75] D. Parish and R. Grabbe. Robust Exterior Autonomous Navigation. In *Proceedings SPIE Conference on Mobile Robots VIII*, volume 2058, pages 280–291, 1993.
- [76] R. Patton. Fault Detection and Diagnosis in Aerospace Systems Using Analytical Redundancy. *IEE Computing and Control Journal*, 2:127–136, 1991.
- [77] R. Patton and J. Chen. Observer-Based Fault Detection and Isolation: Robustness and Applications. *Control Engineering Practise*, 5:671–682, 1997.
- [78] T. Pfeuffer. Application of Model-Based Fault Detection and Diagnosis to the Quality Assurance of an Automotive Actuator. *Control Engineering Practise*, 5:703–708, 1997.

- [79] M. Polycarpou. Automated Fault Detection and Accommodation: A Learning Systems Approach. *IEEE Transactions on Systems Man and Cybernetics*, 25:1447–1458, 1995.
- [80] D. Pomerlau. RALPH: Rapidly Adapting Lateral Position Handler. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 506–511, 1995.
- [81] G. Rae and S. Dunn. On-Line Damage Detection for Autonomous Underwater Vehicles. *IEEE Symposium on Autonomous Underwater Vehicle Technology*, 6:383–392, 1994.
- [82] W. Ribbens and R. Riggins. Detection and Isolation of Plant Failures in Dynamic Systems. In *Proceedings of the American Control Conference*, pages 1514–1521, 1991.
- [83] S. Rogers. Steady-State Kalman Filter with Correlated Measurement Noise - An Analytical Solution. In *Proceedings of NAECON '89*, pages 218–221, 1989.
- [84] N. Sandell, M. Varaiya, M. Athans, and M. Safonov. Survey of Decentralised Control Methods for Large Scale Systems. *IEEE Transactions on Automatic Control*, 23:108–128, 1978.
- [85] S. Scheding, G. Dissanayake, E. M. Nebot, and H. F. Durrant-Whyte. Slip Modelling and Aided Inertial Navigation of an LHD. In *IEEE Conference on Robotics and Automation*, 1997.
- [86] S. Scheding, E. M. Nebot, M. Stevens, H. F. Durrant-Whyte, J. Roberts, P. Corke, J. Cunningham, and B. Cook. Experiments in Autonomous Underground Guidance. In *IEEE Conference on Robotics and Automation*, 1997.
- [87] G. Schenkel. Memory Management, Navigation and Map Building for a Control Architecture of an Autonomous Mobile Vehicle. Master's thesis, University of Stuttgart, 1994.
- [88] H. Sorenson and A. Stubberud. Nonlinear Filtering by Approximation of the a Posteriori Density. *International Journal of Control*, 8:33–51, 1968.
- [89] J. Speyer. Communication and Transmission Requirements for a Decentralised Linear-Quadratic-Gaussian Control Problem. *IEEE Transactions on Automatic Control*, 24:266–269, 1979.
- [90] J. Steele, C. Ganesh, and A. Kleve. Control and Scale Model Simulation of Sensor-Guided LHD Mining Machines. *IEEE Transactions on Industry Applications*, 29:1232–1238, 1993.
- [91] A. Stoerwogel. *The \mathcal{H}_∞ Control Problem*. Prentice Hall, New York, 1992.
- [92] J. Tylee. On-Line Failure Detection in Nuclear Power Plant Instrumentation. *IEEE Transactions on Automatic Control*, 28:406–415, 1983.
- [93] M. Visinsky, J. Cavallaro, and I. Walker. Expert System Framework of Fault Detection and Fault Tolerance for Robots. In *Proceedings of the Fourth International Symposium on Robotics and Manufacturing: Recent Trends in Research, Education and Applications*, pages 793–800, 1992.

- [94] M. Visinsky, J. Cavallaro, and I. Walker. A Dynamic Fault Tolerance Framework for Remote Robots. *IEEE Transactions on Robotics and Automation*, 11:477–490, 1995.
- [95] A. Willsky. A Survey of Design Methods for Failure Detection in Dynamic Systems. *Automatica*, 12:601–611, 1976.
- [96] E. Zalama, G. Mendez, J. Lopez-Coronado, and J. Peran. Mobile Robot for a Ham Industry. In *Proceedings of SPIE Conference on Mobile Robots VIII*, volume 2058, pages 224–231, 1993.
- [97] Z. Zhang. Iterative Point Matching for Registration of Free-form Curves and Surfaces. *International Journal of Computer Vision*, 13:119–152, 1994.
- [98] A. Zolghadri. An Algorithm for Real-Time Failure Detection in Kalman Filters. *IEEE Transactions on Automatic Control*, 41:1537–1539, 1996.