On the modelling requirements for the practical implementation of advanced vehicle suspension control

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: [https://dspace.lboro.ac.uk/2134/12215](https://dspace.lboro.ac.uk/2134/12215)

Publisher: © M.C. Best

Please cite the published version.
<table>
<thead>
<tr>
<th>AUTHOR/FILING TITLE</th>
<th>BEST, M</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESSION/COPY NO.</td>
<td>040121520</td>
</tr>
<tr>
<td>VOL. NO.</td>
<td></td>
</tr>
<tr>
<td>CLASS MARK</td>
<td>REFERENCE ONLY</td>
</tr>
<tr>
<td>12 DEC 1997</td>
<td></td>
</tr>
<tr>
<td>7 MAR 1999</td>
<td></td>
</tr>
<tr>
<td>26 MAR 1999</td>
<td></td>
</tr>
<tr>
<td>date due: 28 JUL 1999</td>
<td>4 APR 2000</td>
</tr>
<tr>
<td>LOAN 3 WKS UNLESS RECALLED</td>
<td>AJ 27507</td>
</tr>
</tbody>
</table>
ON THE MODELLING REQUIREMENTS FOR THE
PRACTICAL IMPLEMENTATION OF ADVANCED
VEHICLE SUSPENSION CONTROL

by

Matthew Best

A Doctoral Thesis

Submitted in partial fulfilment of the requirements
for the award of Doctor of Philosophy of the
Loughborough University of Technology

September 1995

© M.C. Best 1995
for

Rebecca Robyn
Acknowledgments

This Thesis is based on research carried out on a project funded by the FORD Motor Company. The author wishes to thank FORD of Europe, and in particular Ian Crawford and Philip Hand of the Chassis Research Group for their support of the project and for the provision of a research vehicle and testing facilities.

The research was supervised by Dr. Tim Gordon of the Department of Aeronautical and Automotive Engineering and Transport Studies (AAETS) at Loughborough University. Dr. Gordon has provided excellent supervision and valuable guidance. He has always been generous with his time, (even at 6pm on Friday afternoons) and I am extremely grateful for the knowledge he has given me, which will prove invaluable in my future career.

I would also like to thank colleagues in the AAETS department both past and present; Dr. Clive Marsh, (now a lecturer in the Department of Production Engineering, Massey University, New Zealand) was completing his research on vehicle suspension control when I began mine in 1992. I am grateful for the time he spent explaining his work to me when his time was at a premium. More recently, my office compatriots Dr. Mark Howell and Mr. Geoff Frost have been helpful in encouraging my work and providing much needed light relief.

Finally I wish to thank my wife Tessa, for her tireless support and patience during the year it has taken me to complete this work. In that time she has also carried, given birth to, and (almost) single-handedly cared for our first child, Rebecca. I owe them both a great deal of love and attention (not to mention cooking, cleaning, nappy changing, feeding ......)
Summary

This study considers the practical implementation of semi-active suspension control on a test vehicle. The aim is to assess the viability of using simple models to describe the suspension state dynamics, both for simulation purposes and to enable feedback control. To this end, system identification techniques are employed to estimate model parameters, and the design of a suitable real-time observer is considered. The controller design itself is not studied.

The well known quarter-car approach is used to develop suspension models, and a simple time-domain method is presented for parameter identification. Simulated identification tests lead to the development of a new time-domain approach, based on an integrated form of the system differential equations. This is shown to have significant advantages over the direct identification method, under certain disturbance conditions.

In a case study, the new method is applied to identify suspension parameters for the test vehicle, using data acquired during rig tests. Analysis of model errors then motivates a separate modelling exercise on the dampers in isolation. This yields a more complex nonlinear form of the model, which is finally validated against the simplest linear model, using measurements from the rig.

For real-time state estimation, a linear Kalman Filter observer is developed. The observer's sensor requirement is examined, along with other parameters affecting the design, in a factorial experiment based on simulated and rig measured data. This allows an informed choice of the smallest sensor set that affords a high level of state estimation accuracy. The performance of the observer is also examined in the context of simulated closed-loop control.

Finally, the design of an observer for the vehicle on the road is considered, and tests are carried out with the vehicle under semi-active control. Within the accuracy of state estimates derived from the available transducers, the observer performs well.
Contents

1 Introduction .................................................................................................................... 1
  1.1 Practical Application of Control .................................................................................. 2
  1.2 Active and Semi-Active Suspension Control .......................................................... 4
  1.3 Model Identification ................................................................................................. 5
  1.4 Observer Design ....................................................................................................... 7
  1.5 Test Vehicle .............................................................................................................. 9

2 A First Approach to Identifying Suspension Parameters .......................................... 12
  2.1 Quarter Vehicle Ride Model .................................................................................... 12
  2.2 A Simple Identification Strategy ............................................................................ 17
  2.3 Discussion ................................................................................................................ 21

3 Extended Simulation Study ..................................................................................... 23
  3.1 Realisable Test Conditions ....................................................................................... 23
  3.2 Simulating Measurement Noise ................................................................................ 25
  3.3 Unmodelled Dynamics ............................................................................................. 28
    3.3.1 Engine mass vibrations ....................................................................................... 29
    3.3.2 Lateral tyre slip properties ................................................................................ 30
  3.4 Simulation Conditions .............................................................................................. 32
    3.4.1 Measurement error results ................................................................................ 33
    3.4.2 Modelling error results ...................................................................................... 37

4 Identification using Randomised Integration ............................................................ 42
  4.1 Finite Interval Signal Integration .............................................................................. 42
  4.2 Statistical Analysis of Errors .................................................................................... 45
  4.3 Convergence of the RIEC Method ............................................................................ 50
  4.4 Choice of Test Parameters for RIEC Identification .............................................. 51
    4.4.1 Test duration ....................................................................................................... 54
    4.4.2 Input signal .......................................................................................................... 55
    4.4.3 Integration intervals ............................................................................................ 57
  4.5 RIEC Applied to the Extended Simulation Study .................................................. 61

5 Case Study : Identification of the Test Vehicle Suspension ..................................... 64
  5.1 The Four-poster Test Rig and Instrumentation ....................................................... 64
  5.2 Instrument Calibrations and Data Post-processing ............................................... 66
  5.3 Identification Test Procedure, Results and Discussion ......................................... 70
    5.3.1 Mass identification ............................................................................................. 70
    5.3.2 Suspension parameter identification .................................................................. 72

6 A More Detailed Model of the Controllable Damper .............................................. 77
  6.1 Damper and Control Valve Design .......................................................................... 77
  6.2 Component Tests of the Damper ............................................................................. 80
  6.3 Compliant Model for the Damper ............................................................................ 82
  6.4 Optimisation Method for Compliant Damper Model Identification ....................... 83
  6.5 Continuous Force / Velocity Model ......................................................................... 86
  6.6 Compliance Model Identification Results ............................................................... 88
7 A Comparison of Model Structures ................................................................. 91
  7.1 Identification Results .................................................................................. 91
    7.1.1 Completing the case study identification ......................................... 91
    7.1.2 Identification of the linear quarter car model .................................. 95
  7.2 Model Validations and Conclusions............................................................ 95

8 The Digital Kalman Filter Observer ................................................................. 100
  8.1 Kalman Filter Design .................................................................................. 101
  8.2 The Noise Model as a Balance Mechanism .............................................. 104
  8.3 Study of the Kalman Filter Under Non-ideal Noise Conditions ............... 107
    8.3.1 Kalman filter design using nonlinear simulations ............................ 107
    8.3.2 Testing the filter under auto-correlated and idealised error conditions 108
    8.3.3 Results and conclusions ................................................................. 109

9 Factorial Analysis of the Kalman Filter .......................................................... 112
  9.1 Physical and Simulated System Tests ......................................................... 112
  9.2 Fitted Model and Kalman Filter Observer ............................................... 115
  9.3 Factorial Experiment Based on Simulated Responses ............................. 117
  9.4 Simulation Study Results ........................................................................... 119
  9.5 Results from the Rig Data Study .............................................................. 124
  9.6 Choosing a Minimal Kalman Filter ............................................................ 127
  9.7 Discussion and Summary ........................................................................... 130

10 The Effect of Observer Error on Closed-loop Control ..................................... 131
  10.1 Applying Semi-Active Control ................................................................. 131
  10.2 Closed-loop Simulation Study ................................................................. 133
    10.2.1 Kalman filter design and test conditions ....................................... 133
    10.2.2 Simulation results ............................................................................. 134
  10.3 Assessment of Low Frequency Errors ...................................................... 138

11 Observer Design for the Moving Vehicle ...................................................... 141
  11.1 Available Data Sources ............................................................................ 141
  11.2 A Kalman Filter Design Algorithm for the Moving Vehicle ..................... 143
    11.2.1 Choosing a test road ......................................................................... 145
    11.2.2 Nonlinear simulation model ............................................................. 145
    11.2.3 Closed-loop control in simulations and vehicle tests ....................... 147
    11.2.4 Deriving the linear form of the model ............................................. 148
    11.2.5 Estimation of the output noise covariance model ............................ 148
  11.3 Road-testing the Kalman Filter ................................................................. 150

12 Discussion ..................................................................................................... 156
  12.1 Model-identification ................................................................................ 156
  12.2 State Estimation ....................................................................................... 158
  12.3 Conclusions .............................................................................................. 160
  12.4 Recommendations for Further Research ............................................... 161

References ....................................................................................................... 162

Appendix .......................................................................................................... 169
List of Figures

Figure 1 : Application of modern control.................................................................2
Figure 2 : Schematic of suspension deflection transducer orientation ......................10
Figure 3 : Configuration of instrumentation on the test vehicle.............................11
Figure 4 : Quarter vehicle model............................................................................13
Figure 5 : Impulse response of the linear quarter vehicle model.............................14
Figure 6 : Force / velocity characteristic for a passive damper unit..........................15
Figure 7 : Example of 'roof' function, R3 ................................................................16
Figure 8 : Comparison of measured and simulated vehicle responses......................26
Figure 9 : Models for (a) nonlinearity and (b) hysteresis.........................................27
Figure 10 : Quarter vehicle simulation model including engine mass.......................29
Figure 11 : McPherson strut suspension geometry.....................................................30
Figure 12 : Tyre slip model nonlinearities................................................................32
Figure 13 : Damper maps identified from data with measurement error .................33
Figure 14 : The effect of measurement nonlinearity on equation errors.....................35
Figure 15 : Explanation of hysteresis corruption in measured body acceleration, by mass parameters.................................................................36
Figure 16 : Damper maps identified from data with modelling error .......................37
Figure 17 : Comparison of identification methods over varying input bandwidth - engine mass disturbance case ..........................................................40
Figure 18 : Signal gain as a function of integration length.........................................44
Figure 19 : An Illustration of convergence of the RIEC method...............................50
Figure 20 : The influence of test duration on mass identification .............................54
Figure 21 : The effect of input bandwidth on the normalised bias norms..................56
Figure 22 : Variation of the b norm with integration interval; mass identification ..........58
Figure 23 : Variation of the b norm with integration interval; suspension identification .................................................................59
Figure 24 : Variation of fitting errors with integration interval ..................................60
Figure 25 : Damper maps identified from data with modelling errors (RIEC method) .................................................................................................63
Figure 26 : The four-poster rig and test instrumentation ...........................................65
Figure 27 : Kinematic match between body mounted sensors..................................69
Figure 28 : Case study identification results : mass parameters...............................71
Figure 57: Comparison of optimal and achieved damping under semi-active control. ................................................................. 139

Figure 58: Comparison of active systems over a simulated hill. .......................................................... 140

Figure 59: ASD plots of measured wheel acceleration, to determine the wheelhop modal frequency. ................................................................. 146

Figure 60: ASD of simulated and measured accelerations over the test road. .................................................. 149

Figure 61: Kalman filter performance in closed-loop control of the moving vehicle; innovations. ................................................................. 153

Figure 62: Kalman filter performance in closed-loop control of the moving vehicle; suspension states. ................................................................. 154

Figure 63: Kalman filter performance in closed-loop control of the moving vehicle; absolute velocities. ................................................................. 155
List of Tables

Table 1: Damper velocity / force breakpoints for a piecewise-linear model .................. 16
Table 2: Simulation results for a simple identification strategy .................................. 19
Table 3: Results of identification from data with measurement errors ....................... 33
Table 4: Results of identification from data with modelling error ............................. 37
Table 5: Summary of operating variables for suspension identification using the RIEC method ............................................................................. 61
Table 6: Parameters identified from data with modelling errors (RIEC method) ........ 63
Table 7: Mass identification results for a front and rear corner of the test vehicle ...... 72
Table 8: Identified compliance and fitting errors for the damper component test ......... 88
Table 9: Stiffness and fitting error results for three identification approaches .......... 92
Table 10: Average spline damper maps identified under a fixed compliance ............. 94
Table 11: Parameters identified for the linear model .................................................. 95
Table 12: Linear and nonlinear model validation results ............................................ 96
Table 13: Performance comparison of three Kalman filter designs tested under conditions of auto-correlated system noise .......................................... 111
Table 14: Factors chosen for the statistical experiment ............................................. 118
Table 15: Marginal increase in percentage fit from additional sensors; rig test data .......................................................... 127
Table 16: Summary of factors for the minimal Kalman filter .................................... 128
Table 17: RMS performance of simulated closed-loop control systems ................... 134
Table 18: Kinematic performance of the Kalman filter in closed-loop control ........... 137
Table 19: Advantages and disadvantages of three sources of information for design and testing of a Kalman filter for the moving vehicle ......................... 142
Table 20: State reconstruction from the suspension deflection measurement \( ds \), and accelerations \( aw \) and \( ab \) ........................................................................... 151
Chapter 1

Introduction

Recent advances in computational performance, at low cost, have made more and more complex computer controlled systems viable in the mass produced motor car. Engine management systems, anti-lock braking and traction control have all become widely used, and the scope for applications is widening to encompass systems which control the suspension, driveline and engine mounts for example.

This thesis considers the practical application of a suspension control system. The work is not focused on control strategy though, but rather on the modelling and estimation tools that are required to supply information to the controller regarding the dynamic state of the suspension. This area is vitally important in controller design, as the performance of real-time control systems is constrained by the accuracy with which this information can be supplied.

The study is carried out in the context of designing a suspension control system for a prototype vehicle. This car, a FORD Granada, is essentially standard except that it has been equipped with continuously variable damper units, and instrumented with road-going sensors for the application of semi-active control. The hardware configuration for the test car is described in Section 1.5.

Using simulation studies, rig and road based experiments on the test vehicle, and related mathematical analysis, the aim of this work is

1) to quantify the potential for using simple model structures in the practical development of a vehicle suspension control system.

2) to examine the practical consequences of various theoretical limitations in simple identification and observer design techniques.

In this chapter, the reader is first introduced to the main components in the application of model-based control, and the function of an appropriate system model is illustrated. A brief survey is then given of established techniques for suspension control. Methods for model identification are next discussed, and previous studies in the identification of automotive models are investigated in order to establish the state of the art; these studies
motivate the approach that will be used in this work. The literature on, and motivation for observer design is analysed in a similar way. The test vehicle is then described and details are given of its electrical and computer hardware.

1.1 Practical Application of Control

The application of modern automatic control requires a common chain of processes. The principal components of plant, observer and controller are illustrated in Figure 1.

![Figure 1: Application of modern control](image)

Known system inputs $u$, controls $u^*$, and unknown disturbances $\varepsilon_u$ excite the plant, whose dynamics can be described by state variables $x$. A model of these dynamics is obtained, based on parameters $\theta$ and subject to errors $\varepsilon_m$. Measurements $y_s$ are taken which include errors $\varepsilon_s$ and which typically do not include all the states. The observer system uses the known inputs, measurements and model to reconstruct an estimate of the states, $\hat{x}$. The control action is applied on the basis of these state estimates, to regulate the system, usually with regard to some cost function of undesirable dynamics.
For the vehicle ride control systems considered in this study, the known inputs are usually restricted to suspension control actions; steering and braking inputs may also be known if vehicle handling control were under consideration. The disturbing influence of the road is usually assumed to be unknown and is thus the main component of $e_u$.

It is clear from the figure that the system model has a central role in determining control actions. The model must provide a sufficiently accurate representation of the system, capable of predicting state trajectories, and yet must take a simple form for on-line implementation. In the observer, a linear form is most desirable.

Many researchers restrict attention to a quarter vehicle model for the vertical suspension dynamics, for example in the controller and observer design by Raju and Narayanan [34]. This model is the simplest that describes the main ride effects, of isolation of the body mass from the road, and modelling of the tyre contact forces. It has a multi-body structure, such that the vehicle is assumed to be made up of rigid sprung and unsprung masses separated from each other and the road by springs and dampers; the model parameters can thus be physically interpreted.

To identify suitable values for the physical parameters, component tests can be carried out, but these tend to be inaccurate for describing the systems overall dynamics. For example, the total suspension compliance will be due to bushes and component flexure in addition to the main spring. Also, the division between sprung and unsprung masses is difficult to quantify, given the shared mass of the spring, damper and supporting struts.

For these reasons, system identification methods are attractive, whereby the parameters are estimated according to measured dynamics of the whole system under excitation. Identification of physically interpretable parameters has the additional benefit that the analysis of errors can lead to improvements in the model by suggesting extensions to its structure. The improvement in accuracy that this brings can be monitored in relation to the increased complexity of the new model structure.

The formulation of the quarter vehicle model (in a controllable and observable form) typically requires a definition of states describing tyre deflection, suspension deflection and absolute vertical velocity of the wheel hub and vehicle body. However, in practice only one of these is directly measurable by sensors. The observer therefore plays a vital role in determining state estimates from the measurement set which is typically made up of accelerometers, and suspension velocity and displacement transducers.
1.2 Active and Semi-Active Suspension Control

The suspension in any road vehicle is designed to achieve two principal objectives, occupant comfort (ride comfort) and road holding (handling capability), within one main constraint, suspension workspace. With a passive system the two objectives directly conflict with each other; higher spring and damper rates give a stiffer 'sports car' suspension, offering good roadholding, but with poor isolation of the occupant from the road roughness. Softer suspensions however, typical of American luxury sedans, give a smoother ride with poor roadholding.

Automatic suspension control promises an improvement in the trade-off such that disproportionate gains in ride comfort might be had from small concessions in handling or vice versa. Two schemes predominate. Active systems rely on the inclusion of a force generator (typically hydraulically powered) replacing or augmenting the passive suspension components. These actuators typically have a low bandwidth and are therefore restricted to controlling body motions principally; significant reductions in vehicle body oscillations have been claimed in simulation, and achieved using concept cars, for example in Williams and Best [47]. However, active systems tend to be expensive, both in terms of component costs and power usage.

Semi-active systems retain the passive spring elements of the suspension but employ switchable rate or continuously variable dampers. These are typically similar in design to passive hydraulic dampers, with some mechanism for varying valve orifice size, although more recently, electro-rheological (variable viscosity) fluids have been investigated, for example by Petek et al [32]. Semi-active systems are generally cheaper to produce and require little energy to run, but they are only capable of varying the energy dissipated in the system, and this reduces performance. Nevertheless, simulation studies have been presented, e.g. by Margolis [28] that claim performance from semi-active systems approaching that of equivalent active variants.

Since the 1960s, and particularly more recently, many papers have been published on control algorithms for semi-active and active suspensions. An extensive study of the literature is not conducted here however, as the control algorithm plays a minor role in this study. A good review of the design issues for suspension system control can be found in Sharp and Crolla [39].

Considering control strategies in the context of observer design, the important issue concerns the state variables that prospective controllers employ. These range from the use of absolute body velocity alone, in the well known skyhook law proposed by Karnopp et al [24] to use of all four states of the quarter vehicle model in many feedback laws, such as in
Gordon and Best [12]. For semi-active control, the suspension deflection and velocity are also generally required, to convert the feedback law into a control signal for the damper. With the possible exception of tyre deflection, the observer is thus required to estimate all four states defined by the quarter vehicle model.

1.3 Model Identification

System identification is a well known technique in control systems theory, and several methods are described in texts such as Eykhoff [7]. Here we review those that have been applied to automotive suspension and handling systems, and the similar field of rail vehicle suspension.

Almost universally, the methods rely on estimations from data using a least-squares criterion. This has the disadvantage that bias may be induced in the parameter estimate if measurement errors in the data are correlated with the identified dynamics. Also, although it is a secondary effect, bias is induced by any errors in the parameter regressors. The published identification methods can generally be divided into two categories; those which ignore the potential for error, and use a direct least squares approach, and those which modify the data or model to avoid correlation errors. (The secondary errors are usually impossible to avoid.)

One of the simplest approaches is that suggested by Lin and Korttim [26]. The unknowns are identified from an ordinary least squares solution of the equation of motion for the known model structure, which must be linear-in-the-parameters, applied to measured system dynamics. Michelberger et al [29] applies the same technique to identify a z-transform model for suspension components, and this paper also considers a time-domain maximum likelihood method, to estimate parameters in a finite-element model of the vehicle body.

Most studies consider off-line estimation algorithms, but Würtenberger et al [48] and Zhang and Chen [52] present recursive least-squares strategies to identify parameters for handling models which operate on the vehicle. Würtenberger et al transform measured signals into a form suitable for identification using a state variable filter, which is described in a survey of parameter estimation techniques by Young [51]. Zhang and Chen use a block pulse function to approximate the states by integrating measured state derivatives. Both achieve recursive parameter estimates through an observer system similar in structure to a Kalman filter, with slow poles.

Returning to off-line estimation techniques, a popular adaptation of the least squares approach is the use of instrumental variables, described by Söderström and Stoica in [42].
This method has been applied to the identification of nominal stiffness and damping characteristics for the simulated suspension of a magnetic levitation vehicle, by Roether and Müller [38]. The technique employs a transformation of state variables into a form uncorrelated with measurement errors, and Roether and Müller claim a significant improvement in accuracy over the direct least squares approach.

An approach which the author claims is equivalent to an extension of the instrumental variables method, in the frequency domain, is proposed by Kallenbach [21]. Here a linear filter is subjected to the measured signals and the identification is carried out based on the criterion of a stationary cross-covariance between the filter and measured system states. In his paper Kallenbach also includes a summary of identification methods that may be useful in the study of suspension systems.

Other frequency domain approaches range from a simple FFT-based identification of the modal characteristics of a suspension system, e.g. by Shibo et al [41] and Kropáč and Sprinc [25], to detailed analysis of the system's Frequency Response Function (FRF) by Fries and Cooperrider [8]. Here the FRF, described by Bendat and Piersol [3], is shown to offer better rejection of measurement errors than simple power spectral density functions, in the identification of railway suspension parameters.

Finally, an interesting, though perhaps impractical frequency domain method is presented by Hemingway [17], who identifies the immittance matrix between attachment points on individual suspension components, using excitation frequencies up to 250Hz. The expected component interactions are then used to identify system modal masses and from these, stiffness and damping values.

For the quarter car identification in this study the results can be validated on a test vehicle, whereas for most of the studies described above only simulation results are presented, for example [26, 38, 52]. It is therefore possible to investigate the dynamic consequence of parameter bias that may be induced by simple identification techniques, rather than simply opting for a method which theoretically avoids potential problems.

To this end, a direct least-squares identification is first investigated, and this is tested in a simulated quarter car identification in Chapter 2. The model structure adopted in this study takes a simple form, but the principal nonlinear element, the damper characteristic, is described using a piecewise linear-in-the-parameters approximation to a measured damper map.

Excellent results are achieved for the nominal disturbance noise conditions considered, and the study motivates a more formal investigation, in Chapter 3, using simulations of modelling and measurement errors that are more likely, given assumptions about the likely
environment for system identification tests. These simulations reveal severe bias errors under certain conditions however, which motivate a modification to the least-squares approach. A new method is formulated in Chapter 4.

Chapter 5 is devoted to a case study identification of the test vehicle, which serves to confirm earlier findings of the simulation study, but also to cast doubt on the adequacy of the model structure, which has remained fixed up to this point.

A detailed component study is conducted on the damper in Chapter 6, and this leads to development of a highly nonlinear damper model. Although improvements in model accuracy are promised, the added model complexity motivates a comparison of model structures, in Chapter 7; the simplest linear model structure is identified and validated against the new model. Good performance of the linear variant suggests that linear observer designs may be practical.

1.4 Observer Design

Until the 1990s, very little had been published on suspension system observers. Much of the work since then has revolved around the Kalman filter, described by Kalman [22] and Kalman and Bucy [23]. This provides state estimates under stationary or non-stationary conditions which are optimal under specific errors conditions; the system errors \( (\epsilon_u + \epsilon_m) \) and output errors \( \epsilon_y \) are assumed to have zero auto-correlation, zero cross-correlation, and known covariance.

Under most practical conditions however, autocorrelated error conditions exist, largely from the road input disturbance which is dominated by low frequencies, as noted by Robson [37] in his study of measured road profiles.

Many studies have been presented which accept the white noise assumptions in the design and simulated testing of observers. For example, Yoshimura and Ananthanarayana [50] who present an active suspension with preview of the road ahead, and Ray [35] who uses an extended Kalman filter to estimate tyre forces in a nine degree-of-freedom vehicle handling model. Raju and Narayanan [34] assume ideal noise conditions in the Kalman filter design, although the test condition is auto-correlated, with a road profile defined by a second order filter response to white noise. Ulsoy et al [45] investigated a combined LQG observer and controller for an active suspension system, and they considered robustness to changes in cost function and sensor set, though not to correlated noise conditions.

An early attempt at extending the Kalman filter to a robust form was made by Stepinski [43] who used a Bayesian approach to minimise maximum error variances where the p.d.f. of measurement noise is known to be a member of a certain class of functions. More
recently attention has been focused on the specific problem of road profile estimation as a method of whitening the noise processes. Jeong et al [20] utilised a linear shaping filter with white noise input as a model for the road, in an extended full vehicle Kalman filter. Parameters in the filter are identified by matching output measurements with model expectations, and a Pade filter is used to model the time delay between front and rear wheels. Venhovens [46] also suggests a white noise filter for the road and additionally employs an integrated white noise model for measurement noise.

Several alternatives to the Kalman filter approach have been proposed; these include the equivalent, deterministic Luenberger-type observer, used by Alleyne and Hedrick [2], a classical pole-placement approach (Mahajan and Krishnan, [27]) and a nonlinear sliding mode design suggested for semi-active control, by Henry and Zeid [18].

The bilinear representation of semi-active systems has attracted a number of disturbance decoupled bilinear observers. Haç [14] extended earlier work by Hara and Furuta [15] to decouple state estimation errors from \( \epsilon_m \). Hedrick et al [16] claim additionally to decouple the errors from \( \epsilon_u \), although this will only be true if the condition of zero system / output cross-correlation errors is met.

Among these publications many innovative solutions are presented. However, in all cases except two, [16] and [46], the studies are conducted in simulation only. Given the aim of this study and availability of a test vehicle, we attempt here to expose the practical implications of a simple observer, so that an informed judgement can be made on the relevance of, say, non-standard error conditions. Other important practical considerations are made, of sensor requirements and the effect of errors on closed-loop control of the test vehicle.

The observer study commences in Chapter 8 with a description of a simple Kalman filter, designed using linear models and assuming stationary noise conditions. One important modification to the design is considered at the outset however. For observers using accelerometers or other force measurements, the system and output errors are significantly correlated. This issue is ignored in most publications [e.g. 34, 43, 45, 50], yet compensation for the correlation is easily made; a model transformation which is published in Gelb [10] and attributed to Y.C. Ho enables cross-correlation covariance estimates to be included within the well known Kalman filter design. In Chapter 8 the correlated nature of errors is demonstrated using a simple mechanical model, and various error conditions and their effect on Kalman filter designs are investigated.

Having decided on a suitable Kalman filter design structure, Chapter 9 deals with the important issue of sensor requirements. The viability of any practical suspension
controller is equally dependent on the cost of implementation as on its performance, and the sensors represent a large component of total cost. In this chapter, a factorial experiment is conducted, based on rig measurements of the test vehicle and also simulated data, to ascertain the performance benefit of combinations of four possible sensors in conjunction with other factors such as model order. The resulting observer, based on minimal sensor requirements, then forms the basis of further study. Chapter 9 is based on a study by Best and Gordon [4].

Chapter 10 considers simulations of closed-loop control using the filter, and comparisons between semi-active and active systems are made using full state feedback and Kalman filter estimates. For the semi-active observer based system, the closed-loop performance is found to degrade only slightly, despite low frequency errors in the state estimates. This chapter includes some results presented earlier, in Best and Gordon [5].

The study is completed by designing an observer for the test vehicle on the road. This is complicated by the difficulty of identifying a suitable noise model, and obtaining accurate state information without additional sensors. Nevertheless, within the limitations of available data, positive results are achieved for the prototype semi-active system.

1.5 Test Vehicle

The test vehicle for this study is a 2.9 litre, V6 FORD Granada. The car is essentially standard, except that it is fitted with continuously variable dampers, instrumented with sensors at each wheelstation, with electronic control equipment in the boot. The configuration of this equipment is summarised in Figure 3.

The sensor set at each corner of the car comprises two piezo-resistive type accelerometers, and a single-turn rotary potentiometer which is mounted to record suspension deflection, as illustrated in Figure 2. The nonlinear output from the potentiometer is linearised using a mapping to vertical deflection measured on a test rig (see Chapter 5). The accelerometers are rigidly mounted in the vertical plane at each wheelhub, and on the vehicle body at the top pinchbolt of each damper; they have a linear range of $\pm 150\text{m/s}^2$ and $\pm 20\text{m/s}^2$ respectively.

The signal conditioning module amplifies each of the transducer signals, and applies an analogue anti-aliasing filter. This Bessel filter induces a constant time delay of 500 $\mu$secs at pass frequencies up to around 300Hz, suitable for the sampling interval of 500Hz that is used. A multiplexor module is also employed, in order to reduce the signal conversion hardware requirements to a single A/D and a single D/A convertor.
Signal processing is carried out using a PC coupled with a Digital Signal Processor (DSP). The DSP, a TMS320c30, is dedicated to the execution of state estimation and control algorithms. It operates using software which is compiled and downloaded from the PC, a 486 DX2 50MHz machine. The advantage of operating two processors is that data can be acquired and analysed in real-time. As the PC and DSP access shared dynamic memory, the PC is used at run-time to acquire data, both for off-line analysis and to generate real-time graphical displays on a remote screen mounted inside the car.

Completing the control system hardware, the control action for each damper is applied via a damper drive module. This converts control voltage signals into pulse-width modulated square-wave current signals supplying the solenoid valve in each damper. The pulse width modulates the damping rate, and the oscillating drive signal continuously vibrates the solenoid valve; this prevents sticking and improves the dynamic response of the actuator.
Figure 3: Configuration of instrumentation on the test vehicle
Chapter 2

A First Approach to Identifying Suspension Parameters

In this Chapter a mathematical model is presented which can be used to predict and/or simulate vehicle ride dynamics. The quarter vehicle model is a widely used representation of a vehicle's suspension system and it has a simple mathematical structure which is based on physically meaningful parameters. As such, it is an ideal basis for system identification.

Normal modes are used to illustrate the model, although it turns out that to accurately represent damping characteristics, a nonlinear damper map must be introduced.

A simple identification strategy is then proposed, for estimating model parameters from measured time-histories of suspension activity. The aim is to choose parameters which provide the 'best fit' of the model to this data. A simulation study is carried out to investigate the efficacy of the approach under varying levels of noise corruption in the measurements.

2.1 Quarter Vehicle Ride Model

The quarter vehicle model is based on the principal mechanical components of the suspension. The model comprises sprung and unsprung masses connected by a spring and damper assembly, with the unsprung mass isolated from the road through a linear spring modelling the tyre (Figure 4).

The system has two degrees of freedom, describing vertical motion in terms of four state variables:

\begin{align*}
x_1 & \quad \text{Tyre deflection} \\
x_2 & \quad \text{Suspension deflection} \\
x_3 & \quad \text{Absolute vertical velocity of the wheel hub} \\
x_4 & \quad \text{Absolute vertical velocity of the body}
\end{align*}
These states are dynamic variables and are therefore considered in terms of their deviation from a rest state. The only input to the system is the road roughness, described in terms of vertical road velocity $v_r$. Kinematic state relationships can be seen by inspection:

$$\dot{x}_1 = v_r - x_3 \quad /2.1.1/$$

$$\dot{x}_2 = x_3 - x_4 \quad /2.1.2/$$

The dynamic relationships come from consideration of forces acting in the system from the tyre spring and the suspension spring and damper:

$$F_t = K_t x_1 \quad /2.1.3/$$

$$F_s = K_s x_2 + B_s (x_3 - x_4) \quad /2.1.4/$$

and from Newton’s second law, the dynamic state relationships are given by

$$\dot{x}_3 = \{ K_s x_1 - K_s x_2 - B_s (x_3 - x_4) \} / M_w \quad /2.1.5/$$

$$\dot{x}_4 = \{ K_s x_2 + B_s (x_3 - x_4) \} / M_b \quad /2.1.6/$$
The system can be described by two modes, commonly referred to as the *wheel-hop* and *body bounce* modes. Writing the above equations in the matrix form

\[ \dot{x} = Ax + Bv, \]

and taking the linear mass, stiffness and damping constants from Figure 4, these can be characterised by the eigenvalues of \( A \),

\[ \lambda_w = -33.8 \pm 67.4i \]
\[ \lambda_b = -2.36 \pm 6.93i \]

From complex constants \( p \pm qi \), the undamped natural frequency and 98% settling time for each mode is then given by:

\[
\begin{align*}
\omega_n &= \sqrt{p^2 + q^2} / 2\pi \\
T_s &= 4/p 
\end{align*}
\]

<table>
<thead>
<tr>
<th>Mode</th>
<th>( \omega_n ) (Hz)</th>
<th>( T_s ) (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheelhop</td>
<td>12.0</td>
<td>0.12</td>
</tr>
<tr>
<td>Body bounce</td>
<td>1.17</td>
<td>1.7</td>
</tr>
</tbody>
</table>

*Figure 5: Impulse response of the linear quarter vehicle model*
Figure 5 illustrates the system modes, showing time histories for the response of wheel hub and body velocities ($x_3$ and $x_4$) to an impulsive road velocity input, and their PSD response to white noise.

The upper plot illustrates modal settling times, and shows how the energy from an isolated road disturbance is absorbed initially through oscillations of the wheel mass. These in turn excite slower and less extreme oscillations in the vehicle body. The lower plot shows that the body motion is entirely dominated by the body bounce mode, whereas wheel motion is excited in both modes. However, the wheel oscillates across a broader bandwidth and it's energy is spent principally at the higher frequencies. The width of the wheelhop characteristic frequency and the position of its peak is a function of damper rate.

As we wish to identify a model for semi-active control, we are particularly interested in an accurate model for the force / velocity characteristic of the suspension damper. Figure 6 shows measurements of force and velocity taken from a test carried out on a passive damper (details of this test are given in Chapter 6); the relationship is clearly nonlinear. Here the damper expansion velocity is given by

$$V = x_3 - x_4$$

![Figure 6: Force / velocity characteristic for a passive damper unit](image-url)
To improve upon the linear representation of damping in the quarter car model, a simple five segment piecewise linear map is used (as illustrated). The damping is then described in terms of linear functions, based on a knowledge of damper force at five pre-defined damper velocities (Table 1).

<table>
<thead>
<tr>
<th>i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_i$ (m/s)</td>
<td>-0.8</td>
<td>-0.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0.8</td>
</tr>
<tr>
<td>$F_i$ (N)</td>
<td>-800</td>
<td>-300</td>
<td>0</td>
<td>1500</td>
<td>2300</td>
</tr>
</tbody>
</table>

*Table 1: Damper velocity/force breakpoints for a piecewise-linear model*

Damper force $F_d$ can be written in terms of damper velocity $V$ using 'roof' functions $R_i(V)$

$$F_d(V) = \sum_{i=1}^{5} R_i(V) F_i$$  \hspace{1cm}  (2.1.7)

where $R_i$ also depends on the choice of nodal damper velocities. Figure 7 illustrates $R_3$, and the remaining $R_i$ are defined in a similar fashion, centred on node $i$.

*Figure 7: Example of 'roof' function, $R_3$*

The simulation model now uses nonlinear characteristics and cannot strictly be described in terms of modes. However, the impulse response is changed only slightly, and the system retains similar 'resonance' frequencies for the wheel and body masses.
2.2 A Simple Identification Strategy

Assuming the chosen model structure is sufficient to represent the ride dynamics of a test vehicle, we require a method for identifying values for $M_b, M_w, K, K_t$, and the nodal damper forces $F_i, i=1,5$ which best describe these dynamics. Note that the object is not to find masses and spring rates etc. for individual components, but to identify parameters in a global sense, that characterise the whole suspension. The resulting model will be imperfect, as are all models, but it will represent a compromise between accuracy and mathematical complexity that is in some sense optimal.

If a set of measurements can be made from the vehicle while the suspension is excited, these can be applied to the model structure to choose the parameters. This can be done in a number of ways, either directly, using the recorded time histories such as in [26], or after transformation of the data into the frequency domain, as in [8]. Here, we use equations 2.1.3 to 2.1.7 to identify directly from the data. Direct substitution yields an overdetermined set of equations which can be written in the form:

$$y = U\theta + e$$ \hspace{1cm} /2.2.1/

where $U$ is an $N \times p$ regression matrix constructed from sampled data, the regressand $y$ is a similar known vector of length $N$ and $\theta$ is the vector of $p$ unknown parameters. For a given $\theta$, the vector $e$ describes errors which indicate the accuracy with which the identified model describes the data; these are attributable to two factors: model simplicity and measurement error.

We seek a 'best fit' approximation of $\theta$, and the simplest approach is to use the well known method of ordinary least squares (OLS) (see for example [30]). An estimate $\hat{\theta}$ is obtained which minimises the error expression

$$E = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N} (y_k - u_k^T\hat{\theta})^2$$

where $u_k^T$ represents the $k_{th}$ row of $U$. The OLS solution is given by

$$\hat{\theta} = (U^TU)^{-1}U^T y$$ \hspace{1cm} /2.2.2/
To assess the success of this estimate of parameters, we define the \textit{percentage fitting error}, $P$ via the ratio of error norm to the norm of the regressand:

$$P = \left\{ \frac{\sum_{k=1}^{N} e_k^2}{\sum_{k=1}^{N} y_k^2} \right\} \times 100$$

To identify nodal damper forces using this technique, measurements of damper force and velocity are used. From equations 2.1.7 and 2.2.1 we have

$$U = \begin{bmatrix} R_1(1) & R_2(1) & \ldots & R_5(1) \\ R_1(2) & R_2(2) & \ldots & R_5(2) \\ R_1(3) & R_2(3) & \ldots & R_5(3) \\ \vdots & \vdots & \ddots & \vdots \\ R_1(N) & R_2(N) & \ldots & R_5(N) \end{bmatrix}$$

$$y = \begin{bmatrix} F_d(1), F_d(2), F_d(3), \ldots, F_d(N) \end{bmatrix}^T$$

Assuming measurements are also available for $F_s, F_r, x_1, x_2, \dot{x}_3$ and $\ddot{x}_4$, the remaining parameters can each be identified separately, by appropriate rearrangement of 2.1.3 - 2.1.6

$$F_s = \dot{x}_4 M_b$$

$$F_r - F_s = \ddot{x}_3 M_w$$

$$F_t = x_1 K_t$$

$$F_s - F_d = x_2 K_s$$

A simple simulation experiment illustrates the potential of this approach. The required ‘test’ data were simulated by integration of the continuous-time state equations 2.1.1-2.1.7.\textsuperscript{†}

\textsuperscript{†}. For this and other continuous-time nonlinear simulations considered in later chapters, a numerical integration is carried out, using a 5th order Runge-Kutta algorithm, with variable step size. The computer code has been adapted from standard routines described in [33]. The input signal is defined at discrete time intervals $t_d$, which are linearly interpolated. Outputs are then given as vector time histories using the same sampling interval; the simulation thus provides a faithful representation of input and output signals up to the Nyquist frequency $F_N$:

$$F_N = \frac{1}{2t_d}$$

For this simulation a Nyquist of 50Hz was considered adequate, i.e. $t_d = 0.01$ seconds. Choosing a simulation run-time of 20 seconds, a data set of 2000 points is generated.
To simulate the road input, a Gaussian white noise signal was used,

$$v_r \sim N(0, \sigma^2)$$

This is a simple approximation to vertical velocity which has been used in a number of studies, for example [9]. The amplitude of $v_r$ was chosen to excite the suspension at sufficient amplitude to enable identification of all the model parameters. For the nonlinear damper map, this means inducing suspension velocities which exceed $\pm 0.3 \text{m/s}$, and for adequate values in the regressors $R_1$ and $R_3$, peak suspension velocities of $\pm 0.8 \text{m/s}$ are desirable. For the given choice of sampling interval, an amplitude $\sigma = 0.5 \text{m/s}$ was found to be suitable.

To represent errors in the simulated measurements, Gaussian white noise was also added to each of the signals. The RMS amplitude of this noise is set in proportion to the RMS amplitude of the signal, and several ratios are examined.

Results are given in Table 2. To reduce scatter, each parameter estimate and fitting error is a mean value from 20 independent tests.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>5%</th>
<th>10%</th>
<th>20%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$ (N)</td>
<td>-800</td>
<td>-800</td>
<td>-793</td>
<td>-784</td>
<td>-685</td>
</tr>
<tr>
<td>$F_2$ (N)</td>
<td>-300</td>
<td>-301</td>
<td>-305</td>
<td>-309</td>
<td>-252</td>
</tr>
<tr>
<td>$F_3$ (N)</td>
<td>0</td>
<td>6</td>
<td>24</td>
<td>85</td>
<td>279</td>
</tr>
<tr>
<td>$F_4$ (N)</td>
<td>1500</td>
<td>1490</td>
<td>1458</td>
<td>1358</td>
<td>1008</td>
</tr>
<tr>
<td>$F_5$ (N)</td>
<td>2300</td>
<td>2300</td>
<td>2314</td>
<td>2324</td>
<td>2146</td>
</tr>
<tr>
<td>$K_r$ (Kg)</td>
<td>20.0</td>
<td>20.0</td>
<td>19.8</td>
<td>19.3</td>
<td>16.1</td>
</tr>
<tr>
<td>$M_b$ (Kg)</td>
<td>350</td>
<td>349</td>
<td>346</td>
<td>336</td>
<td>280</td>
</tr>
<tr>
<td>$M_w$ (Kg)</td>
<td>30</td>
<td>29.9</td>
<td>29.7</td>
<td>28.9</td>
<td>23.8</td>
</tr>
<tr>
<td>$K_i$ (kN)</td>
<td>160</td>
<td>160</td>
<td>158</td>
<td>154</td>
<td>128</td>
</tr>
</tbody>
</table>

| Fitting error | | | | |
|---------------| | | | |
| P($F_1$)      | 8.1 | 15.9 | 30.3 | 61.1 |
| P($K_r$)      | 12.2| 24.0 | 44.9 | 80.3 |
| P($M_b$)      | 7.0 | 14.1 | 27.5 | 60.1 |
| P($M_w$)      | 12.1| 23.8 | 44.1 | 80.1 |
| P($K_i$)      | 7.0 | 14.0 | 27.5 | 59.8 |

*Table 2: Simulation results for a simple identification strategy*
For low levels of noise, the identification is clearly successful. However, as the noise / signal ratio increases, a distinct negative bias is seen in the parameters, even though the noise is not correlated with any of the signals.

The effect is well known and is due to the presence of correlation between the regressors and the equation error, $\epsilon$. It turns out that some bias is inevitable if the regressors include errors. To see this, consider the identification of a single parameter $\theta$ where the exact relationship

$$\tilde{y} = \tilde{u} \theta$$

exists, and an estimate $\hat{\theta}$ is to be made from measurements which include errors $v$ and $w$

$$y = \tilde{y} + v$$
$$u = \tilde{u} + w$$

We can re-write equation 2.2.3 in terms of the measurements

$$y = u \theta + e$$

where

$$e = v - w \theta$$

The optimal OLS estimate is calculated as

$$\hat{\theta} = (u^T u)^{-1} u^T y$$

which can be re-written using 2.2.5 as

$$\hat{\theta} = \theta + (u^T u)^{-1} u^T e$$

Clearly, if no bias is to occur, the expectation of the term $u^T e$ must be zero. However, equations 2.2.4 and 2.2.6 show that any error in the regressors will result in some correlation between $u$ and $e$ and hence some bias in the estimated parameters. Writing equation 2.2.7 in terms of $\tilde{u}$, and noting that $v$ is uncorrelated with $w$ and $\tilde{u}$;

$$\hat{\theta} = \theta - [ (\tilde{u} + w)^T (\tilde{u} + w) ]^{-1} (\tilde{u} + w)^T w \theta$$

The extent of the bias depends on the relative magnitude of $w$ compared with $\tilde{u}$, along with any correlations that may exist between them. In the simulation study, $w$ is chosen to be uncorrelated with $u$, but related in magnitude such that

$$w_i \sim N(0, \sigma^2), \quad \sigma = \lambda \text{RMS} (\tilde{u})$$
and in these circumstances, the estimate of $\theta$, can be deduced from 2.2.8 as

$$\hat{\theta} = \theta \left( 1 - \frac{\lambda^2}{1 + \lambda^2} \right)$$

The relationships 2.2.8 and 2.2.9 are approximate due to the presence of stochastic terms. More strictly, we should write

$$\text{plim} (\hat{\theta}) = \theta \left( 1 - \frac{\lambda^2}{1 + \lambda^2} \right)$$

where the probability limit $\text{plim}$ of $\hat{\theta}$ is the convergent parameter estimate as the size of the measurement vector increases. For 10% error, $\lambda = 0.1$ and we obtain $\text{plim} (\hat{\theta}) / \theta = 0.99$, whereas for 50% error, $\text{plim} (\hat{\theta}) / \theta = 0.8$.

In a similar way, it is possible to derive an expression for fitting errors in the single parameter case:

$$\text{plim} (P) = \lambda \sqrt{1 + \left( 1 - \frac{\lambda^2}{1 + \lambda^2} \right)}$$

As measurement errors increase, the fitting error does not increase proportionately; OLS seeks to minimise fitting errors, and this is partly achieved by the observed parameter bias.

The single parameter identification results in Table 2 are consistent with equations 2.2.10 and 2.2.11. Note that the increased fitting errors associated with $K_s$ and $M_w$ are to be expected, as noise has been applied independently to $F_d$, $F_s$, and $F_n$, and the total noise will be cumulative in the regressand for these identifications. No increase in bias results from this though, as the additional error does not appear in the regressor.

### 2.3 Discussion

The study presented above provides only an elementary first look at time-domain identification using the nonlinear quarter car model. One of the problems associated with identification from experimental data has been highlighted and explained analytically, but this is not always possible under realistic conditions. In this study we have ignored or over-simplified factors which may have a severe effect on the scope and robustness of direct identification in practice. It would therefore be inappropriate to simply apply OLS in a 'black box' sense without a more thorough investigation of potential pitfalls.
Taking more careful consideration of the issues involved in testing a real vehicle, a more detailed simulation study is entered into in the next chapter. Particular attention will be given to three main issues:

(1) Likely test conditions and availability of measured signals.

(2) Sources of errors in measured signals.

(3) Dynamic modes that are not represented by the quarter car model.
Chapter 3

Extended Simulation Study

In this chapter, the direct least squares identification technique is tested under more realistic assumptions. By considering likely conditions for identification tests on a vehicle, significant sources of measurement and modelling error are isolated. The effect of these errors on parameter identification is then examined using simulation models. The models are nominal, as it is not possible to anticipate precise magnitudes and origins of error; the aim is to investigate trends and examine the technique’s resistance to broad effects.

3.1 Realisable Test Conditions

The first consideration in identifying parameters from a test vehicle is the experimental environment. Can adequate measurements be made on the road, and if not, is a test rig available that can reasonably approximate road conditions? The answers will vary for different research teams, but the experience here seems typical.

Unfortunately, a number of essential variables are impractical to measure from a moving vehicle: no sensors were available for the direct measurement of dynamic tyre deflection and vertical road velocity. (These would require very expensive sonar or laser tracking equipment, and internal tyre pressure or road surface force transducers.) Test rigs were available, that can emulate road inputs to the car, through hydraulic actuators, but none of the available test equipment accommodates rolling wheels; the identified tyre parameters would thus represent an unreliable model of tyre dynamics on the road.

Accepting the implications for tyre modelling, test rigs offer the advantage of low-noise measurements taken in a controlled environment. The most suitable rig comprises four independent hydraulic jacks, fitted with flat plates (wheelpans) on which the car stands. The hydraulic system is typically capable of providing an input bandwidth of at least 50Hz, with a maximum stroke of perhaps 200mm. The latter restriction has implications for the road inputs that can be simulated.

It seems sensible to choose an input that is representative of a true road surface; in fact measured road profiles were available, which could be used as drive signals for the rig. To provide varied test conditions though, it is probably more convenient to use a white noise process, shaped according to a frequency model. One such model suggested by Robson
[37], is widely used. The vertical displacement power spectral density, $S$, of the surface is given by

$$S(f) = kU^{(w-1)}f^{-w}$$

where choosing $w = 2.5$ provides an adequate model in the frequency range of interest. Frequency $f$ is in Hz, $U$ is the forward speed of the car in m/s, and $k$ is a roughness coefficient; this has been estimated by Robson at between $3 \times 10^{-8}$ and $3 \times 10^{-5}$ depending on the type of road. Choosing $k$ for the roughest minor road, at $3 \times 10^{-5}$, forward speed is found to induce the desired suspension velocity peaks (discussed in Chapter 2) at around 30 m/s. To keep the input within assumed stroke limits of the rig, all frequencies below 0.4 Hz are removed via FFT.

Although most of the required dynamic variables can easily be recorded on a test rig, some consideration must be given to the interpretation of force variables. The advantage of identification to a simple model structure is that the identified parameters describe the suspension as a whole. Therefore, the internal variables $F_s$ and $F_d$ describe total suspension force, which is in reality transmitted through several linkages, and total damping force, not exclusively supplied by the damper unit. As a result, these force variables which at first seem measurable using strain gauges, are in fact not directly measurable. Tyre force $F_t$ can be accurately recorded however, as it equates to the total vertical reaction force between the tyre and the wheelpans, and this can be found from a suitably placed load-cell.

Fortunately, the identification problem can be reformulated in terms of external variables alone. Equations 2.1.3, 2.1.5 and 2.1.6 can be combined to allow identification of the sprung and unsprung masses together,

$$F_i = M_w \ddot{x}_3 + M_b \ddot{x}_4$$

The stiffness and damping coefficients can then be identified from equation 2.1.5 using the estimated wheel mass, $\hat{M}_w$

$$F_i - \hat{M}_w \ddot{x}_3 = K_s x_2 + \sum_{i=1}^{5} R_i (x_3 - x_4) F_i$$

or alternatively from equation 2.1.6 using the estimated body mass $\hat{M}_b$

$$\hat{M}_b \ddot{x}_4 = K_s x_2 + \sum_{i=1}^{5} R_i (x_3 - x_4) F_i$$
In both of these equations the nonlinear model of equation 2.1.7 has been substituted to describe damping. If required, a model for the (non-rotating) tyre could be identified, as before, from 2.1.3.

### 3.2 Simulating Measurement Noise

The measurement noise on recorded signals is difficult to quantify. Several distinct effects can be considered however, though no claim is made that these cover all possibilities:

- Poor calibration can result in gain and offset errors in the signal.
- Linearity of the signal can be affected by the geometry of the vehicle, and/or the alignment of transducers attached to it.
- Hysteresis effects might be seen, particularly in strain gauge type transducers.
- Electrical noise and errors caused by analogue to digital conversion will be evident.
- High frequency vibrations which lie outside the dynamic range of interest can be interpreted as measurement noise.

Simple models can be derived for these sources of noise, but consideration should first be given to the evidence that might be directly available to a test engineer. For example, Figure 8 shows the auto-spectral density (ASD) for data collected from a body-mounted accelerometer fitted to a test vehicle. The car was driven at a steady speed of approximately 20m/s on a ‘B’ class road, and the data was sampled at 250Hz. Also plotted, is the same signal from a quarter car simulation using nominal parameter values. (The input signal for this simulation was derived from measurements of the test road profile.)

The suspension dynamics are evident in both traces at lower frequencies. However, above about 20Hz, higher frequency vibrations from the road and engine make up about 40% of the total measured acceleration signal. These vibrations, which are transmitted directly through paths such as the tyre carcass, engine mounts and suspension linkages, are unmodelled by the quarter car model; the simulated acceleration falls off rapidly, with nearly 99% of the total signal power below 20Hz.

We should assume that test conditions will be selected to minimise this error, either by band-limiting the input signal, or filtering the measurements. As the rig actuators have a finite bandwidth in any case, the former option is more desirable. Thus for simulation purposes an upper bandlimit of 20Hz will be assumed in the input process. Above this
Figure 8: Comparison of measured and simulated vehicle responses

frequency, the measurement noise on all transducers can be assumed to be at a low level; it can be modelled by additive white noise at a nominal RMS signal / noise ratio of 5%.

Calibration errors, nonlinearities and hysteresis effects cannot be detected directly from test data. They are highly correlated with, and therefore indistinguishable from the signal itself, and may represent high levels of noise. For ease of comparison, each of the nominal models used here has a 'worst-case' error of 20% of the RMS signal amplitude.

To represent gain and offset errors, denoting the noise-free simulated signal $x$, and the noise added signal $y$, we may write

$$y = \lambda x \quad /3.2.1/$$

and

$$y = x + \varepsilon \quad /3.2.2/$$

where

$$\lambda = 1.2$$

$$\varepsilon = 0.2RMS(x)$$

Figure 9(a) illustrates a simple model for nonlinearity, where the noise added signal is represented by a general, odd-valued cubic function:

$$y = \alpha x^3 + \beta x \quad /3.2.3/$$
α and β are chosen such that within a given operating range [-a,a] of the transducer, the maximum error is 0.2x. The parameters are easily derived by consideration of the error function, as

\[ \alpha = -0.25a^{-2} \]
\[ \beta = 1.15 \]

To model hysteresis, a discretely varying function is used (Figure 9(b)) where y is dependent on the time derivative of x such that the locus of points on the x / y plane follows a closed path as indicated by arrows in the diagram. y is defined by

\[ y = x + \varepsilon, \quad \dot{x} > 0 \]
\[ y = x - \varepsilon, \quad \dot{x} < 0 \]

with \( \varepsilon = 0.2 \) RMS \( \langle x \rangle \) as before.

Figure 9: Models for (a) nonlinearity and (b) hysteresis
3.3 Unmodelled Dynamics

Here we discuss specific effects that are not simulated by the quarter car model, yet which may have a large influence on vehicle dynamics. The aim is to represent those unmodelled dynamic effects which are most significant given the likely test conditions, without recourse to complicated three dimensional analysis of components. To test the success of identification, it must be possible to interpret simulation models in terms of the original quarter car model.

A number of factors exist, which have a greater or lesser effect on dynamics depending on the test conditions. Full vehicle effects such as roll and pitch can be minimised on a test rig, and changes in load distribution caused by accelerating and braking are avoided altogether, along with aerodynamic influences. Other effects are common to road and rig tests alike; the suspension has nonlinearities due to its geometry, and unmodelled friction and damping forces act within and between the suspension components. The engine mass represents an unmodelled degree of freedom as it vibrates on its mounts. Also, impulsive inputs may be felt from the bumpstops at the extremes of suspension travel.

One effect which is specific to the rig arises out of the need to constrain the vehicle’s motion in some way, to prevent the wheels from slipping off the wheelpans. Lateral and longitudinal restraints are generally used, and these will induce impulsive forces to the vehicle, on contact. The behaviour of the tyres on a test rig is also of interest. Most suspension geometries induce lateral movement in the wheels under excitation of the suspension, and for a non-rolling tyre, lateral scrub forces are likely to represent a significant modelling error, even though measures may be taken to reduce friction.

Most of these effects are not examined further. Nonlinearities will have a similar effect on identification whether their origin is mechanical or due to measurement error, so their influences are not separately considered. Impulsive forces are also ignored, as test conditions can be chosen to minimise their frequency, and if necessary impulse events can be located within the recorded data and excluded from the set of identification equations. Restricting attention to problems expected on rig tests, two effects remain - the motion of the engine mass on its mounts, and the influence of lateral tyre forces.
3.3.1 Engine mass vibrations

For a typical car, the orientation of the engine, and positioning of the mounts allows several modes of engine motion. To provide a simple model, the engine is taken as a lumped mass, evenly distributed over the front axle of the car, and assumed to move only in the vertical direction. The engine mounts are modelled as linear springs, with negligible damping; this is known to be a reasonable approximation for small deflections [36].

FORD specifications for the V6 2.9 litre Granada used as the test vehicle, give an engine mass of 300Kg, and figures for natural frequency of the engine's vertical motion are known to lie in the range 10-12Hz. Choosing $\omega_n = 11$Hz, the engine mount stiffness is $K_e = 700$kN/m for each front quarter of the vehicle, for which $M_e = 300/2$Kg. The simulation model (Figure 10) is adapted from the original quarter vehicle model by splitting the body mass into engine mass $M_e$ and the remaining sprung mass, $M_s$.

![Quarter vehicle simulation model including engine mass](image)
The state differential equations for this system are

\[
\begin{align*}
\dot{x}_1 &= v - x_3 \\
\dot{x}_2 &= x_3 - x_4 \\
\dot{x}_3 &= \left\{K_rx_1 - K_sx_2 - F_d\right\}/M_w \\
\dot{x}_4 &= \left\{K_sx_2 + F_d - K_px_3\right\}/M_b \\
\dot{x}_5 &= x_5 - x_6 \\
\dot{x}_6 &= \left\{K_\varepsilon x_5\right\}/M_e
\end{align*}
\]

where states \(x_1 - x_4\) are as defined previously, and \(x_5\) and \(x_6\) are the vertical deflection of the engine mount, and absolute vertical velocity of the engine mass respectively.

### 3.3.2 Lateral tyre slip properties

To derive a nominal model for lateral tyre scrub, consider Figure 11 which illustrates the basic geometry of a quarter vehicle with McPherson strut suspension, (as used at the front of the test vehicle) supported on a hydraulic shake rig.

![McPherson strut suspension geometry](image)

*Figure 11: McPherson strut suspension geometry*
As the suspension expands and contracts, the tyre contact patch will move from side to side. A resistance force $R$ will therefore be felt, and this leads to a vertical reaction, $F$ between the wheel and body, due to the suspension geometry. (The suspension linkages behave in the same way as a rigid link between the tyre contact patch and virtual reaction point, $V_p$ marked. The vertical reaction can be seen by considering parallel and perpendicular components of $R$ along this link. For further details see [11].)

In the linear approximation, the suspension force $F_s$ in the original quarter vehicle model can be altered:

$$F_{s(new)} = F_s + F$$  /3.3.2/

where,

$$F = -k_1R$$  /3.3.3/

and the positive constant $k_1$ is determined by the suspension geometry.

The tyre frictional force will usually be dependent on the size of contact patch, normal force applied, and the side slip deflection and velocity. For simplicity however, a simple viscous frictional force is assumed, dependent on normal force and side-slip velocity alone. (This might be expected if the contact patch is lubricated to allow easier slippage.)

Now, for small changes in angle $\theta$, the lateral deflection of the contact patch, $d$ is proportional to suspension deflection. It follows therefore, that

$$d \propto -V$$  /3.3.4/

where $V$ is the expansion velocity of the suspension. The viscous friction model can thus be approximated by

$$R = \Omega (V) \mu N$$  /3.3.5/

where the function $\Omega(.)$ is illustrated in Figure 12.

The normal force $N$ can be computed in terms of tyre deflection alone for the purposes of estimating friction. Note however that as $x_1$ is defined as a dynamic variable, the point of zero normal force is not at zero tyre deflection. For a lifting tyre model,

$$N = K_r \phi (x_1)$$  /3.3.6/

where the function $\phi(.)$ is also illustrated in Figure 12. In the figure, $x_{1r}$ represents the deflection of the tyre in equilibrium, which can be estimated from the total quarter vehicle mass and linear tyre stiffness as 23mm.

Combining equations 3.3.3, 3.3.5 and 3.3.6 the reaction force is described by

$$F = -\mu k_1 K_r \Omega (V) \phi (x_1)$$  /3.3.7/
As we are interested in a nominal model to illustrate the effect, no attempt is made to identify the values \( \mu \) or \( k_1 \). Instead, the forcing error is modelled at a nominal magnitude, with \( \mu k_1 \) chosen to satisfy the 20\% RMS error to signal magnitude condition used in Section 3.2. The effect is simulated using the standard quarter vehicle model, with the suspension force modified according to equations 3.3.2 and 3.3.7.

![Figure 12: Tyre slip model nonlinearities](image)

### 3.4 Simulation Conditions

Each of the measurement and modelling error conditions were examined separately using the procedure summarised below. To reduce the influence of randomness, each result is a mean value over 20 tests (as in the experiment of Chapter 2).

1. Signals are generated using nonlinear continuous-time simulations of the standard or modified quarter vehicle model. 20 seconds of data is generated, sampled at 100Hz. The road input model of equation 3.1.1 is used, with bandwidth 0.4 - 20Hz.

2. Measurement errors, where applicable, are added to an individual signal.† White noise is also added to all signals at an amplitude giving 5\% RMS signal / noise. A reference case is also considered, where white noise represents the only source of error.

3. Mass parameters are identified using OLS applied to equation 3.1.2. The spring stiffness and damper map are identified from equation 3.1.3 (based on wheel acceleration). Equation 3.1.4 (based on body acceleration) is also employed in the analysis of modelling errors.

† All measurement errors were tested separately on all signals, though only results considered to be of interest will be presented.
3.4.1 Measurement error results

Identified mass and stiffness parameters, and corresponding fitting errors are given for five test cases in Table 3. Identified damper maps are illustrated in Figure 13.

<table>
<thead>
<tr>
<th>Error type</th>
<th>M_w (Kg)</th>
<th>M_b (Kg)</th>
<th>Fitting error (%)</th>
<th>K_s (kN/m)</th>
<th>Fitting error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True parameters</td>
<td>30</td>
<td>350</td>
<td>-</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Reference case</td>
<td>29.9</td>
<td>349.1</td>
<td>7.0</td>
<td>19.9</td>
<td>9.7</td>
</tr>
<tr>
<td>Gain error in V</td>
<td>29.9</td>
<td>349.1</td>
<td>7.1</td>
<td>19.9</td>
<td>10.4</td>
</tr>
<tr>
<td>Offset error in x_2</td>
<td>29.9</td>
<td>349.1</td>
<td>7.0</td>
<td>19.9</td>
<td>9.7</td>
</tr>
<tr>
<td>Nonlinearity in a_w</td>
<td>24.8</td>
<td>347.5</td>
<td>8.0</td>
<td>19.8</td>
<td>10.7</td>
</tr>
<tr>
<td>Hysteresis in a_b</td>
<td>22.3</td>
<td>336.8</td>
<td>13.5</td>
<td>18.5</td>
<td>16.6</td>
</tr>
</tbody>
</table>

Table 3: Results of identification from data with measurement errors

Figure 13: Damper maps identified from data with measurement error
As we might expect from the study of Chapter 2, the control case provides good parameter estimates, with bias typically below 1%. The fitting error of 7.0% and 9.7% in the mass and map equations is typical of 5% white noise added to all the measurements.

The effect of gain error on most signals is obvious; the gain change of a signal is directly compensated by a proportionate change in the parameter it multiplies, and equation fitting errors do not change from the control case. This means that gain errors are virtually undetectable, yet they have a direct influence on parameter accuracy; elimination of gain error, through accurate calibration practices is therefore of vital importance to the validity of any system identification experiment. An exception to the ‘invisible’ nature of gain error is seen in \( V \) however, where slightly increased fitting errors occur for the damper map identification. The identified damper map is unable to fit exactly without a revised choice of \( V \) breakpoints; for this and other reasons, it may be desirable to optimise the positioning of breakpoints.

Offset errors can also be ‘absorbed’ by bias in the identified parameters. Figure 13(b) shows an interesting example, where the effect of a 5mm offset in \( x_2 \) is a shift in the identified damper map. The shift \( Y \) can be predicted from the offset,

\[
Y = \hat{K}_2 \times (0.005) = 100N
\]

and \( \hat{K}_2 \) is unbiased as a result of transferring the constant force error on to the identified damper map. A similar effect is seen for offsets in other signals in the map identification equation, except for \( V \).

The elimination of offset errors may seem trivial in practice, by removing the dynamic mean from all measured signals; in a sense this means adopting a ‘what goes up must come down’ philosophy over the whole test. This is not always true however. Standard passive dampers are asymmetrical, having a lower rate in compression than in extension; this gives rise to a jacking effect, where the net dynamic effect of suspension excitation is a compression of the suspension spring. For these simulations this compression is of the order 20mm, and similar values have been seen in measured signals. One method of estimating the true offset is to adjust all the nodal forces equally such that \( F_3 = 0 \), and equate this shift in force to the jacking effect, as in equation 3.4.1. This does however make the assumption that no significant asymmetrical friction exists in the system at very low damper velocities.

Nonlinearity errors are not compensated entirely by bias in the parameters, and fitting errors are increased in both identification equations as a result. For the nonlinearity model used here, the bias caused is similar to that caused by positive gain error, though not as
severe. The cause can be seen by considering the Gaussian model for the road input signal; much of the data for each signal is at low amplitude, corresponding with those parts of the nonlinearity model where the corrupted signal is greater than the original. Plotting equation error against the affected variable (Figure 14) the cause of error is clearest from correlations at high magnitude, and from the characteristic swirling pattern. Errors may not only correlate with the faulty signal though, as we see in the next example, and later in Chapter 6.

-700 -600 -500 -400 -300 -200 -100 0 100 200 300

Figure 14: The effect of measurement nonlinearity on equation errors

Hysteresis has a significant biasing effect on many parameters and also induces large fitting errors. Of particular interest is the mass identification result, where one acceleration signal is corrupted, yet both mass parameters are biased. In the illustrated case, hysteresis in body acceleration has a biasing effect on the spring and damping parameters, through the biased estimate of wheel mass.

To understand the cause, consider again the structure of the hysteresis model, which applies error to a signal according to the sign of its derivative. For example, hysteresis error in the body acceleration signal depends on $\dot{x}_4$. Differentiating 3.1.4 we can write

$$\dot{x}_4 = \left\{ \dot{K}_x \dot{x}_2 + \sum_{i=1}^{s} R_i (\dot{x}_3 - \dot{x}_4) \dot{F}_i \right\} / \ddot{M}_b$$
As the error is correlated with $\ddot{x}_4$, it is also correlated with wheel acceleration and hence causes a bias in wheel mass. Figure 15 illustrates the ability of biased estimates $\hat{M}_b$ and $\hat{M}_w$ to ‘explain’ (or minimise) this hysteretic error. For clarity, uncorrupted accelerations are plotted against the force that each vibrating mass is expected to contribute to the total tyre force ($30a_w$ or $350a_b$) plus the error from hysteresis in the body acceleration signal; the purple line represents an unbiased parameter estimate and the red line shows an OLS solution.

Figure 15: Explanation of hysteresis corruption in measured body acceleration, by mass parameters

Because the error always has the same sign as the derivative $\ddot{x}_4$, it can most effectively be corrected by a bias in $\hat{M}_w$, the estimate $\hat{M}_b$ being biased only slightly.

In practice, errors are likely to be less clearly defined than in the examples illustrated. A number of effects may occur simultaneously, and correlations between equation error and signals may not be obvious. Plots such as those in Figure 14 and Figure 15 can nevertheless provide useful diagnostic information. Note that fitting errors do not directly measure accuracy in the parameters. High errors may occur with no correlation to the measured dynamics and hence no bias being introduced. Also we have seen large biasing effects with no associated fitting errors.

Up to now we have restricted attention to identification based on the wheel mass acceleration equation. By considering both equations 3.1.3 and 3.1.4, more information about the cause of errors may be obtained; this is investigated in the next section.
3.4.2 Modelling error results

Identified parameters and fitting errors are presented for the tyre slip and engine bounce error conditions in Table 4 and Figure 16.

<table>
<thead>
<tr>
<th></th>
<th>( M_w ) (Kg)</th>
<th>( M_b ) (Kg)</th>
<th>Fitting error (%)</th>
<th>( K_s ) (kN/m)</th>
<th>Fitting error (%)</th>
<th>( K_a ) (kN/m)</th>
<th>Fitting error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True parameters</td>
<td>30</td>
<td>350</td>
<td>-</td>
<td>20</td>
<td>-</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Tyre slip model</td>
<td>29.9</td>
<td>349.2</td>
<td>7.0</td>
<td>19.7</td>
<td>11.4</td>
<td>19.6</td>
<td>10.8</td>
</tr>
<tr>
<td>Engine bounce model</td>
<td>21.8</td>
<td>14.5</td>
<td>60.3</td>
<td>18.5</td>
<td>31.4</td>
<td>0.8</td>
<td>97.3</td>
</tr>
</tbody>
</table>

*Table 4: Results of identification from data with modelling error*

For the tyre slip model, all the parameters except damper map nodal forces have been identified very accurately. The mass identification incurs no more fitting error than the reference case, and this illustrates the independent nature of the two stages of identification; by altering the suspension force model the vehicle dynamics are altered, but the total mass acceleration is still faithfully represented by total force at the wheelpan.
The tyre slip model, although it appears complex, effectively reduces to an additional asymmetrical damping effect; a result which might have been deduced from the original equations. As the tyre rarely lifts off the wheelpan, and most of the dynamic motion is within the range $V = \pm 0.1 \text{m/s}$, functions $\phi(.)$ and $\Omega(.)$ have little influence. The asymmetry results from a positive correlation between $x_1$ and $V$, as damper forces are reacted by the tyre spring. $\phi(.)$ has the effect of inducing greater forces on compression strokes of the damper ($V$ negative), and less force during extension strokes.

As with examples of gain and offset measurement error, this effect may be difficult to detect in practice, as fitting errors remain low. Care should therefore be taken to minimise tyre scrub and reduce the risk of parameter bias, even given the nominal nature of the model used here.

It is immediately apparent from both Table 4 and Figure 16 that the engine bounce model causes an alarming failure in this identification method. The identified mass parameters are completely meaningless and the equations are dominated by errors. We do have an indication of cause though, through the two separate map identifications; low fitting errors and a relatively successful identification of parameters occurs in the wheel acceleration case, whereas the body acceleration case produces meaningless estimates. The error therefore appears to be associated with the body acceleration signal, and closer analysis confirms this.

Forces on the vehicle body are described in the body acceleration analysis of equation 3.1.4 by

$$F_s = M_p \dot{x}_4$$

When the total mass is split however, the suspension force arises from independent vibrations

$$F_s = M_r \dot{x}_4 + M_e \dot{x}_6$$

which are not consistent with equation 3.1.4. Considering the same force in analysis of wheel-hub accelerations however,

$$F_s = F_t - M_w \dot{x}_3$$

where the force at the tyre is measured, so it implicitly includes the cumulative mass-acceleration effects:

$$F_t = M_w \dot{x}_3 + M_r \dot{x}_4 + M_e \dot{x}_6$$

The wheel-hub equation 3.1.3 is thus consistent with equation 3.4.3.
Avoidance of the corrupting influence of engine shake as it is modelled here, is not trivial. The engine motion is concentrated in a coupling with the wheel and body oscillations that cannot be explained by the four state model. For an equivalent linear system, the eigenvector of this mode is

\[-0.0049 \pm 0.0055j\]
\[-0.00068 \pm 0.00045j\]
\[-0.382 \pm 0.338j\]
\[-0.351 \pm 0.385j\]
\[0.0129 \pm 0.0117j\]
\[0.459 \pm 0.506j\]

where the phase angle between the 'in phase' motion of \(x_3\) and \(x_4\) and the engine vibration \(x_6\) is nearly \(180^\circ\). If such a mode is significantly excited, no identification method could provide a reasonable estimation of the parameters without additional instrumentation. The only solution is to prevent such excitation.

As the engine shake mode occurs over a narrow range of frequency, notch filters on the input signal or measurements may be successful. These may be difficult to design though, given that the exact modal characteristics are unknown. Most of the useful dynamic information in the measurements is expected below the engine bounce modal frequency, so a simpler solution is to restrict the input bandwidth.

The purple line in Figure 17 illustrates the effect of varying the maximum frequency of input, and the identification improves progressively as this decreases. However a stable, reliable estimation is not achieved until below 6Hz, and referring back to Figure 5 in Chapter 2, this would include very little of the wheelhop mode, giving a correspondingly uncertain estimate of \(M_w\).

Given the trends in fitting errors and identified body mass above 6Hz, it appears that we require a filter which attenuates errors caused by the unmodelled mode more progressively across frequency. The green line in Figure 17 illustrates the effect of one such filter which gives much improved results. The change in the identification method is simple mathematically, yet fundamentally it is profound; equation 3.1.2 has been integrated, such that the OLS equation set represents a record of momentum changes as opposed to force variations:

\[
\int_0^\tau F_\tau(\tau) d\tau = M_w x_3 + M_b x_4
\]

\%/3.4.5/\]
Figure 17: Comparison of identification methods over varying input bandwidth - engine mass disturbance case
The improvement in results between 6 and 12Hz is easiest to understand in terms of momentum: below resonance the engine bounce dynamics are well bounded, so the total change in momentum they induce in the body is small. This compares with larger momentum changes that occur between the wheel and body over the wider wheel-hop mode.

The alternative approach certainly appears to offer significant benefits under conditions of engine bounce disturbance. In the next chapter, a more thorough analysis is undertaken, resulting in a new system identification methodology.
Chapter 4

Identification using Randomised Integration

We have seen that there may be benefits from integrating the equations of motion originally used for identification, at least to minimise the effect of engine shake. In this chapter, the integrated scheme is investigated more thoroughly, and developed into a formal method. A statistical analysis of errors is carried out, which shows that an optimal solution can be obtained based on the least squares approach, and conditions for avoiding bias are examined. Finally, the simulation study of Chapter 3 is briefly revisited, and results for the new method are compared with those identified using direct OLS.

4.1 Finite Interval Signal Integration

The integrated form of the mass identification equation, 3.4.5 leads to an improved identification under simulated engine-shake errors. In practice however, problems will exist if there are low-frequency errors in measured data which is integrated. The continuous-time transfer function for an integrator can be written in terms of frequency $f$ by

$$H(j\omega) = \frac{1}{j\omega}, \quad f = \omega/2\pi$$

so amplification occurs at frequencies below 0.16Hz. Where the integrated measurement can be replaced by an alternative measurement (e.g substituting measured velocity for integrated acceleration) this does not present a problem. However, very low frequency errors in tyre force or suspension deflection measurements are exacerbated by numerical integration, and it can be shown that subsequent application of OLS would result in a failed identification. Offset errors also clearly represent a significant problem in this respect.

These problems might be avoided using a high pass digital filter, but the more direct approach is to integrate over finite time intervals. Choosing a range of integration $[t, t+T]$, equation 3.4.5 becomes

$$\int_{t}^{t+T} F_i(\tau) \, d\tau = M_w \int_{t}^{t+T} \dot{x}_3(\tau) \, d\tau + M_b \int_{t}^{t+T} \ddot{x}_4(\tau) \, d\tau$$
or
\[ \int_{t}^{t+T} F_i(\tau) d\tau = M_w \Delta x_3 + M_b \Delta x_4 \]  
\[ /4.1.1/ \]
where
\[ \Delta x \equiv x(t+T) - x(t) \]

Similarly, the suspension identification equations 3.1.3 and 3.1.4 become
\[ \int_{t}^{t+T} F_i(\tau) d\tau - M_w \Delta x_3 = K_x \int_{t}^{t+T} x_2(\tau) d\tau + \sum_{i=1}^{5} F_i \int_{t}^{t+T} R_i \{ x_3(\tau) - x_4(\tau) \} d\tau \]  
\[ /4.1.2/ \]
and
\[ M_b \Delta x_4 = K_x \int_{t}^{t+T} x_2(\tau) d\tau + \sum_{i=1}^{5} F_i \int_{t}^{t+T} R_i \{ x_3(\tau) - x_4(\tau) \} d\tau \]  
\[ /4.1.3/ \]

Clearly an overdetermined set of equations could be extracted by increasing \( t \) across the range \([0, L-T]\) for a given choice of \( T \) and data batch length \( L \); an OLS solution for the parameters could then be evaluated as before. However, the frequency domain effect of the new finite interval integration should first be investigated. This can be done as above, using the transfer function, which for a general integration
\[ y_{t+T} = \int_{t}^{t+T} u(\tau) d\tau \]
takes the form
\[ H(j\omega) = \frac{e^{-j\omega T} (1 - e^{-j\omega T})}{j\omega} \]

From this the gain characteristic \( A(\omega) \) can be evaluated
\[ A(\omega) = |H(j\omega)| = T \left| \frac{(\cos \omega t - j \sin \omega t) (1 - \cos \omega T + j \sin \omega T)}{j\omega T} \right| \]
and after some manipulation, the terms in \( \omega t \) cancel, giving
\[ A(\omega) = \frac{2 \sin (\omega T/2)}{\omega} = T \text{sinc} (\omega T/2) \]

This function is illustrated by the green curve in Figure 18 for which \( T \) has been arbitrarily set at 0.1 seconds. There is no magnification of very low frequencies, but note the zero gain nodes at multiples of 10Hz. This could provide a desirable filter if disturbance frequencies are known precisely, but with no a-priori knowledge, it may be unwise to use a single value for \( T \). The figure also illustrates the effect of choosing half the samples of
length 0.1 seconds and half of length 1/3 second (red curve); the attenuation is then less severe until the nodes coincide at 30Hz. Varying $T$ still further results in the purple curve, for which samples have been taken uniformly over the range 0.1 - 1 seconds. This last set has been generated using

$$A(\omega) = \frac{1}{N} \sum_{i=1}^{N} T_i \text{sinc} (\omega T_i / 2)$$

for the $N$ samples taken. Choosing a range of sample lengths smooths the gain response, while retaining broad weighting differences across frequency; hopefully the fact that integration is still being used will retain the earlier improvements in identification. Clearly there is unlimited scope for the choice of $T$, and if a range of values is employed, a larger set of equations can be generated from a given batch of data than under the direct OLS approach.

\begin{center}
\textbf{Figure 18: Signal gain as a function of integration length}
\end{center}

In practice it is not possible to use all time intervals within a range, but the same frequency response is obtained on average if the sample duration is set randomly, using a uniform distribution. The start time $t$, for each frame must also be set, and again it is convenient and natural to use a randomised approach. Thus a set of samples can be extracted according to equations 4.1.1 to 4.1.3, using

$$T_k \sim U(\alpha, \beta)$$

$$t_k \sim U(0, L - T_k)$$

/4.1.4/
where $\alpha$ and $\beta$ are suitable limits on the integration length. An ordinary least squares solution of this overdetermined set will thus identify parameters according to a Randomised Integral Error Criterion (RIEC).

An added advantage of this randomised approach is that it allows a statistical analysis of errors to be carried out. This is the subject of Section 4.2, where a formal proof of optimality is given for the new method.

### 4.2 Statistical Analysis of Errors

Where $N$ random samples are taken according to equation 4.1.4 from a fixed batch of measured data of duration $L$, each of the identification equations 4.1.1 to 4.1.3 can be written in the form

$$U\theta = y + e$$

where the $(N \times 1)$ error vector $e$ has simple statistical properties. Since the order of sampling is irrelevant to the error statistics, the error covariance matrix, $C = E[ee^T]$ is unchanged by any permutation carried out simultaneously on the rows and columns of $C$. The most general form is

$$C = E[ee^T] = \sigma^2 \begin{bmatrix} 1 & e & e & \ldots & e \\ e & 1 & e & \ldots & e \\ e & e & 1 & \ldots & e \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e & e & e & \ldots & 1 \end{bmatrix}$$

where $\sigma^2$ represents the variance of autocorrelated errors, and the off-diagonal elements are non-zero due to expected cross correlations between any two samples. This correlation is due in part to the time correlated nature of measurement and modelling errors, and it is also caused by the expectation of overlap between the integration periods of any two samples. $e$ lies in the range $0-1$.

Now, for OLS to represent an optimal estimate of the parameters, the errors are assumed to be uncorrelated, so direct application of OLS to the solution of equation 4.2.1 yields a sub-optimal solution. It is of interest to note here that the direct OLS method of Chapter 3 suffers from a similar inoptimality, due to time correlations in the errors, though in that case the off-diagonal components will not be equal.
A transformation of variables, based on the eigenvalues and eigenvectors of $C$ can be used to ‘whiten’ the noise process. For each eigenvalue $\lambda_i$ there exists an eigenvector $v_i$ for which

$$Cv_i = \lambda_i v_i$$

By inspection, one such (normalised) eigenvector is

$$v_1 = \frac{1}{\sqrt{N}} [1, 1, 1, 1, 1, 1, 1, 1, 1]^T$$

with associated eigenvalue

$$\lambda_1 = 1 + (N - 1) \varepsilon$$

Vectors orthogonal to $v_1$ also represent eigenvectors. If $N$ is an integer power of two, a suitable set may be written down easily. For example, if $N = 8$ we have

$$v_2 = \frac{1}{\sqrt{N}} [1, 1, 1, -1, -1, -1, -1, -1]^T$$

$$v_3 = \frac{1}{\sqrt{N}} [1, 1, -1, 1, 1, -1, -1, -1]^T$$

$$v_4 = \frac{1}{\sqrt{N}} [1, -1, 1, -1, 1, 1, -1]$$

$$v_5 = \frac{1}{\sqrt{N}} [1, -1, -1, -1, 1, 1, 1]$$

$$v_6 = \frac{1}{\sqrt{N}} [-1, -1, 1, -1, 1, -1, 1]$$

$$v_7 = \frac{1}{\sqrt{N}} [1 -1, 1, 1, 1, -1, 1]$$

$$v_8 = \frac{1}{\sqrt{N}} [1 -1, 1, 1, 1, -1]$$

More generally, the Gram-Schmidt process can be used (described for example, in [49]). It is easily verified that the remaining eigenvalues are equal

$$\lambda_i = (1 - \varepsilon), \quad i = 2, N$$

An orthogonal transformation matrix $Q$ can now be defined ($Q^T = Q^{-1}$)
Identification using Randomised Integration

and equation 4.2.1 can be re-written

\[ \tilde{U}\theta = \tilde{y} + \tilde{e} \]

where \( \tilde{U} = Q^T U \), \( \tilde{y} = Q^T y \), and \( \tilde{e} = Q^T e \)

The transformed problem has a simpler form for error covariances.

\[
R = E[\tilde{e}\tilde{e}^T] = \sigma^2 \begin{bmatrix}
1 + (N - 1) \varepsilon & 0 & 0 & \ldots & 0 \\
0 & 1 - \varepsilon & 0 & \ldots & 0 \\
0 & 0 & 1 - \varepsilon & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & 1 - \varepsilon
\end{bmatrix}
\]

and this diagonal matrix could be used to solve for the parameters by the Generalised Least Squares (GLS) method [30], also known as the weighted least squares method. For a GLS solution, the regressor matrix and output vector are each multiplied by \( R^{-1/2} \), to give an optimal estimate

\[
\hat{\theta} = (\tilde{U}^T R^{-1} \tilde{U})^{-1} \tilde{U}^T R^{-1} \tilde{y}
\]

In this case however, evaluation of equation 4.2.5 relies on explicit knowledge of \( \varepsilon \), as well as computationally expensive calculations using the \( N \times N \) matrix \( Q \). Here an alternative transformation matrix can be derived however, considering the case as \( N \) becomes very large. Rewriting equation 4.2.4:

\[
R = \sigma^2 (1 - \varepsilon) \begin{bmatrix}
1 + \frac{N\varepsilon}{(1 - \varepsilon)} & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & 1
\end{bmatrix}
\]

\[
\lim_{N \to \infty} R^{-1/2} = \frac{1}{\sigma \sqrt{1 - \varepsilon}} \begin{bmatrix}
0 & \ldots & 0 \\
\vdots & \ldots & \vdots \\
0 & \ldots & I_{(N - 1)}
\end{bmatrix}
\]
the limiting covariance matrix applies zero weighting in the direction of the first eigenvector in N dimensional space, and equal weighting in all other directions. Combining $R^{-1/2}$ and $Q$, an alternative transformation can be formed:

$$U^* = H^T U, \quad y^* = H^T y, \quad e^* = H^T e$$

where

$$H = \begin{bmatrix}
0 & \ldots & 0 \\
\vdots & \ddots & \vdots \\
0 & \ldots & I_{(N-1)}
\end{bmatrix} Q^T$$

Now an optimal solution which we will term asymptotic GLS is given by applying a least squares optimisation to these transformed variables. Equation 4.2.1 becomes

$$HU\theta = Hy + He$$

and recalling the OLS solution of equation 2.2.2

$$\hat{\theta} = (U^T H^T H U)^{-1} U^T H^T H y$$

The effect of this transformation is to project the least squares problem into the subspace perpendicular to $v_1$:

$$H^T H = \begin{bmatrix}
0 \\
v_2^T \\
v_3^T \\
\vdots \\
v_N^T
\end{bmatrix}
\begin{bmatrix}
v_2 \\
v_3 \\
\vdots \\
v_N
\end{bmatrix}
\begin{bmatrix}
0 \\
v_2^T \\
v_3^T \\
\vdots \\
v_N^T
\end{bmatrix} = v_2 v_2^T + v_3 v_3^T + \ldots + v_N v_N^T = (I - v_1 v_1^T)$$

and this provides a simple form for transforming the original problem. Given the definition of $v_1$, the effect of equation 4.2.9 in the following definitions

$$\bar{U} \equiv (I - v_1 v_1^T) U, \quad \bar{y} \equiv (I - v_1 v_1^T) y, \quad \bar{e} \equiv (I - v_1 v_1^T) e$$

is the removal of the mean from each column of the regressors and output. It is easily verified, using the property

$$(H^T H)^n = H^T H$$
that the transformation of equation 4.2.10 is equivalent to that of 4.2.8, yielding the same OLS solution. This implies that after mean removal, an OLS solution of the original equations gives an optimal solution in the asymptotic GLS sense, provided that $N$ is large.

Having obtained an optimal parameter estimate based on RIEC, an analysis of errors is required. It can be shown that the optimal GLS solution exists at the minimum covariance estimate of $\hat{\theta}$ [30], and $\text{cov}(\hat{\theta})$ is related to equation error by

$$\text{cov}(\hat{\theta}) = \text{cov}(\tilde{e}) E[(\tilde{U}^T \tilde{U})^{-1}]$$

In this case $\text{cov}(\tilde{e})$ is defined as in equation 4.2.6, with magnitude $\sigma^2 (1 - \varepsilon)$ over the $(N - 1)$ dimensional subspace spanned by $\tilde{U}$. Now, from their definitions $\sigma^2$ and $\varepsilon$ are independent of $N$, yet by inspection $E[(\tilde{U}^T \tilde{U})^{-1}]$ decreases in magnitude as $N$ increases. The theory thus seems to predict a progressively more accurate estimate $\hat{\theta}$ as $N$ increases even though the original batch of data remains fixed, with finite length $L$.

In fact the accuracy of parameter estimates is also dependent on systematic bias, which may be inherent in the data batch, and may also be affected by the choice of integration intervals $T$. As $\tilde{U}$ is non-deterministic, probability limits can again be used to show the convergent bias that may occur. Including the error vector in the OLS estimation:

$$\text{plim} \{ \hat{\theta} \} = \text{plim} \{ (\tilde{U}^T \tilde{U})^{-1} \tilde{U}^T (\tilde{U} \theta - \tilde{e}) \}$$

and making use of the property

$$\text{plim} \{ AB \} = \text{plim} \{ A \} \text{plim} \{ B \}$$

which is true for any two matrices $A$ and $B$ that are functions of the same random variables [30], we can define

$$R \equiv \text{plim} \{ \frac{1}{N} \tilde{U}^T \tilde{U} \}, \quad c \equiv \text{plim} \{ \frac{1}{N} \tilde{U}^T \tilde{e} \}$$

equation 4.2.11 can now be written

$$\text{plim} \{ \hat{\theta} \} = \theta - R^{-1} c$$

which illustrates that one or more parameters will converge to a biased estimate of the true parameter(s) if $R^{-1}$ and $c$ are both non-zero. The $pxp$ matrix $R$ represents the average covariance of regressors, and this must clearly be non-zero for a unique solution to exist. The $px1$ vector $c$ represents the average covariance between regressors and errors, and this will be non-zero if the regressors contain any measurement error - resulting in a second-order bias as discussed in Section 2.2.
4.3 Convergence of the RIEC Method

In the last section we derived an optimal parameter estimate from the RIEC method, provided the number of samples is large. Here we assess convergence of estimated parameters by carrying out repeated identifications using \(N\) samples taken from a single batch of simulated data. Note that as \(N\) increases, and assuming the identified parameters converge, they will converge to optimal estimates in the asymptotic GLS sense, with bias prescribed by equation 4.2.12.

Parameters were identified from three 20 second batches of simulated data taken from the study of Chapter 3, each corrupted with a different source of error. The RIEC identification was carried out using integration intervals uniformly varying in the range 0.05 to 0.5 seconds (this choice is justified in Section 4.4).

The identified wheel mass is illustrated in Figure 19, where \(N\) increases from 20 to 10000 in steps of 20. This parameter is typical of all those identified, and clearly the estimate is convergent, although the rate of convergence differs between tests and also between parameters. The bias error in \(\hat{M}_w\) is similar to that seen under direct OLS in Chapter 3.
To select a suitable default for $N$ in future tests, the trade-off must be considered between reduction of parameter variance, and computation time and capacity. The hardware and software configuration used here allow very large sample sets to be processed within a few seconds, and a maximum sample set of 14000 is possible. Given the small additional benefit of selecting higher values that is indicated by Figure 19, $N = 8000$ is chosen.

4.4 Choice of Test Parameters for RIEC Identification

The number of samples is one of several free parameters in the vehicle test environment and RIEC identification procedure that may affect accuracy in the parameter estimates. We have seen that the frequency content of the input 'road' profile can be altered to reduce the corrupting influence of an engine mode, and a similar influence may be achieved by varying the choices of integration length. In this section, variations in these factors, together with test duration, $L$ are investigated to assess sensitivity in the RIEC identification, and to decide suitable conditions for the general identification of an unknown quarter vehicle system.

Default settings for the conditions of interest are:

- The input signal is taken as a 'Robson' road, as in Section 3.1, with bandwidth $F_{\text{min}} = 0.4\text{Hz}$, $F_{\text{max}} = 10\text{Hz}$
- Test duration, $L = 20$ seconds
- Integration intervals are randomly selected in a uniform distribution, with $T$ in the range 0.05 seconds to 0.5 seconds.† A different randomisation is used for each identification.

The simulated error scenarios of Chapter 3 are used as a basis for the investigation, and to judge likely accuracy in the identification of an unknown system, we require a set of criteria which do not vary significantly across error types. Fitting error and bias from simulated identifications could be used, but alone these may prove misleading. We have seen from Chapter 3 that the former provides a good indicator of the model's ability to explain the data, but not necessarily with unbiased estimates. Direct minimisation of parameter bias is a more appropriate objective, but this may result in selection of operating conditions that are tuned to specific simulated error scenarios.

†. It can be shown that the average absolute change in a sinusoid function

$$|\sin(2\pi f(t + T)) - \sin(2\pi ft)|$$

is maximised by choosing $T = 1/2f$. The choice of $T_{\text{min}}$ and $T_{\text{max}}$ is thus based on maximising the regressors at nominal modal frequencies $f = 1$ and $f = 10\text{Hz}$. 


Recalling from equation 4.2.12

\[ b = R^{-1}c \]

the bias, \( b \) arising from a particular simulated batch of data may be reduced by a choice of operating conditions that emphasise values in \( R^{-1} \) and \( c \) which cancel in their product. In the identification of an unknown system, these cancellations may not be repeated. An alternative approach is to consider the magnitude, or norm of \( R^{-1} \) and \( c \) separately. The inequality

\[ \|b\| \leq \|R^{-1}\|\|c\| \]

always holds, so minimisation of \( \|R^{-1}\| \) and \( \|c\| \) has the effect of minimising the ‘worst case’ bias.

Note however, that absolute values are being considered here, so the parameters (and hence their bias) vary by orders of magnitude. To place equal importance on the same percentage bias in each of the identified parameters, \( R^{-1} \) and \( c \) should be suitably scaled. This is most conveniently done by scaling the regressor matrix according to order of magnitude estimates of the parameters, \( \theta^* \). Writing

\[ \bar{U}_\theta = \bar{U}\Phi \]

where \( \Phi \) is the diagonal matrix of the \( \theta^* \), the normalised probability limits become:

\[ R_\theta^{-1} = \text{plim} \left\{ \frac{1}{N} \Phi^{-1} (\bar{U}^T\bar{U})^{-1} \Phi^{-1} \right\} \]

and

\[ c_\theta = \text{plim} \left\{ \frac{1}{N} \Phi \bar{U}^T e \right\} \]

The normalised worst case bias, termed the \( b \) norm, is then given by the product of the normalised \( R \) norm and \( c \) norm:

\[ \|b_\theta\| = \|R_\theta^{-1}\|\|c_\theta\| \]

A number of definitions exist for the norm of a matrix; here we use the Euclidean norm:

\[ \|R_\phi^{-1}\| = \max \left\{ \lambda_i \right\} \quad i = 1, p \]

and

\[ \|c_\phi\| = \sqrt{c_{\phi 1}^2 + c_{\phi 2}^2 + \ldots + c_{\phi p}^2} \]
where $\lambda_i$ are the singular values of $R^{-1}_\phi$, which represent the magnitudes in each of $p$ orthogonal directions in the parameter space, of the columns of $R^{-1}_\phi$. Given that the regressor matrix $\tilde{U}_\phi$ has been scaled, the $R$ norm can be interpreted as a measure of conditioning in the regressors, since minimisation of $\|R^{-1}_\phi\|$ is equivalent to maximisation of the smallest singular value of $\tilde{U}_\phi$.

Minimisation of the $R$ and $c$ norms seems an appropriate goal in choosing operating conditions, but two problems still remain. As limiting stochastic variables, the norms cannot be explicitly calculated. Conclusions will therefore have to be made on the basis of one or more tests using the choice of $N = 8000$ samples, with variance in the norm estimates taken into account. The choice of $\Phi$ also presents a problem, as the most sensible estimate for $\theta_3^*$ (the third nodal damper force, corresponding to zero suspension velocity in the damper map identification) is zero, which makes $\Phi$ singular. It turns out that in order to make the $R$ norm dependent on the other suspension parameters, the influence of $F_3$ must be artificially reduced by setting a high value for $F_3^*$. The values taken for $\theta^*$ are thus

$$M_w^* = 30, M_b^* = 300, K_s^* = 2000, F_i^* = [1600, 600, 1000, 600, 1600]$$

where the remaining $F_i^*$ follow a linear estimate for damping rate of 2000Ns/m.

This approach provides a reasonable way of choosing conditions to maximise the potential of the RIEC method. It is likely however, that choices of test parameters which minimise $\|R^{-1}_\phi\|$ will not exactly match those which minimise $\|c_\phi\|$. Also, the best parameter choices may vary across a range of simulated measurement and modelling error conditions. Values are thus investigated over a range of simulations, and other indicators such as fitting error are useful in deciding settings for identification of the unknown system.
4.4.1 Test duration

The effect of test duration, $L$ on the identification of mass parameters is illustrated in normalised form, in Figure 20. Samples were taken from the first $\lambda$ seconds of a single reference case simulation test, and values of $\lambda$ between 1 and 20 have been considered at 0.1 second intervals. For each identification, the number of samples was limited in proportion to the test length considered.

![Figure 20: The influence of test duration on mass identification](image)

As $L$ increases, both parameters settle to estimates with low bias, but at different rates. This difference is probably due to the frequency content of each regressor, but intuitively we might have expected the parameters to converge the other way round, with the faster wheelhop mode invoking a more rapid convergence in $\hat{M}_w$. The result suggests that in fact the wheelhop mode is well characterised only over a variety of body bounce conditions, established after around 10 seconds for the illustrated case.

The bias norms settle over a similar period, although the $c$ norm increases slightly between 18 and 20 seconds; this may be in response to an uncharacteristically large body bounce event that was seen in the time history. To achieve suitably characteristic parameters in the identification, $L$ might sensibly be set at 30 or 40 seconds. Clearly this could be reduced if many tests with different road inputs can be carried out, and the results averaged.
4.4.2 Input signal

Until now the input signal has been chosen to emulate real road conditions. However, the RIEC method identifies parameters from a frequency modulated form of the suspension’s dynamic response, and we have a reason for band limiting the input well below the expected frequency content of any physical road. As we have moved away from the simulation of typical roads, we might therefore consider a new frequency model for the input signal.

A general spectral model could be optimised according to the bias norms, but as we are also interested in examining the effect of varying integration length $T_k$, a simpler model is favoured for the road. Here we adopt a Gaussian white noise process for road velocity, and restrict the investigation to considerations of bandwidth. The choice of $F_{\text{min}}$ and $F_{\text{max}}$ depends on four possibly conflicting assumptions:

1. The identified model will be required to describe activity of the suspension throughout its dynamic range. The identification should thus be representative of all relevant frequencies.

2. Rig test conditions limit the choice of $F_{\text{min}}$ (see Section 3.1). For the Robson road model, the minimum value for $F_{\text{min}}$ was 0.4Hz; for white noise this limit is reduced to 0.25Hz.

3. Preliminary investigations of higher order modal disturbances indicate that a choice of $F_{\text{max}}$ in excess of 10Hz is undesirable.

4. Accuracy in the identification may be increased by choosing $F_{\text{min}}$ and $F_{\text{max}}$ according to minimisation of the bias norms.

Within these limitations there may be scope for increasing $F_{\text{min}}$ and/or decreasing $F_{\text{max}}$, depending on the relative importance placed on points 1 and 4 above. In Figure 21, the effects of such changes on the norms is considered; in the left series $F_{\text{min}}$ is increased with $F_{\text{max}}$ fixed at 10Hz, and on the right, $F_{\text{max}}$ is decreased keeping $F_{\text{min}}$ at 0.25Hz.

For each simulation the road velocity is obtained by first digitally filtering a white noise signal, then adjusting the gain to maximise the simulated suspension velocity (as described in Section 2.2) within the assumed rig actuator limits on road displacement. The mass and suspension identifications (based on the wheel momentum equation 4.1.2) are both considered for each of two disturbance conditions; the white noise (reference) case is contrasted with the engine vibration case which is known to be bandwidth-sensitive. The results have been normalised according to their value for the default bandwidth road.
Strong similarities are evident between the norms across disturbance conditions. This reinforces the assumption that the disturbance effect of the engine mode is low below 10Hz, and suggests a level of robustness in the use of bias norms to choose these conditions. As $F_{\text{min}}$ increases, the effect is to decrease excitation of the body bounce mode, and this increases the relative wheelhop content. A small improvement in conditioning for the mass identification is registered by a decrease in the $R$ norm, but the opposite trend occurs in the $c$ norm, which makes the effect in worst case expected bias minimal. Similar trends appear for the suspension identification, except that here, the $c$ norm dominates; the obvious conclusion is to retain a setting of $F_{\text{min}} = 0.25\text{Hz}$. Another good reason for doing this is that the bandwidth of the body bounce mode is narrow (see Figure 5 in Chapter 2), so the success of attempts to diminish total body excitation by 'eroding' the modal excitation will be strongly dependent on the true mass, stiffness and damping parameters.

The effect of decreasing $F_{\text{max}}$ is very different between mass and suspension identifications. The trend in the $R$ norm for mass identification is understandable, given the significant reduction in wheel motion that a lower value for $F_{\text{max}}$ causes, but this is now supported by a similar trend in the $c$ norm. In suspension identifications, there is a minimum in the $b$ norm at 6Hz. This suggests there may be some benefit in running separate tests for mass and suspension identifications, each with a different setting for $F_{\text{max}}$. 

Figure 21: The effect of input bandwidth on the normalised bias norms
4.4.3 Integration intervals

Although the RIEC method dictates a uniformly random choice for the start time, \( t_k \) of each integrated sample, the probability statistics of the integration interval \( T_k \) is unrestricted. The expected modal frequencies of the quarter car give some insight into suitable choices, yet the scope remains limitless; the range of \( T_k \) could be varied, or even split into two ranges according to the modal frequencies, and various probability density models could be investigated. Here, we restrict attention to the default model of a single uniform distribution, and a thorough investigation is made of variations over the range \([T_{\text{min}}, T_{\text{max}}]\).

Because the system's modal frequencies are well separated, we might expect significant variations in the quality of identification from values of \( T_{\text{min}} \) of the order 0.01 seconds through to values of \( T_{\text{max}} \) of several seconds. For this reason, a nonlinear scale of values, \( T_{\text{set}} \), is chosen to efficiently cover the range of interest, from 0.01 to just over 2 seconds:

\[
T_{\text{set}} = [10, 20, 30, 40, 50, 50\alpha, 50\alpha^2, 50\alpha^3, \ldots 50\alpha^{15}] \times 10^{-3} \text{ seconds}
\]

where, for the logarithmic progression,

\[
\alpha = 1.28, \text{ and } 50\alpha^{15} = 2028.
\]

Every combination of

\[
\tau_{\text{min}} = \tau_{\text{set}}(i) \quad j > i, \quad i, j \in [1, 20]
\]

has been considered, and the results are illustrated by 3D surface plots and 2D contour plots. Results were obtained for both the white noise assumption (reference case) and engine vibration case; only the reference results are presented here though, as the findings were similar in both cases (as in Section 4.2.2). Figure 22 shows variation of the \( b \) norm for mass identification. Figure 23 shows the same variable for the suspension identification, and Figure 24 illustrates mass and suspension identification fitting errors.

The dominant effect is a consistent high value for the \( b \) norm and fitting errors where the integration interval is restricted by low values in both \( T_{\text{min}} \) and \( T_{\text{max}} \). Intuitively this can be explained by the very small momentum changes that occur over such intervals in the body bounce mode, and this is borne out by the source of these high values being in the \( R \) norm (not illustrated). The \( c \) norm is approximately constant across all combinations except at very high values of \( T_{\text{max}} \) in the suspension identification case, where it dominates the \( b \) norm to create the high worst case bias seen in Figure 23 above about \( j=17 \).
Figure 22: Variation of the $b$ norm with integration interval; mass identification
Figure 23: Variation of the $b$ norm with integration interval; suspension identification
Identification using Randomised Integration

**Mass identification**

![3D graph showing Mass identification with maximum and minimum T values]

**Suspension identification**

![3D graph showing Suspension identification with maximum and minimum T values]

Figure 24: Variation of fitting errors with integration interval
An interesting feature of the suspension identification $b$ norm is the ridge of high values where $T_{\text{min}} = T_{\text{max}}$. The effect is caused by high values in the $R$ norm, suggesting poor conditioning in the regressors, yet no similar characteristic is seen in fitting errors.

Careful consideration of the suspension identification equation 4.1.2 explains this phenomenon. If exactly the same interval $T$ is chosen for every sample, the result is a rank deficiency in the regression matrix, caused by interdependence of the regressors of the $F_i$. Physically, the integrated summation of $F_i$ terms quantifies change in total damping force over time, and this change is equal if all the $F_i$ shift en bloc by any constant. Taking all the samples over similar time intervals thus affects conditioning, without affecting fitting errors; this also explains the reduction in the $b$ norm along the ridge, as consecutive members of $T_{ser}$ have a wider time difference as $i$ (and $j$) increase.

The choice of suitable integration intervals for a general identification is made easier by minima in both the $b$ norms and fitting errors in the mid-range of $j$. A suitable coordinate was found using a normalised sum of the $R$ norm, $c$ norm and fitting error, and this is illustrated by a black circle on each plot, at

$$[T_{\text{min}}, T_{\text{max}}] = [9, 14] = [135, 460] \text{ ms}$$

4.5 RIEC Applied to the Extended Simulation Study

The final choice of operating variables is summarised in Table 5, and for an immediate indication of the capability of the new method, its performance has been tested on the full range of error conditions simulated in Chapter 3.

<table>
<thead>
<tr>
<th>Test length, $L$</th>
<th>30 seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples taken from each test</td>
<td>12000</td>
</tr>
<tr>
<td>Input 'road' process</td>
<td>Gaussian white noise in velocity, bandlimited in the range 0.25-10Hz</td>
</tr>
<tr>
<td>Integration intervals</td>
<td>Random, over a uniform distribution on $[135, 460]$ milli-seconds.</td>
</tr>
</tbody>
</table>

Table 5: Summary of operating variables for suspension identification using the RIEC method
Unfortunately direct comparisons cannot be made for the measurement error scenarios, as the RIEC method relies on a different measurement set to that of direct OLS (DOLS). Results have been generated based on gain, offset, nonlinearity and hysteresis corruption of the measurements critical to RIEC but these are not presented here. Apart from the excellent rejection of offset errors that RIEC affords, the new method is generally no more or less capable than DOLS in rejecting random noise effects or those correlated with a direct measurement.

Table 6 and Figure 25 show results for the three modelling error scenarios, and here the RIEC method performs better than DOLS. For ease of comparison, the results take the same form as those presented in Section 3.4.2; note however that fitting errors are not directly comparable, as the equation set and hence equation errors for RIEC are entirely different to those under DOLS.

For the tyre slip case the new method achieves a similar level of accuracy to that of DOLS, except that a slight improvement can be seen in the estimate of extreme nodal forces for the damper map. Given the discussion of this form of error in Chapter 3, a significant improvement in accuracy was not expected for this data. The influence of engine bounce has been largely rejected by RIEC however. Fitting errors are generally lower for RIEC compared with DOLS, so the mass identification error of 17.9 is high, yet parameter bias is limited to a 4Kg reduction in $\hat{M}_w$. This improvement in the mass identification also has a 'knock on' effect, allowing highly accurate parameter estimates in the suspension identifications.

In summary, the RIEC method provides a high level of accuracy in identification, performing as well as, or better than DOLS for the error scenarios considered. RIEC operates under conditions dictated by the setting of variables such as integration interval length, and it has been shown that these can be 'tuned' to reduce worst case expected bias in the identified parameters. It is reasonable to expect that these variables could be tuned to improve accuracy in other applications of system identification, where data is corrupted by an unmodelled mode.
Table 6: Parameters identified from data with modelling errors (RIEC method)

<table>
<thead>
<tr>
<th></th>
<th>( M_w ) (Kg)</th>
<th>( M_b ) (Kg)</th>
<th>Fitting error (%)</th>
<th>( K_s ) (kN/m)</th>
<th>Fitting error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True parameters</td>
<td>30</td>
<td>350</td>
<td>-</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Tyre slip model</td>
<td>30.2</td>
<td>348.8</td>
<td>5.6</td>
<td>19.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Engine bounce model</td>
<td>25.9</td>
<td>349.5</td>
<td>17.9</td>
<td>19.9</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Figure 25: Damper maps identified from data with modelling errors (RIEC method)
Chapter 5

Case Study: Identification of the Test Vehicle Suspension

In Chapter 4 a new identification methodology has been derived, and simulations seem to confirm its earlier promise. In this chapter, the RIEC method is tested in practice, and a case study identification of the suspension at the front nearside wheelstation of the test vehicle is presented. The four-poster test rig and instrumentation is first described, and this is followed by a summary of calibration procedures and a description of the post-processing operations that have been adopted to maximise data accuracy. Identification results are presented for three fixed settings of the continuously variable damper, and the results are analysed to reveal two significant sources of modelling error. One of these motivates a more detailed investigation of the damper unit, which is carried out in Chapter 6.

5.1 The Four-poster Test Rig and Instrumentation

The vehicle was tested on a four poster hydraulic shake rig as shown in the photographs of Figure 26. The car is supported by circular wheelpans which are mounted on independently driven, high bandwidth hydraulic actuators; these have a dynamic stroke of 250mm. To minimise friction at the tyres, the wheels were liberally greased and during testing the handbrake was left off. To prevent the car from falling off the rig, a minimal constraint set was employed; lateral constraints are attached to the wheelpans at diagonally opposite wheelstations, and a single longitudinal constraint is employed at the rear. Only light contact between the tyre wall and constraints was observed during testing.

As a further safety precaution, the car’s fuel tank was removed, and this allows a clear view of the vehicle underside, in the lower photograph. The prototype damper units are clearly visible; note the solenoid valve assembly which protrudes from the damper casing, and on the nearside unit, cable ties which hold a thermocouple in place to provide an indication of damper temperature variations during testing.

For future application of control on the road, the test vehicle was permanently instrumented at each wheelstation with accelerometers mounted at the top and bottom damper pinch bolts, and a suspension deflection transducer. The deflection measurement is
Case Study: Identification of the Test Vehicle Suspension

Figure 26: The four-poster rig and test instrumentation
not used for system identification however. Instead, suspension and tyre deflections are obtained from absolute displacement measurements of the vehicle body, wheel-hub and wheelpan.

The wheelpan displacement measure was obtained from the rig’s internal transducers, permanently installed within each actuator housing. Absolute displacements of the wheel hub and vehicle body were recorded by sensors attached to an overhead gantry. These ‘string’ type instruments employ a long wire which is wound around a counter-sprung barrel within the sensor casing. As the wire is withdrawn and the barrel rotates, a connected potentiometer records variations in displacement. The wires were attached to the vehicle by a screw at the top of each wheel arch, and by extended brackets attached to each wheel.

To measure vertical force at the tyre, heavy duty load cell devices were installed between each wheelpan and its hydraulic ram. Designed specifically for the purpose, they are rated at 50kN, and use a ‘triad’ of strain gauges to measure vertical forces independently from the bending moments that will result from any eccentric loading from the tyre.

5.2 Instrument Calibrations and Data Post-processing

In Chapter 3 the detrimental effect of measurement errors on identified parameters was clearly demonstrated. The calibration of test instrumentation was thus carried out very carefully, with the transducers in their test positions, operating over representative frequencies and involving the full measurement chain. To assess drift from amplifiers, and other errors caused by changes in ambient conditions, the calibration tests were carried out both before and after testing.

The rig’s internal wheelpan displacement transducer is a highly accurate non-contacting LVDT device, which is regularly calibrated, so this reading was used to calibrate the other displacement transducers. With the vehicle in position on the rig, low-strain extension wires were connected from the wheelpan to the transducer measurement wires at their appropriate operating heights. Then with the car excluded from the measurements, the rig was driven by bandlimited white noise, as for an identification test.

To calibrate the tyre force measurement, the vehicle was replaced by a series of known weights, $M_i$. Driving the actuator with sinusoids inducing maximum accelerations just below ±1g at various frequencies in the range 1-10Hz, the dynamic force on the load cell is described by

$$\lambda F_i = a_w (M_i + M_p)$$
At each fixed frequency, the peak acceleration of the loaded wheelpan $a_w$ was deduced from the wheelpan displacement transducer. $F_i$ is the uncalibrated force measure from the load cell, and by varying the calibration mass $M$, the load cell gain $\lambda$ and wheelpan mass $M_w$ were easily deduced.

For the identification tests, the integrated tyre force is calculated from $F_i$ by

$$\int F_i = \lambda \int F_i - v_w M_p$$

where $v_w$ is the wheelpan velocity, obtained by numerically differentiating the signal from the rig's internal wheelpan displacement sensor. Note that this direct evaluation of integrated force at the tyre avoids the need to double-differentiate the wheelpan displacement sensor.

To complete the data required for RIEC identification, wheel and body vertical velocities are required, yet velocity transducers have not been included in the measurement set. This omission was in part caused by difficulties in obtaining suitable velocity transducers, and in part deliberate; it is possible to obtain a reliable measure of velocity by virtue of the fact that both acceleration and displacement measurements are available. Velocity can be generated either by numerically differentiating the recorded displacement time histories, or by integrating the acceleration records. Here we propose to do both, and generate a 'best fit' velocity measure by a process which we will call kinematic interpolation. The advantage of this approach is that the quality of both of the measurements can be assessed, and the acceleration and velocity records can both be calibrated according to the displacement measure, thereby relating all the kinematic information to a single source of calibration.

The analysis is conducted in the frequency domain; if sampled acceleration time-histories $y(k)$ and displacement $x(k)$ exist for the same point on the vehicle, the following relationship is expected between the Fourier transforms:

$$X(j\omega) \rightarrow H_{xy}(j\omega) \rightarrow Y(j\omega)$$

where

$$E[H_{xy}(j\omega)] = -\omega^2$$
\(X(j\omega)\) and \(Y(j\omega)\) are easily evaluated from the data, using fast fourier transforms (FFT), and standard methods exist for estimating \(H_{xy}(j\omega)\) from these to see how well the expectation holds (for example, in [3]).

The standard frequency response function estimate is

\[
\hat{H}_{xy}(j\omega) = \frac{\hat{G}_{xy}(j\omega)}{\hat{G}_{xx}(\omega)}
\]

evaluated from 'smooth' estimates of the input auto-spectral density function, and the input-output cross spectral density function. Assuming that the data represents a stationary process, these are given from one-sided FFT's:

\[
\begin{align*}
\hat{G}_{xx}(\omega) &= \frac{2}{n_s N\Delta t} \sum_{i=1}^{n_s} |X_i(j\omega)|^2 \\
\hat{G}_{xy}(j\omega) &= \frac{2}{n_s N\Delta t} \sum_{i=1}^{n_s} [X_i^*(j\omega) Y_i(j\omega)]
\end{align*}
\]

where \(X^*_i(j\omega)\) is the complex conjugate transpose of \(X_i(j\omega)\), and \(\Delta t\) is the sampling interval. The smoothness of these estimates comes about by averaging over \(n_s\) different sections of the time-history, each section having length \(N\Delta t\) seconds. The averaged result has a coarser interval between frequencies, but a smoother representation across frequency to the Nyquist, \(\omega_N = \pi/\Delta t\).

The gain and phase relationships between \(x(k)\) and \(y(k)\) can now be estimated as a function of frequency by standard manipulations of the frequency response function estimate. Taking into account the kinematic shift of 5.2.1:

\[
\hat{A}(\omega) = \left|\hat{H}_{xy}(j\omega)\right| / (\omega^2)
\]

and

\[
\hat{\theta}(\omega) = \tan^{-1} \left[ \frac{\text{Im} \{\hat{H}_{xy}(j\omega)\}}{\text{Re} \{\hat{H}_{xy}(j\omega)\}} \right]
\]

where \(\text{Re} \{\cdot\}\) and \(\text{Im} \{\cdot\}\) refer to real and imaginary parts.

It is not appropriate to judge the match between displacements and accelerations from these functions alone however; the transducers have different sensitivities to noise over frequency, so the gain and phase might not be expected to match exactly at very low or
very high frequencies. To describe the match between sensors independently from gain and phase, a coherence function [3] is employed:

$$\gamma_{xy}^2(\omega) = \frac{|\tilde{G}_{xy}(j\omega)|^2}{\tilde{G}_{xx}(\omega) \tilde{G}_{yy}(\omega)}$$

At each frequency, this value can lie between 0 (no correlation) and 1 (perfect correlation), with values less than 1 indicating the presence of measurement noise or nonlinearities in the system relating the two measurements. The latter may exist due to the fact that the two transducers are not measuring from exactly the same point on the vehicle.

![Figure 27: Kinematic match between body mounted sensors](image)

Figure 27 illustrates the coherence, gain and phase between a body mounted accelerometer, and string pot displacement transducer. The coherence is clearly excellent between about 1 and 10Hz, and in this range the gain is also reasonably stable. Small phase errors are apparent though, and the peak value here of around 12° is typical of the sensor pairs tested; at this magnitude the effect on identification is not thought to be significant. Note the rapid degradation of coherence above 10Hz, which is a direct result of the sharply delimited input signal bandwidth.
Typically, high coherence was found between sensor pairs in the range 2 to 10Hz, so the accelerometers were calibrated according to the average gain in that range. The velocity signal is constructed in the frequency domain, as a combination of the correctly calibrated integrated acceleration signal, and the differentiated displacement signal:

\[
V_{lo}(j\omega) = S(j\omega) \times j\omega, \quad \omega < \omega_s
\]

\[
V_{hi}(j\omega) = A(j\omega) / j\omega, \quad \omega \geq \omega_s
\]

\[
V(j\omega) = V_{lo}(j\omega) + V_{hi}(j\omega)
\]

where \( S(j\omega) \), \( A(j\omega) \) and \( V(j\omega) \) are the displacement, acceleration and resultant velocity fourier transforms. \( \omega_s \) is chosen at the midpoint of the range of high coherence, for a 'switch' frequency of 6Hz. The resulting 'hybrid' velocity generally has better noise characteristics than either differentiated displacement or integrated acceleration alone.

5.3 Identification Test Procedure, Results and Discussion

To promote a consistent characterisation of the suspension across tests, a standard procedure was adopted. First the dampers were either excited, using a 10Hz sinusoidal drive signal to the actuators, or allowed to cool, such that they remained within a defined operating range of temperatures (thermocouple readings between 40°C and 60°C were allowed). Before each test, the vehicle was repositioned to centre the tyres on the wheelpans, and data collection was started a few seconds prior to excitation of the rig actuators, to record transducer offset levels. Data was collected from all the sensors at a sampling rate of 204.8Hz, while all four rig actuators were driven by the same bandlimited white noise signal, as dictated in Table 5 of Chapter 4. Four tests were carried out, with each input signal based on a different pseudo-random number sequence.

After reconstruction of the velocity signal, as described above, the identification of mass and suspension parameters was carried out under the RIEC operating conditions prescribed in Table 5.

5.3.1 Mass identification

Figure 28 illustrates the mass identification results in normalised form. For each batch of test data, 20 RIEC identifications have been carried out; the low variance within each test verifies that sufficient samples have been taken. Across the tests at each damper setting the variance is higher, particularly in identified wheel mass. These findings are consistent with the simulation results described in Chapter 4.
By far the largest influence on the results is damper setting. As damping increases, the identified wheel mass decreases, fitting errors become larger, and at 100% damping the wheel mass estimate also has significantly greater variance. Increased damping clearly affects the vehicle dynamics substantially, and it is possible that the errors originate at least in part from generally higher levels of measurement noise induced by increased damper forces. However, the parameter bias is severe, and this suggests that the model is performing badly, unable to describe the ride dynamics at higher damping; from earlier simulations we might suspect that the free bounce motion of the engine is having an effect. This conclusion is consistent with the expectation that engine modes will be excited more as the wheelhop bandwidth increases with damping.

One simple check of this diagnosis of the main source of error, is to consider the mass identification at a rear corner of the vehicle, where engine vibration corruption should be negligible. Table 7 shows identified masses and fitting errors over seven damping levels for a front and rear corner.
The front corner is certainly less consistent and the wheel mass estimates more severely biased, but notice that the same trends occur in the rear corner identification as in the front, only to a lesser extent. It may be considered reasonable to allow mass to vary with damping rate in the model, on the grounds that the system dynamics change so much that the ‘best fit’ quarter vehicle solution should not be restricted. This additional freedom is not favoured though, as it represents a non-physical solution, and would involve a complicated nonlinear function of the masses when modelling the full range of continuous damper levels.

Without further instrumentation of the vehicle, and complication of the model, the reduced accuracy of the identified mass model is unavoidable for the front of the vehicle at high damping. To identify a fixed estimate for the masses, an average value is taken over the three lowest damper settings, where the mass estimates are most consistent and fitting errors are lowest. Thus for the front, case study corner we obtain:

\[ M_w = 28.1 \text{Kg} \]
\[ M_b = 367 \text{Kg} \]

5.3.2 Suspension parameter identification

To identify the suspension parameters, both momentum into the wheel and into the vehicle body have been investigated. The identified damper maps and spring stiffness values are illustrated in Figure 29 along with fitting errors. Suspension velocity breakpoints for the damper map were chosen according to an informal optimisation of fitting errors.

In Chapter 3, the effect of engine bounce error on suspension parameter identification was discussed and the conclusion drawn that identification from body motion was inaccurate compared with identification from the wheel. This conclusion is verified in these results,
Figure 29: Case study identification results: spring stiffness and damper maps
as fitting errors are generally higher in the body momentum case, but the most significant feature of this study is that fitting errors are again dramatically higher at the highest damper setting. Although some error may be caused by imperfect identification of the masses, the effect is unlikely to be principally as a result of engine vibration, as both wheel and body momentum identifications are significantly affected; this is in contrast to the findings of Chapter 3.

The high fitting errors are not associated with any systematic bias in $K_s$, but it is not possible to gauge the bias that may have occurred in the damper map itself. It thus seems sensible to investigate the fitting errors further to isolate any correlations that may exist with the measured system dynamics. We have seen from simulations that plotting fitting errors against the measurements can give clues as to the origin of systematic errors, though this cannot be done using the randomised equation errors of the RIEC method.

By applying the non-integrated form of the suspension identification equation 3.1.3 to the identified parameters, a more suitable suspension force error measure can be derived:

$$F_e = F_t - 30a_w - (23000x_2 + f(x_3 - x_4))$$

where $f(.)$ denotes the averaged 100% damping map indicated in Figure 29, and the mean wheel mass and spring stiffness estimates are written explicitly.

It turns out that $F_e$ is most strongly correlated with suspension acceleration. The link is illustrated by the scatter plot of Figure 30(a) which shows a strong negative correlation with an apparent phase lag between errors and acceleration, and also a weaker positive correlation within the ‘hysteretic loop’. The immediate response might be to arrange some model modification to explain these errors in terms of lagged accelerations, but although this may reduce fitting errors, it would have no physical meaning; in the context of explaining forces, a multiplier of acceleration would represent mass, and clearly there is no significant mass in the suspension which moves in this way.

The cycling trend suggests a form of hysteresis in the system, with forces physically being generated out of phase of those predicted by the current model. Indeed, further investigations reveal that this hysteresis may be associated with force generation within the damper unit. Figure 30(b) shows a scatter plot of damper force $F_d$, estimated as

$$F_d = F_t - 30a_w - 23000x_2$$

against damper velocity for the 100% damping case. This shows a physical representation of damper forces compared with the modeled piecewise linear damper map. The scatter represents suspension errors $F_e$, and its regular form, with distinct groupings of points.
Figure 30: Illustration of errors at the high damping setting
indicates the presence of a consistent unmodelled effect. To reduce suspension force errors, it is necessary to consider the presence of compliance in the damper itself.

In Chapter 6, a more thorough examination of the damper unit is undertaken, using tests carried out on the damper in isolation. These tests are essential to understand the capabilities of the continuously variable unit for control purposes, but they also provide an opportunity for comparing detailed models with those of a more simple structure. In Chapter 7 we will return to the completion of this case study.
Chapter 6

A More Detailed Model of the Controllable Damper

In Chapter 5 we discovered modelling limitations of the quarter vehicle approach which motivate a more thorough examination of force generation in the damper unit. Modern vehicle suspension damping is generated by fluid flow through complex systems of passages connecting reservoirs via different-rated sprung loaded valves, so it is not surprising that a simple force / velocity relationship fails to explain the physical dynamics of the system.

In this chapter, the internal operation of the damper is first examined, and the operation of the continuous solenoid valve is described. Force and velocity measurements from component tests are then used to confirm the findings of Chapter 5, and physical justifications are given for an extension of the damper model to include a series compliance. Although the extended model is shown to be highly effective, large alterations to the original nonlinear force / velocity model are required, and the solution described here does not account for time-transients in the actuation of variable control.

Although the compliant damper model may be critical in the application of closed-loop control, its benefit in improved modelling accuracy must be measured against the increased number of free parameters identified, and the degree of complication incurred in the identification process. This issue is the subject of Chapter 7.

6.1 Damper and Control Valve Design

The prototype, controllable dampers used in the test vehicle are uni-directional flow units, with a three cylinder construction. Their operation is illustrated in schematic form, in Figure 31, and a more detailed cross sectional scale drawing of the solenoid valve is given in Figure 32. Note the exterior placement of the valve, which can also be seen in the photograph of Figure 26 in Chapter 5.

The schematic shows how fluid flow within the unit is arranged such that the flow through the solenoid valve is always in the same direction; this allows a simpler design of the variable valve. To achieve this flow however, it is necessary to equip the main body of the damper with several uni-directional valves, and this invokes a complex flow path for the oil in both directions.
Figure 31: Schematic of the variable damper
Figure 32: Detailed cross-sectional view of the continuously variable solenoid valve
The operation of the solenoid valve is dependent on a fine balance of forces across the pressure plate and also across the solenoid slug. In the default 'hardest' setting (illustrated), with no current acting in the solenoid, the balance of forces in springs $s_2$ and $s_3$ keeps the slug fully away from the solenoid. Note that the force from $s_3$ can be adjusted using the preload screw. Oil pressure forces the pressure plate open against spring $s_1$ to open a flow path through vents $v_1$ and $v_2$. Although oil can pass through to the solenoid side of the pressure plate, a channel through the solenoid slug prevents the slug from being lifted.

As current flows in the solenoid, the slug is gradually lifted, permitting a small flow through vent $v_3$. The size of the orifice in the pressure plate and the gap at the slug nose is chosen such that vent $v_3$ provides a progressive relief of pressure. $v_1$ and $v_2$ are still opened by the oil pressure, although to a progressively reduced extent until the current in the solenoid is at its maximum level, and with the damper at its softest setting, the oil flow is principally directed through $v_3$.

6.2 Component Tests of the Damper

To develop and validate a more accurate model for the damper, component tests were carried out. A hydraulic, ‘INSTRON’ rig of the type generally used for strain testing, was used; the rig is driven by a displacement demand signal, and was instrumented with a load cell and velocity transducer.

As the object of these tests is to develop a force / velocity model that reflects forces generated on the vehicle, the input displacement signal for the rig was taken to be the measured suspension deflection time history of the vehicle on the four-poster rig. Tests of 20 seconds duration were performed, and records of force and velocity were taken at a sampling rate of 200Hz.

Scatter plots of measured force against velocity are given for the extreme hard and soft settings of the damper, in Figure 33. The hardest setting displays a distinct hysteretic effect, which is similar to, yet more pronounced than the damper force and error plot of Figure 30 in Chapter 5. The soft setting shows no noticeable hysteresis however, so these tests are consistent with the earlier diagnosis in the case study identification, that the suspension errors were concentrated in the damper model.

The expectation is that this hysteresis can be modelled by a compliant element in series with the nonlinear force / velocity damper map. We can have some confidence in this assumption, as a number of publications employ this modification in support of more
A More Detailed Model of the Controllable Damper

accurate suspension models (e.g. [19]) and several physical reasons exist for the compliance, given the damper’s design. For example:

- There may be cavitation causing compressibility in the oil as it is forced at high velocity through small orifices in the valves of the damper.
- The internal and/or external damper casings may be subject to radial flexing under the high pressures acting at high damping levels.
- Bushes in the top and bottom mounts of the damper may be contributing a compliant effect.

Additionally, the concept of a stiff spring in series with viscous damping fits well with the observed behaviour. The change in effect between high and low force generation in the unit is in keeping with higher or lower levels of excitation in the spring causing more or less deviation from the expectations of a viscous model. The consistent hysteretic pattern is also in keeping with a force contribution that is out of phase with the principal forcing effect.

Figure 33: Force/velocity characteristic for hard and soft settings of the variable damper
6.3 Compliant Model for the Damper

Figure 34 shows a graphical representation of a compliant damper model.

Under the assumption of no mass in the spring / damper model, the force in the compliance and damping elements is equal:

\[ F_c = F_d \]  

Therefore, in terms of the damper in the vehicle, we can describe the system dynamics as before, employing a damping force \( F_d \) acting on both the wheel and the body. However, this force is now a function of the non-physical variable \( v_d \) rather than the physical suspension velocity, \( V \). For a constant control action on the damper, \( F_d \) can be modelled as before, in terms of a suitable nonlinear function:

\[ F_d = f(v_d) \]  

A model for \( v_d \) can be found by consideration of the spring force. Equations 6.3.2 can be written in terms of compliant element stiffness:

\[ K_c x_c = f(v_d) \]

and by differentiating:

\[ K_c \dot{v}_c = \frac{df}{dv_d} \dot{v}_d \]
Rearranging 6.3.3 and substituting for $v_c$, the 'effective' damper velocity $v_d$ obeys the differential equation

$$v_d = \left(\frac{df}{dv_d}\right)^{-1} K_c (V - v_d)$$

which can be represented by a fifth state in the quarter vehicle model. Dynamically, the relationship between the effective damper velocity and suspension velocity is thus a nonlinear first order lag, with a time constant that varies with damping rate.

If the control variable is included, equation 6.3.2 becomes

$$K_c x_c = f(v_d, c)$$

and equation 6.3.4 becomes

$$v_d = \left(\frac{df}{dv_d}\right)^{-1} \left\{ K_c (V - v_d) - \frac{df}{dc} \frac{\partial f}{\partial \dot{c}} \right\}$$

Then the variation of the damper map in the control direction needs to be defined at all possible effective damper velocities, to quantify $\frac{df}{dc}$, and some model of the damper’s transient response to control demand must be developed, to quantify $\dot{c}$.

In Section 6.4 and 6.5 we will see that identification of the compliant damper model presents some interesting challenges even in the simpler, fixed control form of equation 6.3.4. We will therefore treat the control transient form of the model as a potential further improvement of the model that may be significant in the modelling of closed-loop control, but is beyond the scope of this thesis.

### 6.4 Optimisation Method for Compliant Damper Model Identification

The damper model identification problem is now more complex than before in two respects. Firstly, the kinematic variables $\dot{v}_d$ and $v_d$ are non-measurable, making direct identification impossible. Secondly, the model is dependent on a nonlinear function of the damper map and compliance, and both of these need to be identified.

It follows that conventional identification from linear regressors is impossible, and best estimates for the parameters will be found by a multi-parameter optimisation of fitting errors. Thankfully, this is more easily achieved than it seems, because the processes of damper map identification, and compliant element optimisation can be separated.
The iterative approach we shall consider, involves alternate optimisation of $\hat{K}_c$ and identification of the damper map. Assuming that an estimate of the time history $v_d(t)$ is available, the damper map can be identified by a least squares approach which is equivalent to the RIEC suspension force identification of equation 4.1.2

$$\int_{t}^{t+T} F_d(\tau) \, d\tau = \sum_{i=1}^{S} F_i \int_{t}^{t+T} R_i \{ v_d(\tau) \} \, d\tau$$

6.4.1

The identified damper map model is then used along with an estimated value of $K_c$, to simulate a time history $v_d(t)$ according to equation 6.3.4. Now with the damper map fixed, a single parameter optimisation can be carried out to find the value for $\hat{K}_c$ that generates a sequence $v_d(t)$ which minimises errors in equation 6.4.1. In this way, the $\hat{F}_i$ and $\hat{K}_c$ are alternately optimised under the same cost criteria. The algorithm is summarised by the flowchart of Figure 35.

---

**Figure 35**: Flowchart of the compliance model optimisation algorithm
A test of the algorithm applied to force and velocity measurements taken on the INSTRON rig, with the damper at its hardest setting, is illustrated in Figure 36. In the figure, the scatter plot of measured forces and velocities is shown in purple, and the piecewise linear model which was fitted to this data is shown in black. The second stage, single parameter optimisation of $K_c$ yields a lagged velocity time history, which is plotted against the original measured force values, in red.

Figure 36: Comparison of compliance and original damper models for the hard damper setting

As a result of this first run through of the algorithm, fitting errors have been reduced from 23% to 7%, and the capability of the compliance model is clear. The figure also illustrates a problem with the piecewise linear damper model however. The model includes discontinuities at the breakpoints, where instantaneous damper rate is undefined, and this presents a problem in calculation of $v_d$. Evaluation of $v_d$ is impossible at such a discontinuity, and the consequence of numerical integration of equation 6.3.4 is a noticeable ‘bunching’ of data points around the most severe changes in slope. Repeated identifications of the linear damper map, based on successive optimisations of $K_c$ serve to exacerbate the problem. The result is a ‘stagnation’ of the best fit damper map, with greatest errors occurring around the arbitrarily chosen nodal breakpoints. To avoid the problem, an alternative, continuous form of damper map must be employed, ideally without arbitrary breakpoints.
6.5 Continuous Force / Velocity Model

In the search for a continuous form for the damper model, the main restriction is that in order to identify parameters using the RIEC or direct OLS methods, the map must be linear in the parameters. Various polynomial functions were tried, but these were found to be ineffective, as they required too many terms to accurately describe the exponential nature of decay in extreme forces. Exponential functions were not themselves suitable due to the nonlinear relationship between their describing parameters.

Although they rely on velocity breakpoints, a successful model can be based on natural cubic spline functions (described in [33]); these offer the required flexibility in curvature within a reasonably small parameter set. The structure of the finalised model is symmetric, and the negative velocity half is illustrated in Figure 37.

![Illustration of a continuous damper force map using cubic splines](image)

Figure 37: Illustration of a continuous damper force map using cubic splines

Four nodal points are employed in each half of the map, with a shared node fixed at zero suspension velocity, $v_1$. A discontinuity of slope was permitted across $v_1$, this having a physical grounding in that different valving is employed in compression and extension strokes. Indeed a discontinuous force may also be present, if the unit has significant friction, though this was not found to be the case for these dampers. All of the nine nodal damper forces are freely identified.

The first three sections in each half of the model, coloured green, are described using a natural cubic spline, and the most extreme section is modelled with a straight line segment, coloured red in the figure. It was found that at extreme velocities, where there are fewer test points, the straight segment was more consistently identified than a continuation of the spline model, and the straight line also provides a clearer indication of the implications of extrapolating the model.
The spline model is defined, between the first four nodal velocities, by

\[ F = A(v) F_i + B(v) F_{i+1} + C(v) F''_i + D(v) F''_{i+1} , \quad i = 1, 3 \quad /6.5.1/ \]

where \( F \) is the force at a known velocity \( v \) between nodal velocities \( v_i \) and \( v_{i+1} \). \( F_i \) are the identified nodal forces, and \( F''_i \) are the identified nodal force / velocity curvature values:

\[ F''_i = \left. \frac{d^2 F}{dv^2} \right|_{v = v_i} \]

Functions \( A(v) \) and \( B(v) \) perform the same rôle as the roof functions in Chapter 2. Alone, they describe a piecewise linear model:

\[ A(v) = \frac{v_i + v - v_i - v}{v_i + v - v_i - v} \quad B(v) = A(v) - 1 \]

Functions \( C(v) \) and \( D(v) \) then provide an additive curvature, based on cubic functions of \( v \) which have the property

\[ C(v_i) = C(v_{i+1}) = D(v_i) = D(v_{i+1}) = 0 \]

\[ C(v) = \frac{1}{6} (A^3 - A) (v_i + v - v_i - v)^2 \quad D(v) = \frac{1}{6} (B^3 - B) (v_i + v - v_i - v)^2 \]

To ensure a continuous damping rate \((dF/dv)\) across the nodes, we require that, in terms of two adjacent splines \( F \),

\[ \left. \frac{dF}{dv} \right|_{v_i} (v_i - v_{i-1}, v_i) = \left. \frac{dF}{dv} \right|_{v_i} (v_i, v_{i+1}) \]

and by differentiation of equation 6.5.1, this can be written as an additional constraint equation in the parameters:

\[ \frac{v_i - v_{i-1}}{6} F''_{i-1} + \frac{v_{i+1} - v_i - v_{i-1}}{3} F''_i + \frac{v_{i+1} - v_i}{6} F''_{i+1} = \frac{F_{i+1} - F_i}{v_{i+1} - v_i} - \frac{F_i - F_{i-1}}{v_i - v_{i-1}} \quad /6.5.2/ \]

Boundary values are also required to define curvature at the extreme nodes; adopting a natural cubic spline,

\[ F''_1 = F''_4 = 0 \]

and given the definition of functions \( A(v) \) and \( B(v) \), regressors for the straight line segment are included in the identification by setting

\[ F''_5 = 0 \]
A More Detailed Model of the Controllable Damper

Direct solution for the parameters by OLS can be achieved by generating a node-specific form of the general equation 6.5.1, and conversion to identification by RIEC is easily accommodated; each column of the regressor matrix and the regressand is integrated over the random time intervals, $T$ to form a new set of equations of the form:

$$
\int_{t}^{t+T} F_d(\tau) d\tau = F_1 \int_{t}^{t+T} g_1 \{v(\tau)\} d\tau + F_2 \int_{t}^{t+T} g_2 \{v(\tau)\} d\tau + \ldots
$$

$$
\ldots + F_{2n} \int_{t}^{t+T} h_1 \{v(\tau)\} d\tau + \ldots
$$

where the $g_i$ are node-specific combinations of $A(v)$ and $B(v)$, and $h_i$ are similar combinations of $C(v)$ and $D(v)$.

The constraint is enforced by appending equation 6.5.2, multiplied by a suitable weighting factor, to the regression matrix. By choosing a higher weighting factor, the constraint is met with greater accuracy; here a weighting factor of 10 was found to be adequate.

### 6.6 Compliance Model Identification Results

Using the RIEC map identification algorithm of Section 6.5, and the optimisation algorithm of Section 6.4, combinations of spline damper map and compliance were identified from the force and velocity measurements obtained from tests on the INSTRON stroking rig. The results for five settings of the damper are presented in Table 8. In each case the damper map and compliance were optimised together, so separate estimates were generated for $K_c$ at each damper setting in the ‘Optimisation’ column.

<table>
<thead>
<tr>
<th>Damper setting</th>
<th>No $K_c$</th>
<th>Optimisation</th>
<th>Standardised result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fit %</td>
<td>$K_c$ (kN/m)</td>
<td>Fit %</td>
</tr>
<tr>
<td>100%</td>
<td>22.6</td>
<td>435</td>
<td>6.7</td>
</tr>
<tr>
<td>75%</td>
<td>18.4</td>
<td>413</td>
<td>8.9</td>
</tr>
<tr>
<td>50%</td>
<td>12.7</td>
<td>457</td>
<td>7.5</td>
</tr>
<tr>
<td>25%</td>
<td>10.6</td>
<td>523</td>
<td>8.8</td>
</tr>
<tr>
<td>0%</td>
<td>12.4</td>
<td>10800</td>
<td>12.4</td>
</tr>
</tbody>
</table>

*Table 8: Identified compliance and fitting errors for the damper component test*
The identified compliance is consistent for those tests where damping is harder, and here fitting errors are reduced significantly. However, for the softest damper setting, where the compliance has become physically insignificant, a meaningless, high value for $K_c$ is identified. As damping forces are by definition relatively low for the 0% level, the quality of map identification is effectively independent from $K_c$ here, provided a suitably high stiffness is chosen. Figure 38 illustrates the optimisation process for hard, medium and soft damper settings, and the decreasing significance of compliance is evident.

To complete the identification, a single value for $K_c$ was chosen, based on those tests where fitting error is most improved; here the average value over the three hardest damper settings was chosen. The damper maps were then re-optimised in sympathy with this fixed value, by two or three further iterations of the optimisation algorithm, and the revised fitting performance is given in the 'Standardised result' column of Table 8.

The imposition of a greater compliance in the system has made fitting errors higher at the softer damper settings, but this reduction in the quality of fit is slight compared with the improvement in fit at harder settings.
Figure 38: Compliance identification with cubic-spline damper map
Chapter 7

A Comparison of Model Structures

In this chapter the case study identification of Chapter 5 is completed, using the compliant damper model of Chapter 6. The parametrised model is then validated by carrying out simulations over a typical identification test, and comparing simulated state time histories with those measured on the rig.

Given the relative complexity of the new damper model however, it is of interest to quantify the degree of modelling accuracy that it achieves on the car, compared with a more simple alternative. To this end an RIEC identification of the linear quarter vehicle model is carried out, and although high levels of fitting error and bias occur, the linear model performs well in validation tests. Within the context of fixed control settings of the damper, the comparison of model structures illustrates that marginal increases in model accuracy may be small when compared with the computational expense of nonlinear identification and simulation.

7.1 Identification Results

7.1.1 Completing the case study identification

To complete the case study identification, the compliance optimisation algorithm of Figure 35 in Chapter 6 is modified to include identification of the suspension spring stiffness, by substituting equation 4.1.2 for equation 6.4.1. Employing the continuous damper model of equation 6.5.3, the regression equation takes the form:

\[
\int_{t}^{t+T} F_{1}(\tau) d\tau - M_{w}\Delta x_{3} = K_{s} \int_{t}^{t+T} x_{2}(\tau) d\tau + \ldots
\]

\[
\ldots F_{1} \int_{t}^{t+T} g_{1}\{v_{d}(\tau)\} d\tau + F_{2} \int_{t}^{t+T} g_{2}\{v_{d}(\tau)\} d\tau + \ldots + F_{m-2} \int_{t}^{t+T} h_{1}\{v_{d}(\tau)\} d\tau \ldots
\]

where \( M_{w} \) is assumed to be known \textit{a priori}, via the earlier mass identification.

The inclusion in the parameter set of the main suspension spring presents some problems however, as we can see from the results for \( \hat{K}_{s} \) and \( \hat{K}_{c} \) presented in Table 9. Using the
spline model for the damper, with no compliant element, the result for \( \hat{K}_s \) is similar to that in Chapter 5, but the optimised identification of \( K_c \) induces a significant variation across damper settings in \( \hat{K}_s \). There appears to be a commonality between the function of the compliant element and suspension spring in modelling forces on the car, such that \( \hat{K}_s \) and \( \hat{K}_c \) are not fully independent.

For a physically meaningful interpretation of the model, we require fixed values for \( K_s \) and \( K_c \) over all damping levels. This could be achieved by identifying at all levels simultaneously, with a single regressor for \( K_c \), or perhaps by fixing the estimate for \( K_s \) and re-identifying other parameters in sympathy, as we did for \( K_c \) earlier. However, as it is not clear what a good global estimate for \( K_s \) would be, a more pragmatic approach is applied.

To test the hypothesis of a dependence between \( K_s \) and \( K_c \), a series of identifications were carried out with \( K_c \) fixed at different levels. The results showed that more consistent estimates are obtained for \( K_s \) if a stiffer level is adopted for \( K_c \). This compromises the improvement in fitting errors that \( K_c \) can achieve in the 100% damping case, but allows a fixed estimate to be made more easily for both spring stiffnesses. Setting

\[
K_c = 800 \text{ kN/m}
\]

we obtain the results listed in the ‘Compromised \( K_c \)’ section of Table 9. Both the optimised and compromised results are also illustrated in Figure 39.

<table>
<thead>
<tr>
<th>Damper setting</th>
<th>( K_s ) (kN/m)</th>
<th>( K_c ) (kN/m)</th>
<th>Fit (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>22.8</td>
<td>5538</td>
<td>9.4</td>
</tr>
<tr>
<td>50%</td>
<td>22.6</td>
<td></td>
<td>9.2</td>
</tr>
<tr>
<td>100%</td>
<td>23.0</td>
<td></td>
<td>24.7</td>
</tr>
<tr>
<td>0%</td>
<td>22.7</td>
<td></td>
<td>10.0</td>
</tr>
<tr>
<td>50%</td>
<td>21.8</td>
<td>800</td>
<td>8.8</td>
</tr>
<tr>
<td>100%</td>
<td>19.1</td>
<td></td>
<td>13.7</td>
</tr>
<tr>
<td>0%</td>
<td>5538</td>
<td></td>
<td>55.3</td>
</tr>
<tr>
<td>50%</td>
<td>1204</td>
<td>800</td>
<td>8.9</td>
</tr>
<tr>
<td>100%</td>
<td>396</td>
<td></td>
<td>16.6</td>
</tr>
<tr>
<td>0%</td>
<td>22.6</td>
<td></td>
<td>10.0</td>
</tr>
<tr>
<td>50%</td>
<td>21.6</td>
<td>800</td>
<td>8.9</td>
</tr>
<tr>
<td>100%</td>
<td>21.1</td>
<td></td>
<td>16.6</td>
</tr>
<tr>
<td>0%</td>
<td>800</td>
<td></td>
<td>8.9</td>
</tr>
<tr>
<td>50%</td>
<td>800</td>
<td>800</td>
<td>16.6</td>
</tr>
<tr>
<td>100%</td>
<td>800</td>
<td></td>
<td>16.6</td>
</tr>
</tbody>
</table>

*Table 9: Stiffness and fitting error results for three identification approaches*
A Comparison of Model Structures

Figure 39: Case study identification results: spring stiffness and compliant model damper maps

Under the optimised $K_c$ estimate, fitting errors at the hard damper setting are around 10% lower than in Chapter 5. The compromised $K_c$ still provides a 6% improvement at hard damping, but the results from the vehicle are not as encouraging as the damper component tests promised.

An interesting aspect of the damper component test is that comparisons can be made between damping estimated on and off the vehicle. Figure 40 shows damper maps identified under the compromised $K_c$, and the figure also illustrates the result of the component based study. Apart from a larger difference at high damping in compression, there is an approximately constant damping contribution from other elements in the suspension of around 200N. This is a neat illustration of the self-compensating nature of parameters identified according to a simplified model structure.
Figure 40: Vehicle and component identified damper maps

One effect of the compliant model that can be seen in both Figure 39 and Figure 40 is that the variance has increased between samples from the same test, though this effect is still exceeded by the variance between tests and the bias effects associated with changes in damping. The variance and bias remain reasonably low however; a mean level is assumed for the identified spring stiffness and damper maps, given below and in Table 10.

\[ K_s = 21.8 \text{ kN/m} \]

<table>
<thead>
<tr>
<th>0% level</th>
<th>50% level</th>
<th>100% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>( v_i )</td>
<td>( F_i )</td>
</tr>
<tr>
<td>-0.85</td>
<td>-0.93</td>
<td>-1278</td>
</tr>
<tr>
<td>-0.76</td>
<td>-0.83</td>
<td>-1198</td>
</tr>
<tr>
<td>-0.35</td>
<td>-0.38</td>
<td>-702</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.15</td>
<td>-252</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>0.13</td>
<td>0.15</td>
<td>311</td>
</tr>
<tr>
<td>0.35</td>
<td>0.36</td>
<td>820</td>
</tr>
<tr>
<td>0.76</td>
<td>0.85</td>
<td>1291</td>
</tr>
<tr>
<td>0.86</td>
<td>0.94</td>
<td>1327</td>
</tr>
</tbody>
</table>

Table 10: Average spline damper maps identified under a fixed compliance

a. In this identification, and also in the component identification of Chapter 6, the velocity breakpoint values were chosen according to the magnitude and spread of the collected velocity data. For each half of the map, the \( v_i \) were set as the 35th, 75th, 98th and 99th percentiles of the lagged velocity sequence, \( v_d(t) \).
7.1.2 Identification of the linear quarter car model

The simplest, (linear) form of the quarter vehicle model is that described in Figure 4 of Chapter 2. The mass parameters identified in Chapter 5 remain valid, and here the RIEC suspension force model takes a form based on equation 2.1.4:

\[ F(t) = K_s \int_{t-T}^{t} x(t) \, dt - B_s \int_{t-T}^{t} \{x(t) - x_4(t)\} \, dt \]

Applying this model structure to the four-poster rig data, the linear spring and damper rates \( K_s \) and \( B_s \) are identified in Table 11. The values are averaged over results from all four tests at each damping rate.

<table>
<thead>
<tr>
<th>Damper setting</th>
<th>( K_s ) (kN/m)</th>
<th>( B_s ) (N/m)</th>
<th>Fit (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>17.9</td>
<td>4045</td>
<td>41.1</td>
</tr>
<tr>
<td>50%</td>
<td>21.7</td>
<td>1797</td>
<td>13.9</td>
</tr>
<tr>
<td>0%</td>
<td>22.6</td>
<td>746</td>
<td>11.6</td>
</tr>
</tbody>
</table>

Table 11: Parameters identified for the linear model

Not surprisingly for such a simple model, very high fitting errors are apparent for the high damper setting; the unmodelled linearity and compliance also cause a bias in \( K_s \). At the softer settings, both the identified parameters and fitting errors are similar to those for the compliant damper model.

In accordance with earlier physical justifications for a single spring stiffness, an average stiffness is used in the validation:

\[ K_s = 20.7 \text{ kN/m} \]

7.2 Model Validations and Conclusions

In this section, the accuracy of both models is tested by comparing time histories measured on the four poster rig with simulation results. By considering performance in this way, at both hard and soft damper settings, the significance of parameter errors can be investigated in terms of relevant response variables.

To generate simulated data, a modified form of the differential equations presented in Chapter 2 can be used. As we have not identified parameters for a tyre model, equation 2.1.3 is substituted into equation 2.1.5 and the simulations are carried out using the rig
measured tyre force as an input rather than the ‘road’ velocity. The simulation model differential equations are summarised below

**Nonlinear compliance model**

\[
\begin{align*}
\dot{x}_2 &= x_3 - x_4 \\
\dot{x}_3 &= \{ F_t - \tilde{K}_s x_2 - \tilde{f}(x_5) \} / \dot{M}_w \\
\dot{x}_4 &= \{ \tilde{K}_s x_2 + \tilde{f}(x_5) \} / \dot{M}_b \\
\dot{x}_5 &= \left( \frac{df}{dx_5} \right)^{-1} \tilde{K}_c (x_3 - x_4 - x_5)
\end{align*}
\]

**Linear simplified model**

\[
\begin{align*}
\dot{x}_2 &= x_3 - x_4 \\
\dot{x}_3 &= \{ F_t - \tilde{K}_s x_2 - \tilde{B}_t (x_3 - x_4) \} / \dot{M}_w \\
\dot{x}_4 &= \{ \tilde{K}_s x_2 + \tilde{B}_t (x_3 - x_4) \} / \dot{M}_b
\end{align*}
\]

Tyre force measurements were obtained from typical identification tests, at damper settings of 100% and 0%. The simulated suspension deflection, \( x_3 \) and wheel and body vertical velocities, \( x_4 \) are compared with the measured states over typical sections of the tests, in Figure 41 and Figure 42; note that different time-bases were chosen to illustrate each state.

For an objective measure of performance, an index is defined which relates estimated state \( \hat{x}_i(k) \) to measured state \( x_i(k) \) over each sampling point \( k \). This measure is the coefficient of determination, referred to here as the ‘percentage explanation’ of the data, \( p_i \):

\[
p_i = \left\{ 1 - \frac{\sum_{k=1}^{N} [x_i(k) - \hat{x}_i(k)]^2}{\sum_{k=1}^{N} [x_i(k)]^2} \right\} \times 100 \%
\]

Table 12 gives the values \( p_2 \) to \( p_4 \) over the whole 30 second test for both model variants.

<table>
<thead>
<tr>
<th>Damper setting</th>
<th>state</th>
<th>Linear model</th>
<th>Nonlinear model</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>( x_2 )</td>
<td>82.1</td>
<td>95.7</td>
</tr>
<tr>
<td></td>
<td>( x_3 )</td>
<td>84.7</td>
<td>86.8</td>
</tr>
<tr>
<td></td>
<td>( x_4 )</td>
<td>90.9</td>
<td>91.4</td>
</tr>
<tr>
<td>0%</td>
<td>( x_2 )</td>
<td>96.5</td>
<td>96.9</td>
</tr>
<tr>
<td></td>
<td>( x_3 )</td>
<td>87.7</td>
<td>88.0</td>
</tr>
<tr>
<td></td>
<td>( x_4 )</td>
<td>98.4</td>
<td>98.3</td>
</tr>
</tbody>
</table>

*Table 12: Linear and nonlinear model validation results*
Figure 41: Validation of the identified linear and nonlinear models - soft damping
Figure 42: Validation of the identified linear and nonlinear models - hard damping
The performance is generally excellent and it verifies successful identification from both models. Results are better for the soft damper than for the hard, though this is to be expected, given that increased damping gives rise to higher fitting errors. An interesting feature is the over-estimation of suspension deflection $x_2$, by both models, where the measured deflection approaches the suspension work-space limit of 100mm. The vehicle is fitted with progressive bumpstops which have not been modelled, but are surely having an effect here.

In both cases the wheel velocity is less well explained than the body velocity, although the suspension model is the same in both differential equations. This can easily be attributed to the tenfold difference in masses; errors in the modelled suspension force lead to higher acceleration (and hence, velocity) errors in the wheel than in the vehicle body. Although there may be some bias in the mass estimates, they are likely to have a small effect in comparison.

Perhaps the most important result is the accuracy of the linear model, which almost matches that of the nonlinear model, at the softer setting. Percentage explanation is only significantly reduced in the estimation of suspension deflection on the high damping test.

It should be borne in mind that these simple validations do not fully illustrate the requirements of a model in the application of closed-loop control, where accuracy in the damper model is likely to be more significant. Unfortunately it has not been possible to validate the models further; although tests were carried out on the four-poster rig under transient control conditions, the data was later found to be corrupted. Conclusions can therefore only tentatively be drawn based on these results and those for fixed settings of the damper in Chapter 6. Provided the effect of time-transients is small in the application of control, the compliance damper model might be considered sufficiently accurate for the execution of a semi-active control strategy.
Chapter 8

The Digital Kalman Filter Observer

In Chapters 2 to 7 we have discussed system identification, and established and parametrised a simple model to describe quarter car ride dynamics. One of the principal reasons for deriving a model in this form is to enable the development of a suitable real-time observer to track the system states, given signals from transducers which do not explicitly measure those states. Kalman filter observers are thought to be suitable for this purpose.

The observer relies on two sources of information, both of which are imperfect. Sensor readings include electrical and other measurement noise, and the linear system model, on which the filter is based, suffers from inaccuracies caused by simplifications. Because of this, an essential component in the Kalman filter design is the noise model, which quantifies the expected noise from each source and the correlations that may exist between noise sources. This model enables the filter to provide the most accurate state estimates from the information available; it can be thought of as a balance mechanism, weighting model information against the information from each sensor, on the basis of expected accuracy.

In this chapter, the design of a general discrete time-invariant Kalman filter observer is discussed. The role of the noise model is described and its effect on the filter's *modus operandi* is illustrated. Particular attention is given to the case where force detecting transducers such as accelerometers are included in the sensor set. These attract special attention both in the design and subsequent performance of the filter. The important issues are illustrated in the context of a simple mass, spring and damper model, for which simulated observers are designed and tested.
8.1 Kalman Filter Design

The digital Kalman Filter observer is based on a known, or identified linear model describing the states of the plant, $z(k)$

$$z(k + 1) = A_d z(k) + B_d u(k)$$
$$y(k) = C z(k) + D u(k)$$

where the output and transmission matrix are chosen such that $y(k)$ emulates the available measurements of the plant. In terms of the information available to an observer however, equation 8.1.1 represents an idealisation of the system. In terms of the 'true' states to be observed, $x(k)$, and the recorded measurements $y_s(k)$, a more appropriate description is given by

$$x(k + 1) = A_d x(k) + w(k)$$
$$y_s(k) = C x(k) + v(k)$$

where the system noise $w(k)$ comprises unknown inputs and errors in the system model, and the output noise $v(k)$ comprises measurement noise and errors in the output model. Recalling Figure 1 in Chapter 1, these are the combined influences $\varepsilon_s$, $\varepsilon_a$, and $\varepsilon_m$. Note that for the systems considered here, the input sequence $u(k)$ is assumed unknown.

The Kalman Filter provides an optimal estimate of the states $\hat{x}(k)$, assuming $w(k)$ and $v(k)$ are white noise processes, with known covariance matrices:

$$Q_{xx} = E[w(k) w(k)^T]$$
$$Q_{xy} = E[w(k) v(k)^T]$$
$$Q_{yy} = E[v(k) v(k)^T]$$

A predictive, time-invariant observer can be designed according to the block diagram in Figure 43. The corresponding observer system equation is (see [10])

$$\dot{\hat{x}}(k + 1) = (A_d - KC) \hat{x}(k) + K y_s(k)$$

Here $\hat{x}(k)$ is the estimated state vector at the $k_{th}$ time step, $\hat{y}(k)$ is the corresponding estimate of output, the model estimate of sensors, and $y_s(k)$ is the $k_{th}$ measurement from the sensors. The algorithm operates in two stages; the first applies a correction to $\hat{x}(k)$ by feeding back the error in estimated output, through a suitable gain matrix $K$. The second stage applies the system model to predict the state estimates at the next time step, $\hat{x}(k + 1)$. Here $\hat{e}(k)$ is the error between measured and estimated outputs, commonly
The Digital Kalman Filter Observer

Figure 43: Kalman Filter Block Diagram

referred to as the innovation. The innovations sequence is often the only on-line measure of Kalman filter performance.

The observer gain matrix $K$ is found through calculation of the anticipated state error covariance matrix $P$, by application of the discrete-time Riccati equation:

\[
P = GPG^T - GPC^T [CPC^T + R]^{-1} CPG^T + \bar{Q}
\]

\[
K = (A_dPC^T + Q_{xy})(CPC^T + Q_{yy})^{-1}
\]

\[
G = A_d - Q_{xy}Q_{yy}^{-1}C
\]

\[
\bar{Q} = Q_{xx} - Q_{xy}Q_{yy}^{-1}Q_{xy}^T
\]

These formulae are derived in the Appendix, which also shows that the observer can be simplified by making the assumption that system and output noise are uncorrelated, i.e. that $Q_{xy} = 0$. This assumption is widely accepted as standard in textbooks, e.g. [1], and has previously been adopted in Kalman filter designs for vehicle suspension systems, e.g. [45]. However, careful consideration of the system and output errors shows that there exists a quantifiable correlation, if accelerometers or force transducers are included in the sensor set.

The issue is best discussed in the context of a simple mechanical system, as in Figure 44.
The mass on the left is assumed to be driven by a velocity signal $u$, which is assumed unknown to the observer. The right hand mass is fitted with an accelerometer. Ignoring frictional effects ($F_f = 0$), the system can be described by deflection and velocity states as labelled, with the following state equation

$$\dot{x} = \begin{bmatrix} 0 & -1 \\ \frac{K_s}{M} & -\frac{B}{M} \end{bmatrix} x + w$$ \hfill (8.1.6)

Here $w$ is a function of the unmodelled kinematic, $u$. From the definition of $y_s$ and $x_2$, we can write

$$y_s = x_2 + \epsilon_s$$

where $\epsilon_s$ represents independent measurement noise on the sensor. Thus

$$y_s = \begin{bmatrix} K_s/M & -B/M \end{bmatrix} x + v$$

where

$$v = w_2 - \epsilon_s$$

If $\epsilon_s$ and $w_2$ are uncorrelated, the cross-correlation between $v$ and $w_2$ then follows:

$$E(w_2 v^T) = E(w_2 (w_2 - \epsilon_s)^T) = E(w_2 w_2^T) \neq 0$$
A more general continuous-time form for equation 8.1.2 can therefore be written, whenever force sensing devices are included in the sensor set:

\[
\begin{align*}
\dot{x} &= Ax + w & (8.1.7) \\
y_s &= Cx + Lw - \epsilon_s
\end{align*}
\]

where \( L \) is some simple transfer matrix. After discretisation, equation 8.1.7 becomes

\[
\begin{align*}
x(k+1) &= A_d x(k) + I_d w(k) & (8.1.8) \\
y_s(k) &= Cx(k) + Lw(k) - \epsilon_s(k)
\end{align*}
\]

where \( I_d \) is the discretised form of the identity matrix. Although the process of discretisation has altered the precise correlation between system and output noise, \( Q_{xy} \) clearly remains non-zero.

### 8.2 The Noise Model as a Balance Mechanism

To examine the influence of the matrices \( Q_{xx} \), \( Q_{xy} \), and \( Q_{yy} \) on the designed Kalman filter, the example system of Figure 44 is again used, here with parameters arbitrarily chosen as

\[
M = 10\text{Kg} \quad K_s = 10\text{kN/m} \quad B = 20\text{Ns/m}
\]

By consideration of the model and equations 8.1.6, if modelling errors are assumed to be limited to the unknown input, the system and output noise variables can be written explicitly

\[
w = \begin{bmatrix} u \\ Bu/M \end{bmatrix} = \begin{bmatrix} u \\ 2u \end{bmatrix}
\]

and

\[
v = 2u - \epsilon_s
\]

To simulate ideal noise conditions for the filter design, the input and sensor error sequences are taken to be zero-mean Gaussian white noise processes, uncorrelated with each other:

\[
u = N(0, \sigma^2) \\
\epsilon_s = N(0, \sigma_s^2)
\]
The discrete-time noise covariance matrices can thus be written explicitly, given equations 8.1.8 and 8.1.3

\[ Q_{xx} = l_dQ_cI_d^T \]
\[ Q_{yy} = LQ_cL^T + \sigma_s^2 \]
\[ Q_{xy} = l_dQ_cL^T \]

where
\[ L = [0, 1] \]

and
\[ Q_c = E[ww^T] = \begin{bmatrix} \sigma^2 & 2\sigma^2 \\ 2\sigma^2 & 4\sigma^2 \end{bmatrix} \]

Example Kalman filters can therefore be designed for given choices of \( \sigma \) and \( \sigma_s \) according to equations 8.1.4 and 8.1.5. Here, a discrete-time estimate for the system matrix is used, obtained for a sampling interval of 0.005 seconds, using an exponential discretisation method, [31] as

\[ A_d = \begin{bmatrix} 0.988 & -0.005 \\ 4.954 & 0.978 \end{bmatrix} \]

Consider two cases of high and low sensor noise, fixing \( \sigma = 1 \):

(a) For an arbitrary, high sensor noise amplitude, \( \sigma_s = 50 \), we might expect the filter to rely on the model in preference to sensor information which is highly corrupted by noise. The closed-loop filter system matrix \( (A_d - KC) \) and sensor gain matrix \( K \) in this case, are

\[ K = \begin{bmatrix} 8.56 \times 10^{-5} \\ 4.29 \times 10^{-4} \end{bmatrix}, \quad A_d - KC = \begin{bmatrix} 0.902 & -0.005 \\ 4.525 & 0.979 \end{bmatrix} /8.2.1/ \]

and as expected, the sensor gain matrix is small. Indeed as \( \sigma_s \) increases further, the filter tends towards the system matrix \( A_d \).
(b) For low sensor noise (\(\sigma_s = 0\)), the filter matrices become

\[
K = \begin{bmatrix}
0.0017 \\
0.0091
\end{bmatrix}, \quad A_d - KC = \begin{bmatrix}
-0.6672 & -0.0016 \\
-4.1628 & 0.9959
\end{bmatrix}
\]

By assuming perfect sensor data we might expect the filter to estimate the states entirely from sensor information, ignoring the model. However the strategy of numerical integration and double integration of the acceleration sensor to give \(\dot{x}_2\) and \(\dot{x}_1\) respectively, is simplistic. The model itself includes an approximate kinematic relationship, with associated non-zero modelling error \(w_1\). Although this closed-loop system is less transparent than that of equation 8.2.1, both elements of \(K\) are much larger, and the system clearly relies heavily on sensor information.

Note that the filter design of equation 8.2.2 depends on the cross correlation matrix \(Q_{xy}\). If \(Q_{xy}\) is set to zero, the filter changes to a design which employs lower gains in \(K\):

\[
K = \begin{bmatrix}
0.0009 \\
0.0047
\end{bmatrix}, \quad A_d - KC = \begin{bmatrix}
0.1243 & -0.0032 \\
0.3050 & 0.9870
\end{bmatrix}
\]

Although \(\sigma_s = 0\), this design is constrained by the component of system error in the output equation, which is interpreted as independent sensor noise. The resulting filter thus relies too heavily on a relatively unreliable model. It is clear that the exclusion of cross correlation terms has greatest detrimental effect to observer accuracy when force detecting sensors are used which have low noise characteristics compared with the system noise.

More generally, the cost of an incorrect noise model in terms of observer accuracy, is dependent on the relative disruptive capacity of each noise source. Clearly, if low levels of both measurement and modelling error exists, any choice of \(Q\) matrices will yield a successful observer. Accuracy in the noise model is most vital when one source of error is very much greater than the others, not only between measurements and the model, but also between sensors, and within the model, where kinematic relationships are generally more accurate than modelled dynamics.

Another important effect which is frequently ignored in the practical application of Kalman filters, is that of auto-correlated errors. The statistical assumption of zero auto-correlation is rarely met when modelling physical systems, so the Kalman filter design algorithm described here will usually provide a sub-optimal observer. The accuracy of state prediction under such 'non-ideal' conditions is discussed in the next section.
8.3 Study of the Kalman Filter Under Non-ideal Noise Conditions

From the study of errors in Chapter 3 it is clear that in practice, system and output errors are likely to be auto-correlated. In this section, the simple mass, spring and damper model is again used, this time to illustrate the performance of the Kalman filter both under ideal and auto-correlated noise conditions.

8.3.1 Kalman filter design using nonlinear simulations

Auto-correlated modelling errors can be simulated by the inclusion of a nonlinear frictional force, $F_f$ in the model of Figure 44. Simple Coulomb friction is assumed:

$$F_f = -F, \quad |F| \leq \mu N$$

$$F_f = -\text{sgn}(F) \mu N, \quad |F| > \mu N$$

where $F$ is the force exerted by the spring and damper on the block, and the normal force is constant:

$$N = 10g$$

The linear model on which the Kalman filter is based remains unchanged, but in order to estimate a suitable noise model, and to generate test data, nonlinear simulations of the system are used.

The velocity input for these simulations is taken as a normally distributed pseudo-random number series, with variance $\sigma^2 = 1$ which induces peak deflections between the blocks of approximately ±0.1m. The coefficient of friction is then set at a high value, $\mu = 1.5$, to induce noticeable periods of zero acceleration on the block at this level of input. Transducer noise is again simulated by a Gaussian process. Its variance, $\sigma^2_r$ is matched to the modelling errors perceived by the filter (see below), as we wish to avoid the simplistic case of the filter producing excellent state estimates by over-reliance on sensor information.

The noise covariance matrices are estimated from simulated state and acceleration data. Time histories for $w$ and $v$ are calculated according to the linear model matrices using equations 8.1.2:

$$w = x(k+1) - A_px(k)$$

$$v = y_s(k) - Cx(k)$$
The noise matrices are then estimated by

\[
Q_{xx} = \frac{\{w(k)w(k)^T\}}{4000}
\]
\[
Q_{xy} = \frac{\{w(k)v(k)^T\}}{4000}
\]
\[
Q_{yy} = \frac{\{v(k)v(k)^T\}}{4000}
\]

based on the 4000 sample points of the simulation. A suitable estimate for the transducer noise variance, \(\sigma_t^2\), is chosen by setting the modelling error and transducer error components of \(Q_{yy}\) equal, such that

\[
\sigma_t^2 = \frac{\{w_2^Tv_2\}^T}{N} = \frac{1}{2}Q_{yy}
\]

8.3.2 Testing the filter under auto-correlated and idealised error conditions

Auto-correlated errors are generated by simulating a 'test batch' of state and measurement data from the nonlinear simulation model, with \(\sigma, \sigma_s, \text{ and } \mu\) as above. The 'real-time' performance of the designed filter can then be examined by applying the simulated measurement sequence \(y_s(k)\) to generate state estimates \(\hat{x}(k)\) based on zero initial errors:

\[
\hat{x}(0) = x(0)
\]

In the idealised case, the noise should have zero auto-correlation, yet conform to the expected noise covariance matrices. This is achieved in simulation by using the noise model as a transformation matrix, acting on a white noise input. The desired error sequence is first constructed according to

\[
\begin{bmatrix}
w_1(k) \\
w_2(k) \\
v(k)
\end{bmatrix} = P^2
\begin{bmatrix}
\xi_1(k) \\
\xi_2(k) \\
\xi_3(k)
\end{bmatrix}
\]

where \(P = \begin{bmatrix} Q_{xx} & Q_{xy} \\ Q_{xy}^T & Q_{yy} \end{bmatrix}\) and \(\xi_i(k) = N(0, 1)\), \(i = 1, 3\)

Test time histories for the states and measurements are then reconstructed from a known initial condition \(x(0)\) using the linear model matrices according to equation 8.1.2, with the noise added explicitly. The 'real-time' performance is then examined as above.
8.3.3 Results and conclusions

The filter performance is illustrated by plots of a small section of the simulated and estimated state trajectories in Figure 45 and Figure 46. The ‘percentage explanation’, introduced in Chapter 7, has also been calculated and is included in the figures.

Under the influence of friction (auto-correlated model errors), the performance of the filter is reasonably good; the block velocity is very accurately estimated, with \( p_2 \) only slightly higher under idealised noise conditions. The accuracy with which deflection is estimated is significantly reduced in the auto-correlated case however, with \( \hat{x}_1 \) consistently underestimating the true deflection.

To an extent, the propagation of modelling errors into error in the state estimate is inevitable; in this example the model error appears in both the discretised system model \( A_d \) and the output model \( C \), so no choice of observer gain \( K \) can eliminate the error. This is clear from the filter system equation 8.1.4, written in the form

\[
\dot{x}(k+1) = A_d \dot{x}(k) + K(y_s(k) - C\hat{x}(k))
\]

\[
\begin{align*}
\text{Spring extension } (x_1) \\
\text{Block velocity } (x_2)
\end{align*}
\]
The filter design is nevertheless sub-optimal under the influence of friction, because the disruptive influence of the frictional force has been misrepresented; the design assumption is that the errors occur over a broad bandwidth when in fact they are concentrated at the system dynamic frequencies. It turns out that if the system noise model is inflated, better performance is achieved under the same test conditions. This is illustrated in Table 13, which compares the original filter design with that of an alternative observer, designed under the assumption that modelling errors are 20% greater than for the original. $p_1$ and $p_2$ are both improved under this nominal design adjustment. To illustrate the earlier discussion in Section 8.2, the table also shows the significant effect of setting $Q_{xy} = 0$. 

Figure 46: Kalman filter performance under idealised noise conditions
In this chapter we have shown that Kalman filter observers can provide accurate state estimates for a mechanical system. However, if forces or accelerations are used as output the observer accuracy depends on inclusion of a system / output cross correlation model, and on the accuracy of the output model.

In the next chapter, the application of a Kalman filter observer to a quarter vehicle suspension system is considered. Given the above conclusions, the choice of a suitable sensor set is critical to the success of this observer, and this choice is considered, along with other factors affecting the design, in the context of a formal statistical experiment.

Table 13: Performance comparison of three Kalman filter designs tested under conditions of auto-correlated system noise

<table>
<thead>
<tr>
<th>Design Description</th>
<th>$K_T \times 10^{-4}$</th>
<th>$Q_1$ (%)</th>
<th>$P_2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original design</td>
<td>[2.09, 33.2]</td>
<td>83.3</td>
<td>95.6</td>
</tr>
<tr>
<td>Design with inflated model error expectation</td>
<td>[2.52, 33.7]</td>
<td>84.4</td>
<td>96.2</td>
</tr>
<tr>
<td>Design with inflated model error expectation, and $Q_{xy} = 0$</td>
<td>[2.64, 5.28]</td>
<td>81.7</td>
<td>89.3</td>
</tr>
</tbody>
</table>
Chapter 9

Factorial Analysis of the Kalman Filter

This chapter considers the design of Kalman filter observers for the vehicle suspension system. The main aim of the study is to assess some of the very many options available to the system designer, and to do this, a ‘full factorial’ statistical experiment is carried out. This allows detailed conclusions to be drawn concerning the relative importance of various operational parameters in the Kalman filter design, particularly regarding the type and location of feedback sensors fitted to the vehicle. From the findings of the study, an optimised Kalman filter is designed, which achieves high accuracy in state estimation using the smallest sensor set.

The study is based on two sources of information; from vehicle tests on a four poster hydraulic shake rig, and also from quarter vehicle simulations. Kalman filters are designed from both, using a direct, data based identification method. The advantages of using simulation are that precise values are available for the ‘actual’ vehicle states, and a relatively accurate discrete-time model may be found. The major disadvantage is that noise corruption can only be applied in a simplistic manner. Conversely, the stochastic disturbances from empirical data taken from the vehicle are much more realistic, although the accuracy of both the system model and measured internal suspension states becomes less certain. The objective is thus to arrive at conclusions which are compatible between the two approaches.

9.1 Physical and Simulated System Tests

To obtain data from the test vehicle, a rear corner of the car was instrumented with road-feasible sensors; four are considered in the factorial study, measuring vertical acceleration of the body and wheel, and suspension deflection and velocity. These are denoted $y_1, \ldots, y_4$ respectively. With the car on the test rig, the additional instrumentation described in Chapter 5 was used to record system states $x_1, \ldots, x_4$.

The test conditions were similar to those described in Chapter 5. Tests of 20 seconds duration were carried out for ‘hard’ and ‘soft’ settings of the dampers, with all four units on the vehicle set the same. Band-limited white noise was again used for the ‘road’ input velocity and this was delivered in-phase to all four wheels at an RMS level of 0.27m/s.
The band-pass range for these tests was 0.25-25Hz though, with the higher upper limit imposed to prevent excessive excitation of higher frequency modes in the system, such as body structural resonances. The input and test data were sampled simultaneously at 204.8Hz as for the identification tests.

To acquire simulated data, $x$ and $y$ were calculated using precisely the same input applied to a quarter vehicle model with compliant damper, using the following nominal parameters:

$$
M_b = 250\text{Kg} \quad K_s = 20\text{kN/m}
$$

$$
M_w = 30\text{Kg} \quad K_i = 170\text{kN/m}
$$

To assess the importance of an accurate damper model in the Kalman filter design, the nonlinear spline model of Figure 47 is used; this was identified using the optimisation method described in Chapter 6. Note that this model does not correspond with that identified in Chapter 6 or Chapter 7 however, as this experiment was carried out using a different set of prototype dampers. The 'hard' and 'soft' settings are thus nominal, and consequences of a wider damping range are discussed later, in terms of their effect on closed-loop suspension control (see Chapter 10).

![Figure 47: Damper force map for the factorial study](image-url)
Simulated data-sets were generated by carrying out continuous-time simulations of 20 seconds duration using the model equations summarised in Section 7.2. Driven by the road input velocity $u(t)$, the five state vectors $x(t)$, were evaluated, and the four 'sensor' output vectors $y(t)$ constructed according to their definition above, by

$$
y_1(t) = \dot{x}_3(t) \\
y_2(t) = \dot{x}_4(t) \\
y_3(t) = x_2(t) \\
y_4(t) = x_3(t) - x_4(t)
$$

To assess the significance of noise in the design of the filters, two sets of simulation data were prepared. The first is noise free, and the second incurs system noise $w(t)$ added to the state equations, and output noise $v(t)$ added to the output equations. Both noise processes are defined as broad band Gaussian white noise with RMS nominally related to $u(t)$ as follows. Let

$$\sigma_i = RMS(y_i)$$

denote the RMS output excitation in the noise-free case, $v(t) = 0$, $w(t) = 0$. The RMS values for $v(t)$ were defined simply as

$$RMS(v_i) = \lambda \sigma_i$$

where the noise to signal ratio was chosen, somewhat arbitrarily, as

$$\lambda = 0.05$$

Similarly, system noise RMS values were chosen to satisfy

$$\max_j (RMS(y_j)/\sigma_j) = \lambda$$

under each of the following input conditions

$$w_j(t) \equiv 0, \quad (j \neq i)$$

$$w_i(t) \neq 0$$

where in this case $v(t) \equiv 0$, $u(t) \equiv 0$.

\[\text{† In order to compare four and five state models for both simulated and rig-based data, an equivalent fifth state is also calculated for the rig data set. This is simulated from the model } f(l) \text{ and compliance } K_c \text{ of Figure 47.}\]

using measured values $x_3$ and $x_4$ according to equation 6.3.4,

$$\dot{x}_5 = (df/dx_5)^{-1} K_c (x_3 - x_4 - x_5)$$
9.2 Fitted Model and Kalman Filter Observer

For each set of operational parameters (factors) considered in the factorial experiment, a sample of data \{u(k), x(k), y(k)\} is used to identify a linear discrete-time model for the system, and hence design a Kalman filter. The choice of sensors, the order of the model identified, and other design parameters are determined by these factors, but the procedure described below applies generally.

The sample data is split into two 10 second batches, one of which is used to identify the model and design the filter, the 'design batch', the other 'test batch' being retained for performance testing. The single exception to this rule is in the use of noise-free simulations, where the model is identified on a separate, noise-free batch.

A discrete-time linear model is identified by applying a least-squares procedure to the design batch. The model takes the standard form

\[
\begin{align*}
x(k + 1) &= Ax(k) + Bu(k) \\
y(k) &= Cx(k) + Du(k)
\end{align*}
\]

and matrices \(A\) to \(D\) are found by minimising the error criterion:

\[
J = \sum_k \left\| \begin{pmatrix} x(k + 1) \\ y(k) \end{pmatrix} - S \begin{pmatrix} x(k) \\ u(k) \end{pmatrix} \right\|^2
\]

where \(S\) is the system matrix

\[
S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}
\]

Although the system model is obtained with knowledge of the road input \(u(k)\), the observer must operate without knowledge of this signal, which acts on the Kalman filter as a source of noise. Equation 9.2.1 can therefore be written explicitly, in terms of exact states, outputs and noise terms.

\[
\begin{align*}
x(k + 1) &= Ax(k) + w(k) \\
y(k) &= Cx(k) + v(k)
\end{align*}
\]

where \(w(k)\) and \(v(k)\) describe all sources of system and output noise respectively. Using the identified matrices \(A\) and \(C\), the noise signals \(w(k)\) and \(v(k)\) can be estimated from the design batch, by rearranging equation 9.2.4. Estimates for the covariance matrices of these processes are then made as described in Chapter 8:

\[
\begin{align*}
Q_{xx} &= \{w(k)w(k)^T\}/N \\
Q_{xy} &= \{w(k)v(k)^T\}/N \\
Q_{yy} &= \{v(k)v(k)^T\}/N
\end{align*}
\]
The observer is obtained in one of two ways; either as a predictive estimator, estimating $x(k+1)$ given measured outputs $y(k)$, or as a current-time estimator, estimating $x(k)$ given outputs $y(k)$ (see [10]). In control applications the former is required, since a finite computation time is required to evaluate the control action. Here we consider both designs in order to assess the significance of any additional error incurred in predicting the states over the sampling interval $T$ seconds. Here we have chosen the sampling rate

$$T^{-1} = 204.8 \text{Hz}$$

and this value was used throughout the factorial study.

The predictive filter is designed as in Chapter 7. The current-time estimator is designed in a similar way, except that the prediction step is omitted. The states $\hat{x}_{ci}(k)$ can be written in terms of $\hat{x}_p(k)$ by applying the correction step only:

$$\hat{x}_{ci}(k) = (I - K_0C)\hat{x}_p(k) + K_0y(k)$$

and in this case, the optimal estimator gains are given by

$$K_0 = PCT(CPC^T + Q_{yy})^{-1}$$

Having obtained the appropriate Kalman filter, its 'real-time' operation is examined using the test batch of data. The known output sequence $y(k)$ is used to generate state estimates $\hat{x}(k)$ based on zero initial errors, and the percentage explanation $p_i$ is used as the performance index for each of the $\hat{x}_i$, $i = 1,5$.

Here $p_5$ may be considered especially important, in that it describes the accuracy with which a semi-active control system can apply required control forces.

Note that the order of the discrete-time model is variable in the factorial experiment; where a four state model is adopted, and $\hat{x}_3$ is not explicitly evaluated, it is approximated by

$$\hat{x}_3 = \hat{x}_3 - \hat{x}_4$$
9.3 Factorial Experiment Based on Simulated Responses

Factorial experiments are a well-known technique from applied statistics, for example see [13] or [6]. The basic idea is to measure responses of a system (e.g. growth of crops) under a variety of conditions, described by the 'levels' of a discrete set of factors (e.g. type of seed, use of fertilizer). This is formalised into a statistical model, for example

$$y = \alpha_0 + \alpha_{f_1} + \alpha_{f_2} + \ldots + \alpha_{f_n} + \varepsilon$$

Variables $f_i$ define the various factors, numerically coded in some way. $\varepsilon$ is a random variable representing the effects of un-modelled noise, and $y$ is the measured response. $\alpha_0$ is a mean response level, and the values of the model parameters $\alpha_i$ define the relative effectiveness of each of the factors, provided the $f_i$ are coded in a consistent manner. It is the goal of a factorial experiment to obtain data to estimate the $\alpha_i$ accurately and efficiently, in spite of the effects of random noise.

Five response variables will be considered here, these being the Kalman filter performance indices $p_i$ to $p_5$. The factors will be essentially the same in each case, and these are defined in Table 14. Each factor takes a binary form, and a standard ±1 coding is adopted. The first six factors are control factors; $f_1$ to $f_6$ define the suspension output sensors, and $f_5$ and $f_6$ specify details of the filtering algorithm (see Section 8.1). The remaining factors $f_7$ to $f_9$ are so-called noise factors; for $f_8$ and $f_9$ we would expect the responses to be insensitive to such factors, and they essentially provide a check on the experimental method. Factor $f_7$ defines the damper control setting (restricted to a steady-state value in this study), and it is again essential that state estimation performance be insensitive to this factor. Note that the term 'experiment' is being applied here equally to data obtained from simulation and vehicle measurements.

The simple linear model of equation 9.3.1 is much too restrictive for the state estimation problem, in that it ignores any possible interactions between the factors. For example we would expect the effect of adding a new sensor to depend on which sensors are already fitted, or it might be that the estimation is only insensitive to damper setting if a particular sensor is used. Such interactions are accounted for in the more complex 'linear plus interactions' model [13]:

$$y = \alpha_0 + \sum_i \alpha_i f_i + \sum_{i<j} \alpha_{ij} f_i f_j + \sum_{i<j<k} \alpha_{ijk} f_i f_j f_k + \ldots + \varepsilon$$

Here for example $\alpha_{12}$ represents the interaction between factors $f_1$ and $f_2$, i.e. between wheel and body accelerometers. In the analysis below, it is convenient to denote all model parameters, including interaction terms, by the labels given in Table 14; thus $W$ is used to
<table>
<thead>
<tr>
<th>FACTOR</th>
<th>LABEL</th>
<th>VARIABLE</th>
<th>LEVEL</th>
<th>LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheel Accelerometer</td>
<td>W</td>
<td>$f_1$</td>
<td>Not fitted</td>
<td>Fitted</td>
</tr>
<tr>
<td>Body Accelerometer</td>
<td>B</td>
<td>$f_2$</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>Suspension Deflection Transducer</td>
<td>S</td>
<td>$f_3$</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>Suspension Velocity Transducer</td>
<td>V</td>
<td>$f_4$</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>Prediction or current-time filter applied</td>
<td>P</td>
<td>$f_5$</td>
<td>Current-time</td>
<td>Prediction</td>
</tr>
<tr>
<td>Number of state variables used in on-line model</td>
<td>X</td>
<td>$f_6$</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Damper setting(^{(1)})</td>
<td>$N_1$</td>
<td>$f_7$</td>
<td>Soft</td>
<td>Hard</td>
</tr>
<tr>
<td>Data batch used for performance evaluation</td>
<td>$N_2$</td>
<td>$f_8$</td>
<td>0-10 seconds</td>
<td>10-20 seconds</td>
</tr>
<tr>
<td>Data type used for model identification(^{(2)})</td>
<td>$N_3$</td>
<td>$f_9$</td>
<td>Noise-free</td>
<td>Noise-corrupted</td>
</tr>
</tbody>
</table>

\(^{(1)}\) Refer to Figure 47  
\(^{(2)}\) Applies to simulation study only

*Table 14: Factors chosen for the statistical experiment*
denote $\alpha_i$, termed the ‘main effect’ of the wheel accelerometer, $WB$ denotes $\alpha_{12}$ (two-way interaction) etc. A suitable interpretation of these interaction parameters will be made below.

When dealing with all nine factors, there are $2^9 = 512$ model parameters that need to be estimated from the experiment. These are uniquely determined by carrying out the corresponding ‘$2^9$ full factorial experiment’ wherein all available parameter combinations are tested. Note that a more efficient approach, using ‘fractional factorials’ is also possible, but this was not considered necessary.

To analyse the experimental results, the first step is to solve the system of simultaneous equations obtained by substituting experimental results into equation 9.3.2. The factors $f_i$ can then be interpreted in terms of the relative importance of each fitted parameter $a_i$. We are mainly interested in parameters that are relatively large in magnitude, but the question of statistical significance cannot be ignored. The parameters have well-known statistical properties, and can be analysed using the standard ANOVA method [6], though here we use a more convenient, graphical approach of normal probability plotting.

To understand the basic idea behind normal probability plots, consider the case where all parameters $\alpha$ in equation 9.3.2 are actually zero, so there is only a noise term on the right hand side. This fact is unknown to the experimenter, who estimates parameter values as described. These values will generally be ‘small’ but non-zero, adopting a roughly Gaussian distribution. Plotting fitted parameters against an adjusted rank scale called the normal score leads to an approximately straight-line graph, whose slope is defined by the standard deviation of the ‘background’ noise term $\varepsilon$. In the general case, statistically significant parameters easily show up as deviations from this underlying straight line. An advantage of this graphical analysis is that we may easily apply judgement as to whether the parameters are also significant from an engineering point of view - see the discussion below.

### 9.4 Simulation Study Results

Probability plots for the factorial study conducted on the simulated suspension data are shown in Figure 48. In each case, since the mean $\alpha_0$ is the dominant parameter, this is removed and the remaining parameters are plotted (such a step is allowed within the probability plot method). Of the remaining 511 parameters, most are insignificant - in fact most are indistinguishable from zero. Strictly, any parameters that deviate from the underlying ‘noise’ line are statistically significant, and in (b) there is a fairly clear distinction. On the other hand, Figure 48(a) displays many points that are significant but of small magnitude (of the order $\pm 2\%$) and can be considered insignificant from an operational point of view.
Figure 48: Normal probability plots of factors - simulation experiment
First consider Figure 48(a) - estimation of $x_1$. Clearly $W$ predominates, and in relation to the underlying mean $p_1 = 71\%$, the fitting of a wheel accelerometer increases $p_1$ by 18% to give $p_1 = 89\%$. Due to the convention of assigning ±1 to factor levels, this also means that specifying the removal of this accelerometer reduces $p_1$ by 18%, to 53%. Interactions are also very important; adding a velocity sensor ($V$) increases $p_1$ by approximately 12%, but if $W$ is already fitted, the interaction term $WV = -11\%$ almost completely negates the improvement. In similar fashion, the 6% improvement due to $B$ would be reduced through the $VB$ or $WB$ interactions in the case $V = +1$ or $W = +1$. Clearly sensors $B$ and $S$ are relatively insignificant to $x_1$ estimation.

Another important parameter is $P$, and here its effect is not so complicated in terms of interactions. Compared to the mean filter performance, the use of a predictive filter ($P=+1$) reduces $p_1$ by 9%. This is due to the fact that the prediction step in Figure 3 must take place without any knowledge of the road disturbance. Intuitively, it is reasonable to suppose that $x_1$ would be the most severely affected state, and this is easily verified from plots (b)-(e). The effect of $P$ is actually dependent on sampling rate, and it turns out that
this source of improvement in the accuracy of predictive estimation is the only major benefit of increasing the sampling rate above about 200Hz.

The remaining plots in Figure 48 can be similarly assessed. In particular (c) shows the factorial dependence of $x_3$ estimation to be very similar to that of $x_4$, with the notable absence of $P$ however. $x_2$ estimation (plot b) appears to require one of the sensors $S$, $V$ or $B$, but with little benefit of fitting more than one of these; although direct measurement of suspension displacement ($S$) is favoured, there is little to choose between the three sensors.

In plot (d), the estimation of $x_4$ reveals some problems. In particular, the appearance of the 'noise' term $N_2$ indicates that the results should not be taken at face value. This is considered further below. Plot (e) shows that the estimation of effective damper velocity ($x_5$) is substantially improved by the inclusion of the fifth state in the discrete-time model; $X$ interactions do not really compromise this fact. Another notable feature here is that once any sensor is fitted, the effect of adding any other is very much diminished by interactions. The noise term $N_1$ also plays a part with $XN_1 = -N_1$, suggesting that proper estimation of the fifth state seems to make damper velocity estimation insensitive to damper setting.

Returning to the problem of $x_4$ estimation, it is possible to assess the large effects of $N_1$ and $N_2$ more closely. Referring to Figure 49, the upper plot compares 'actual' state $x_4$ with the estimate $\hat{x}_4$ under the most favourable condition ($N_1 = +1$, $N_2 = +1$); although there are some errors, the estimation appears highly successful.

By contrast, the lower plot is for the least favourable condition ($N_1 = -1$, $N_2 = -1$) and there is a clear problem with errors at low frequencies. This may be traced back to the $z$-plane pole locations of the Kalman filter. A typical case is as follows; choosing $f_i = +1$, ($i=1, \ldots, 9$) the poles are all real, located at

$$Z_i = -0.007, 0.384, 0.699, 0.847, 0.996$$

The last of these is very close to the unit circle, and has a time constant

$$t_p = \left| \frac{1}{T \log |Z_i|} \right|^{-1} = 1.26 \text{ seconds}$$

$9.4.1$
Consideration of the corresponding eigenvector

\[ \mathbf{v} = (0.004, 0.0002, -0.679, -0.734, 0.015)^T \]  

indicates that estimation errors which involve near rigid-body drift of the whole suspension

\[ \hat{x}_3 - x_3 = \hat{x}_4 - x_4 \]

are very slow to decay. This is true of all variants of the Kalman filters designed in this simulation exercise. Thus the excitations in the second data batch \((N_2 = +1)\) just happen to excite this mode more than in the first batch; a low damper setting encourages the problem further.

It is possible to resolve this issue by iterating on the noise processes assumed for the simulation model, in order to drive the offending filter pole away from the unit circle. However, a more pragmatic approach is to simply set \(N_1\) and \(N_2\) to their most favourable conditions \((+1,+1)\) and tentatively draw conclusions regarding the other factors. This has been carried out in Figure 50.
Factorial Analysis of the Kalman Filter

Not surprisingly, the body accelerometer has a major influence, again with interaction terms reducing the effectiveness of adding further sensors.

9.5 Results from the Rig Data Study

The same five response variables are considered for the rig test data, and the results are given in the probability plots of Figure 51(a)-(e). The plots are presented in the same way as in Section 9.4, but here only eight of the factors are considered (factor 9 of Table 14 being redundant).

Comparing Figure 51(a) with the simulation study result of Figure 48(a), we see a broad agreement between the most significant terms $W$, $V$ and their interaction $WV$. However, these terms are all about 3% less significant in the rig test study. Also note the reduction in significance of factor $P$, from 9% in the simulation study to about 4% for the rig test data. These changes are almost certainly due to the noise added to the simulated $x_1$, which is at a higher level than in the rig test study; $p_1$ is thus less sensitive to the critical sensors $W$ and $V$ now, and the predictive step of $P = +1$ incurs a smaller error. Also, we see a 5.5% improvement in the mean value of $p_1$, again attributable to the noise differences.
Factorial Analysis of the Kalman Filter

Figure 51: Normal probability plots of factors - rig experiment
The suspension deflection sensor \( S \), shows an increased significance in Figure 51(a), but as before, interaction terms with \( W \) or \( V \) will greatly diminish this effect. Note also the absence here of factor \( X \), which is discussed below.

The rig test results for estimation of \( x_2 \) and \( x_3 \), in plots (b) and (c) are almost identical to those in the simulation study. Small differences can be seen in the distribution of the points, and there is a slightly higher significance of sensor \( S \) in \( x_3 \) estimations; this is in line with the test rig results for \( p_1 \).

Figure 51(d) shows a very encouraging result for estimation of \( x_4 \). The factors are spread in a similar way to Figure 50, and the only 'noise' factor present is \( N_2 \), with an effect of just \( \pm 2\% \). Examining the slowest filter pole as before, we now find

\[
 t_p = 0.47 \text{ seconds} \quad /9.5.1/
\]

diminishing the influence of low frequencies in the filter. Again the corresponding eigenvector indicates a near rigid-body mode.
Returning to the influence of factor $X$ (four or five state model), consider Figure 51(e). The conclusions of the simulation study for $p_5$ are also valid here, except that the significance of $X$ has dropped dramatically from $\pm 10\%$ to just $\pm 1\%$. This trend is also seen when comparing plots (a) and (d) for the two studies; in the rig test experiment, factor $X$ has disappeared completely. The cause may well lie in the formulation of $x_5$ for the rig test data; these values were simulated from the compliant damper model, and we know from the vehicle identification results of Chapter 7 that the series spring model still represents a simplification of damper dynamics on the vehicle.

### 9.6 Choosing a Minimal Kalman Filter

Hardware costs for a semi-active suspension control system are heavily dependent on sensor requirements, and although the relative cost and reliability of each sensor is not discussed here, the factorial experiment results can be used to select an optimal sensor set, capable of providing an acceptable level of accuracy using the lowest number of sensors. From both the simulation and empirical studies, the following conclusions can be drawn:

- The wheel accelerometer ($W$) has a large influence on the accuracy of estimating $x_1$ and $x_3$. It should therefore be fitted.
- The addition of at least one of the other three sensors is significant in estimations of $x_2$, though it is not critical which one.
- The most significant sensor in estimating $x_4$ is the body accelerometer ($B$).
- If the wheel accelerometer is fitted, the addition of any other sensor is insignificant in estimating the critical variable $x_5$.

It appears that the use of wheel and body accelerometers alone represents a suitable set, with additional sensors adding little to the level of accuracy. The marginal effects of $S$ and $V$ can be quantified, again using the factorial approach, by setting $W$ and $B$ to $+1$; in Table 15 this has been carried out for the rig test data.

<table>
<thead>
<tr>
<th>Benefit of adding sensor(s)</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>0.17</td>
<td>3.37</td>
<td>0.47</td>
<td>0.23</td>
<td>0.75</td>
</tr>
<tr>
<td>$V$</td>
<td>0.08</td>
<td>2.62</td>
<td>0.55</td>
<td>0.20</td>
<td>0.85</td>
</tr>
<tr>
<td>$S$ and $V$</td>
<td>0.07</td>
<td>3.41</td>
<td>0.55</td>
<td>0.21</td>
<td>0.86</td>
</tr>
</tbody>
</table>

*Table 15: Marginal increase in percentage fit from additional sensors: rig test data*
Not surprisingly, the greatest effect is on $p_2$, where $S$ and $V$ provide a direct measure of the variable, $x_2$. Nevertheless, sensors $W$ and $B$ alone provide a percentage fit greater than 90% for $p_2$ in the test shown below, so the marginal 3% does not warrant the addition of $S$ or $V$. The remaining values in the table are within the range of background noise.

Other factors in the final choice of Kalman filter are less critical. Real-time control applications require that a predictive estimate is made of each state ($P = +1$), though we know this will reduce the accuracy of $x_1$ estimation. The influence of factor $X$ is not clear, due to the nature of $x_5$ simulation; the cost in processing time, of including the fifth state is negligible though, so the five state model is chosen. Finally, the ‘noise’ terms $N_1$, $N_2$ and $N_3$ are set arbitrarily, to +1.

The minimal Kalman filter is summarised in Table 16. Its kinematic performance is illustrated in Figure 52, where time-histories for the last two seconds of each ten second batch of rig data is shown, as a typical sample.

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>Wheel accelerometer</th>
<th>Body accelerometer</th>
<th>Suspension deflection Transducer</th>
<th>Suspension velocity Transducer</th>
</tr>
</thead>
<tbody>
<tr>
<td>LABEL</td>
<td>$W$</td>
<td>$B$</td>
<td>$S$</td>
<td>$V$</td>
</tr>
<tr>
<td>LEVEL</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>Fitted</td>
<td>Fitted</td>
<td>Not fitted</td>
<td>Not fitted</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>Prediction or current-time</th>
<th>Number of state variables</th>
<th>Damper setting</th>
<th>Data batch used for evaluation</th>
<th>Data type used for model identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>LABEL</td>
<td>$P$</td>
<td>$X$</td>
<td>$N_1$</td>
<td>$N_2$</td>
<td>$N_3$</td>
</tr>
<tr>
<td>LEVEL</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>Prediction</td>
<td>4</td>
<td>Hard</td>
<td>10-20 seconds</td>
<td>Noise corrupted</td>
</tr>
</tbody>
</table>

*Table 16: Summary of factors for the minimal Kalman filter*
Factorial Analysis of the Kalman Filter

Figure 52: ‘Real-time’ performance of the minimal Kalman filter
9.7 Discussion and Summary

It is worth emphasising that the Kalman filter state estimates in Figure 52 were obtained from sampled accelerometer measurements taken on an essentially standard vehicle, excited at relatively high amplitude on a four poster hydraulic rig. The simple conclusion is that state estimation is very effective, given that only two accelerometers are being used.

In Figure 52, \( x_1 \) and \( x_4 \) estimation errors occur mainly at low frequencies. The problem can again be attributed to the slow filter pole in the observer, causing rigid body drift; careful inspection of \( \hat{x}_3 \) shows that similar low frequency errors occur here, though these are less obvious due to the generally higher levels of excitation in the wheel-hub.

As similar errors occur in \( \hat{x}_2 \), it may be thought that low frequency drift is a fundamental problem with an ‘accelerometer only’ sensor set, but this need not be so. Assuming that a reasonably accurate spring / damper model is available within the Kalman filter, any steady-state error in \( x_2 \) leads to a discrepancy in spring-force, and hence in acceleration. This discrepancy should be fed back within the filter to correct the \( x_2 \) estimates. Hence the application of a dynamic model in real-time avoids the worst excesses of low frequency drift that inevitably arise from a simplistic ‘double integration’ of accelerometer signals.

Conversely, the use of dynamic models in real-time may act as a limitation to the Kalman filter. In practice, any vehicle is subject to variations in a number of characteristics, notably sprung mass and vertical tyre stiffness; it may be that these parameters will need to be estimated on-line in practice. On a more positive note, suspension nonlinearity has not seriously degraded the results, suggesting that the effectiveness of the Kalman filter is not particularly sensitive to the somewhat idealised assumptions implicit in the design method. This is in keeping with the encouraging validation result for the linear model identified in Chapter 7.

The minimal Kalman filter clearly performs well in terms of the conditions of this experiment, where separate observers have been designed for each fixed damping level. The critical test of an observer lies in its interaction with an appropriate controller however, under closed-loop operation; the most suitable criterion of observer success is the extent to which it affects controller performance compared with the use of perfect state knowledge. In Chapter 9 this issue is discussed, and the minimal Kalman filter is tested in simulation along with a suitable semi-active suspension control law.
Chapter 10

The Effect of Observer Error on Closed-loop Control

Here the performance of the Kalman filter observer is assessed in terms of its efficiency in the application of closed-loop control. Using the observer design for the minimal Kalman filter discussed in Chapter 8, a simulation study is carried out. Although simplistic noise conditions are assumed, this enables the closed-loop filter system to be accurately judged in terms of an equivalent system having full state feedback.

The performance of active and semi-active systems is considered, using two very different road input conditions, and the results motivate a further analysis of low frequency errors in the filter. It is demonstrated that the dynamic influence of filter errors is dependent on the control actuation used, and also on the magnitude of low frequencies that are present in the disturbing road profile.

10.1 Applying Semi-Active Control

To apply semi-active control using the continuously variable damper discussed in Chapter 6, we require a feedback law to evaluate the damper control \( c(t) \) in terms of the states \( \ddot{x}(t) \). These may be perfectly known, in the case of full state feedback \( \ddot{x} = x \), or they may be estimated, \( \ddot{x} = \hat{x} \) for Kalman filter feedback.

A standard method of controller design is to use a ‘clipped’ active control law, as in [44]. This takes advantage of the linear form of active suspension systems, where the spring and damper assembly are replaced by a force generator. The optimal suspension force \( F_s^* \) is then suitably described by a linear feedback of the four principal states:

\[
F_s^* = k_1 \ddot{x}_1 + k_2 \ddot{x}_2 + k_3 \ddot{x}_3 + k_4 \ddot{x}_4
\]

To apply this algorithm in the semi-active sense, the optimal damper force is first obtained, by subtracting the spring force:

\[
F_d^* = k_1 \ddot{x}_1 + (k_2 - K_s) \ddot{x}_2 + k_3 \ddot{x}_3 + k_4 \ddot{x}_4
\]

and the appropriate control \( c \) is then deduced by inverting the damper map, given \( \ddot{x}_s \).
For the simulations carried out here, the model parameters identified in Chapter 7 are used. The optimised damper map for all five of the fixed damper settings that were considered is recalled in Figure 53.

If \( F_d^* \) lies within the damping force envelope at \( \ddot{x}_s \), \( c \) lies between 0 and 100, and is deduced by linear interpolation. Where \( F_d^* \) lies outside the achievable range, the control is 'clipped' to 0 or 100 as appropriate. The damper force that is then applied by each control is dependent on the true damper velocity,

\[
F_d = f(x_s, c)
\]

It has been shown [12] that this clipped inversion of the damper map provides performance comparable with a more formal nonlinear optimal semi-active controller design.

There is an implicit assumption made here, that any force within the damping envelope can be delivered instantly, i.e. that the effect of control transients is negligible. This is reasonable provided that the damper control solenoid operates significantly faster than the fastest system dynamic mode. Here we have a 12Hz wheelhop resonance, with a period of approximately 80 milliseconds. Simple switch tests that were carried out on the test
dampers suggest that transient times of 20-25 milliseconds are achievable. The assumption therefore appears reasonable.

The active control algorithm that is adopted here employs LQG optimal control [1], where the gain vector $k$ in equation 10.1.1 is found by solution of the algebraic Riccati equation, to minimise the time integral of the following cost function:

$$ J\{x(t)\} = \alpha_1 x_1^2 + \alpha_2 x_2^2 + x_3^2 $$  \hspace{1cm} (10.1.3)

The body acceleration term costs passenger discomfort, tyre deflection is included as an indication of road-holding, and these two are optimised within the constraint of suspension deflection workspace. The weightings $\alpha_1$ and $\alpha_2$ are chosen to tune the controller, and here these are set according to a study by Gordon and Best [12]:

$$ \alpha_1 = 116000, \alpha_2 = 1190 $$

which, applied to the present model yields the feedback gain vector:

$$ k = \{-20032, 8624, 1071, -2195\} $$

10.2 Closed-loop Simulation Study

10.2.1 Kalman filter design and test conditions

The observer is designed as in Section 9.2, using a batch of data simulated over a band-limited white noise input to identify the linear system model and noise matrices. The simulation is conducted using the variable damping envelope of Figure 53, and closed-loop controls are applied using full state feedback. Each control is held constant within the data sampling period of 0.005 seconds; this emulates the control action that is applied when digital Kalman filter state estimates are used for the feedback.

As the control action is known, the observer could be designed with damping force as an input, estimated from the damping envelope and $\dot{x}_3$. However, as we wish to consider the simplest possible observer, no control input is used, and the effect of damper control on the Kalman filter design is merely to increase $w$.

Dynamic performance of the quarter vehicle model was assessed using three system variants:

$S1$ active system assuming perfect suspension force actuation, and employing full state feedback.

$S2$ semi-active system, with full state feedback.

$S3$ semi-active system, with feedback via Kalman filter state estimates.
Each system was tested using two types of road input. 'Road 1' is a typical 20 second 'test batch' sample of band limited white noise in velocity having RMS 0.27 m/s. This has the same spectral properties as the sample used to design the filter. 'Road 2' is a 400m section of measured road profile from a B class road in Leicestershire, traversed at 20 m/s, giving an RMS vertical velocity of 0.38 m/s. Simulated output noise was added to the acceleration trajectories in both cases, at the same level as that used for Kalman filter design.

**10.2.2 Simulation results**

In view of the cost criteria of the controller, the performance of each system is measured in terms of $x_1, x_2, x_4$. RMS values are given in Table 17, and time histories for the final two seconds of each test are also given, in Figure 54 and Figure 55.

<table>
<thead>
<tr>
<th>System</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x_1$ (mm)</td>
</tr>
<tr>
<td>Road 1</td>
<td>S1</td>
</tr>
<tr>
<td></td>
<td>S2</td>
</tr>
<tr>
<td></td>
<td>S3</td>
</tr>
<tr>
<td>Road 2</td>
<td>S1</td>
</tr>
<tr>
<td></td>
<td>S2</td>
</tr>
<tr>
<td></td>
<td>S3</td>
</tr>
</tbody>
</table>

*Table 17: RMS performance of simulated closed-loop control systems*

As we would expect from earlier research on suspension control systems, (e.g. [40]) the active system S1 provides the best performance in terms of passenger comfort for Road 1, having low body accelerations, through the use of higher levels of suspension and tyre deflection. S1 also achieves significantly lower accelerations than the semi-active systems for Road 2, although here the suspension workspace usage is considerably greater, due to higher levels of low frequencies in the road profile.

The most notable result is the close match in performance between systems S2 and S3, which shows that the use of a Kalman filter causes only a slight reduction in performance compared with full state feedback. The performance reduction is certainly small compared with the difference in performance between S1 and S2. For Road 1 the performance of S3 may be attributed to the fact that test conditions were chosen to be exactly the same as the filter design conditions. However, the performance of S3 on Road 2 seems to show robustness of the filter to very different input conditions.
Figure 54: Dynamic performance under closed-loop control; Road 1
The Effect of Observer Error on Closed-loop Control

Figure 55: Dynamic performance under closed-loop control; Road 2
In both cases, the differences between $S2$ and $S3$ occur at low frequencies. This is in keeping with the findings of the kinematic study, although the dynamic effect of feeding back kinematic errors is only slight. The performance of $S3$ on Road 2 is remarkable, as it is largely unaffected by a significant drop in estimation accuracy of the filter, as shown by the kinematic performance figures in Table 18.

<table>
<thead>
<tr>
<th></th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road 1</td>
<td>86.54</td>
<td>96.53</td>
<td>97.68</td>
<td>96.11</td>
<td>94.51</td>
</tr>
<tr>
<td>Road 2</td>
<td>47.32</td>
<td>61.23</td>
<td>71.26</td>
<td>80.02</td>
<td>68.04</td>
</tr>
</tbody>
</table>

*Table 18: Kinematic performance of the Kalman filter in closed-loop control*

The poor kinematic results for Road 2 are once again due to low frequency inaccuracies. Here, errors in $\hat{x}_3$ and $\hat{x}_4$ are again caused by a rigid body drift effect, though at a higher magnitude than seen in Chapter 9; the errors correlate with a high level of excitation around 0.3Hz in the road input.

The source of errors in $\hat{x}_2$ is illustrated in Figure 56. Here, the closed-loop observer / controller system appears to have invoked a gain error which we have not seen before. It turns out that this is also the principal contributor to the errors in estimation of $x_5$. A careful review of the results of Chapter 9 (in Figure 52) shows that this effect may have been present in the earlier study, although in that case it is difficult to differentiate gain errors from drift.

*Figure 56: Errors in the estimation of suspension deflection, $x_2$*
10.3 Assessment of Low Frequency Errors

The simulation results seem to show that estimation errors at low frequencies are of no concern for the closed-loop system, yet this is counter-intuitive considering the control algorithm. It is clear from equation 10.1.1, that low frequency errors in the states will be passed on directly to the estimated optimal suspension force, and if this force were consistently delivered by the dampers, the suspension system would steadily be driven away from optimal regulation.

It is thought that the performance of System S3 on Road 2 does not suffer significantly from low frequency errors, because the actuation of control does not allow consistent delivery of the feedback force. This point can be illustrated by comparing optimal and achieved damping forces for system S3 on Road 2. The upper plot of Figure 57 shows the 'optimal' damper force, based on estimated states ($\hat{F}_d^*$) compared with the 'true optimal' damper force, based on true states ($F_d^*$); the effect of drift in the states estimates is considerable.

In the lower plot, the true optimal damper force is repeated, and this is compared with the force achieved by system S3, in attempting to track $\hat{F}_d^*$, after the clipping effect of the damper map. Also plotted is the equivalent clipped force that could have been achieved if perfect state knowledge were available.

Note that for much of the time, neither achieves a close approximation to the true optimal, due to the dampers limitations. For short sections, e.g. between 7.7 and 7.9 seconds, the optimal force can be achieved under both true and estimated states, but clearly the dynamic effect of drift over these periods will be small. Presumably it would take a large and consistent error to make a significant difference to performance under semi-active actuation. The same would not be true if an active system is used.

Until now we have interpreted poor low frequency tracking of the states in terms of drift, corresponding with the slowest filter pole, and related to the filter's lack of dynamic information about the vehicle's global velocity. At very low frequencies however, these errors may be manifest as a consistent under-estimation of the true states; the filter does not 'recognise' large scale geographical features such as hills, because they have no effect on the suspension dynamics. The underestimation of suspension deflection, $\xi_2$ in Figure 56 bears this out, as the errors are concentrated at 0.3Hz where the road amplitude is high, yet the system response is low. In terms of applying closed-loop control, this can actually have a beneficial effect.

† Note that the trajectory of instantaneous forces $F_d^*$ is that generated by applying equation 10.1.2 to the trajectory of instantaneous true states for system S3.
Figure 57: Comparison of optimal and achieved damping under semi-active control
Consider for example, a suspension system with active feedback control traversing a small hill. The hill can be modelled by a raised cosine function, and choosing a maximum gradient of 10%, the height is obtained as 6.4m over a horizontal distance of 400m. The road surface is made more realistic by superimposing a band-limited white noise velocity (Road 1) on this model.

When the active system $S_1$ is simulated on the hill, the low frequency vertical velocities are fed back, resulting in extreme suspension compression, as illustrated in Figure 58. The effect is similar to that seen on cars in children's cartoons, where the car body is driven along at a steady altitude, while the wheels negotiate huge undulations, on telescopic legs. In reality of course, the suspension would be compressed onto its bumpstops. When Kalman filter estimated states are substituted in the active feedback, the low frequencies are under-estimated, and the system performs more realistically as a result. Note that the filter estimates are still subject to drift however, as shown by the positive trend in suspension deflection for the estimated feedback system.

![Figure 58: Comparison of active systems over a simulated hill](image)
Chapter 11

Observer Design for the Moving Vehicle

Until now the discussion of Kalman Filters has been limited to theoretical concepts, and applications using rig-based and simulated data. In this chapter a design methodology is elaborated for observers to operate on the test vehicle, and these are assessed on the road. This procedure represents a practical solution to the vehicle Kalman filter design problem; it is not intended as a rigorous methodology with proven optimality.

In contrast to previous designs, this observer must take into account differences in the model and noise characteristics that occur when the tyre is rotating, the engine is running and full vehicle motion occurs. There are also problems to overcome in obtaining reliable test data with which to compare observer state estimates.

After considering the sources of data which are available to design the observer, it becomes clear that no single source provides all the information required. The design procedure is thus based on a combination of simulation results and measurements from the vehicle. The filter is tested, in the absence of additional instrumentation on the vehicle, by comparing observer state estimates with 'true' states reconstructed from the available acceleration and deflection measurements.

11.1 Available Data Sources

Three sources of information are available for the system and noise models, and also for testing the resultant observer. Table 19 summarises their advantages and disadvantages.

Considering first the design of discrete-time model matrices $A_d$ and $C$, the simple least-squares identification approach used in Chapter 9 has advantages in its simplicity and probably also in accuracy for the particular design and test conditions of the rig. Notes D₁ and D₂ make rig-based models impractical for the road-going observer though. A more general model based on road data would be suitable, but the problems of notes D₃-D₅ are insurmountable within this study. Another problem is the inflexibility of a directly identified discrete-time model; the matrix elements are non-physical, so for example, changes in the sprung mass due to different loading conditions cannot easily be introduced to the observer design. The continuous-time simulation model represents a more robust
### Table 19: Advantages and disadvantages of three sources of information for design and testing of a Kalman filter for the moving vehicle

<table>
<thead>
<tr>
<th>Rig Measured Data</th>
<th>Road Measured Data</th>
<th>Simulations based on identified parameters</th>
</tr>
</thead>
</table>
| **A₁** Accurate measurements, taken in a controlled environment | **D₁** Non-rotating tyre. Excludes effect of vibrations caused by rotation of transmission | **D₆** No rolling tyre model.  
**D₇** High frequency and full vehicle effects not modelled.  
**D₈** Additional data needed for estimating noise models |
| **A₂** Varied road conditions can be simulated (within the bandlimits of the rig actuators). Engine can be run in idle | **D₂** High frequency and very low frequency events not represented |  |
| **A₃** Tests are carried out in the required operating environment, so all relevant dynamic effects are implicitly included.  
These include aerodynamic forces, driveline torques and vibration, electrical interference. | **Difficulties in obtaining accurate state measurements:**  
**D₃** Additional sensors (e.g. laser tracking devices) would be very expensive.  
**D₄** Post-processing from the current sensor set requires integration, giving low frequency errors.  
**D₅** For tests to be representative of general driving conditions, several speeds and road types need to be included. Test conditions may vary. |  |
| **A₄** Accurately identified suspension model is available.  
**A₅** Cheap and convenient test environment.  
**A₆** Principal noise effect of the unknown road input can be quantified.  
Linearisation errors can be assessed.  
**A₇** Measured road profiles are available. |  |  |
basis for the observer, based on physically interpreted parameters with proven accuracy. The lack of a tyre model is not even a significant problem, as a simple tyre spring stiffness can easily be identified from road tests. The principal problem is that of $D_8$.

Unfortunately there is no obvious way of estimating the noise statistics without state measurements from the vehicle on the road. It appears that the most suitable approach is to restrict the estimate of system errors to those effects that can be simulated, and find an alternative vehicle-based source for estimating output errors. It turns out that $A_7$ is a critical factor here, enabling a link between simulated and measured accelerations from which the level of output errors can be estimated. Although $D_5$ suggests that many different road conditions would be required to define a system model, it is thought reasonable to use a single road for this nominal estimate.

### 11.2 A Kalman Filter Design Algorithm for the Moving Vehicle

The identified continuous-time quarter car model of the test vehicle is used as a basis for the Kalman filter. Assuming that a suitable estimate is made for rolling tyre stiffness, a linearised form for this model can be derived. The noise model can then be formulated from the continuous-time representation of true states and sensor measurements given in equation 8.1.7

\[
\dot{x} = Ax + (I)w \\
y_s = Cx + Lw - \varepsilon_s
\]

where, for the minimal sensor set:

\[
L = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

The discretisation interval, and hence prediction time-step of the filter is minimised, within the limitations of observer and controller execution times for the hardware used on the test vehicle. The findings of Chapter 9 suggest that accuracy in $\hat{x}_1$ will be improved under the sampling frequency of 500Hz that is achievable.

Recalling the discrete-time form for equation 11.2.1,

\[
x(k+1) = A_dx(k) + l_dw(k) \\
y_s(k) = Cx(k) + Lw(k) - \varepsilon_s(k)
\]
and using the equations

\[ E[\mathbf{w}(k)\mathbf{w}(k)^T] = Q_c \]
\[ E[\mathbf{e}(k)\mathbf{e}(k)^T] = Q_s \]
\[ E[\mathbf{w}(k)\mathbf{e}(k)^T] = 0 \]

the discrete-time noise matrices required for observer design are

\[ Q_{xx} = E[(I_d\mathbf{w}(k))(I_d\mathbf{w}(k))^T] = I_dQ_cI_d^T \]
\[ Q_{xy} = E[(Lw(k) + \mathbf{e}(k))(Lw(k) + \mathbf{e}(k))^T] = LQ_cL^T + Q_s \quad /11.2.3/ \]
\[ Q_{xy} = E[(I_d\mathbf{w}(k))(Lw(k) + \mathbf{e}(k))^T] = I_dQ_cL^T \]

Assuming that the principal disturbing effect arises from the unknown road, \( w(k) \) and hence \( Q_c \) may be estimated from a simulated time series for the states. \( \mathbf{x} \) and \( \mathbf{x} \) are generated using a five-state nonlinear simulation over the 'design' road profile (see Section 11.2.1), and a system noise time history is generated using

\[ w(k) = \dot{x}(k) - Ax(k) \]

For the output noise estimate, auto-spectral density plots for sensor measurements taken from the vehicle on the road are compared with those for simulations of the sensors over the same road profile. A comparison is therefore made between

\[ G\{y_s(k)\} \quad \text{and} \quad G\{Cx(k) + Lw(k)\} \]

where \( G(\cdot) \) represents the smoothed estimate of the auto-spectral density function described in Chapter 5.

This graphical procedure provides a nominal estimate for the variance of \( \mathbf{e} \) and hence the principal elements of \( Q_s \). The comparisons also provide an indication of the accuracy of the system noise model. Note that although the sequence \( w(k) \) is the best available estimate of system noise, it is compromised in that it ignores unmodelled full vehicle effects such as roll and pitch, and also the effects of engine bounce and other low frequency modes of the vehicle on the road. These are reflected in differences between the simulated and measured sensor characteristics in the system's dynamic frequency range.
The design algorithm is summarised below, and details of the procedure are given in subsections 11.1.1 to 11.1.5.

1) A suitable road is chosen for the provision of acceleration data from the vehicle. A measured profile of this road should be available to provide an input for comparative simulation tests.

2) A nonlinear simulation model is derived, based on parameters identified from the four-poster rig study. To simulate rolling tyre dynamics, a linear tyre stiffness is estimated from a separate vehicle test.

3) A nonlinear simulation is carried out over the test road to yield $\dot{x}$, $x$ and $y$. A vehicle test is carried out on the same section of road, to yield $y_v$. In both cases, closed-loop control of the damper is applied.

4) A linear form for the continuous-time model matrices $A$ and $C$ is defined, based on the identified nonlinear model.

5) $w_k$ is evaluated from the simulation results and linear model, and $Q_c$ is estimated. The elements of $Q$, are estimated directly by comparing simulated $y$ with measured $y_v$ in the frequency domain.

6) The noise matrices are converted to a discretised form using equation 11.2.3, and the steady state Kalman filter is constructed according to equations 8.1.5.

11.2.1 Choosing a test road

In the course of this research, five road profiles have been measured for roads around Loughborough, in Leicestershire, UK. The section used for this study connects the B5350 with the B591 and is known as the Breakback Road. It is particularly suited to the design and testing of suspension observers and controllers, as it excites both the body bounce and wheel-hop modes of the system. The road is 1.2km in length, and is divided into two equal sections, a design track and a test track. Each is traversed at 20m/s, to give 30 seconds of data.

11.2.2 Nonlinear simulation model

To obtain a representative model for the vertical stiffness of the rolling tyre, a simple frequency domain analysis is conducted. The test vehicle was driven along a B class road with all dampers switched to their softest setting. By recording wheel accelerations over four tests carried out at different fixed speeds, the characteristic frequency of wheel-hop for the vehicle can be estimated.
Figure 59: ASD plots of measured wheel acceleration, to determine the wheelhop modal frequency.
Figure 59 shows auto-spectral density plots of the measurements, and the wheelhop frequency is clearly discernible. Interestingly, the plots also show a speed varying excitation on the wheel-hub that is presumably associated with the wheel rotation. The disturbance frequencies occur at twice the wheel rotation frequency, and are probably due to mass unbalance, runout or tyre stiffness variations.

Estimating the frequency of wheelhop from the plots at 12.5Hz, the rolling tyre stiffness can be estimated from a simple linear analysis of the quarter-car system. It can be shown that the effect of damping at the softest setting is negligible, so the wheel-hop mode can be described by

\[ \omega = 2\pi f = \sqrt{K/M} \]

Here the total stiffness \( K \) is given from the parallel action of tyre and suspension springs, and the effect of body mass is assumed to be negligible, giving the required estimate

\[ K_t = (2\pi f)^2 \times M_w - K_s \]

Using the known wheel mass and suspension spring stiffness identified in Chapters 5 and 7, the rolling tyre stiffness is estimated at 138.5kN/m.

### 11.2.3 Closed-loop control in simulations and vehicle tests

As with the filter design of Chapter 9, the vehicle Kalman filter treats control force changes in the suspension system as a source of noise. Both design tests therefore employ a nominal closed-loop control strategy in order to exercise changes in the control action, and enable suitable noise estimates to be made. Here, the simple 'skyhook law' feedback of absolute vertical body velocity \[24\] is used, as subjectively this was found to provide good ride comfort in the vehicle. The required suspension force is given by

\[ F_s^* = -Kx_4 \]

where \( K \) is chosen as 2000, and this is 'clipped' to provide a semi-active control action, as described in Section 10.1.

To maintain compatibility between the simulated and vehicle measured data, the same control law is executed in simulation as on the vehicle. For the simulation study this control is applied assuming perfect state knowledge, whereas on the vehicle a prototype Kalman filter (based on nominal noise assumptions) is employed. The simulated and measured controls are not expected to correlate exactly, just as the road inputs to each system will not correlate exactly; the aim is just to achieve a nominal compatibility between responses.
11.2.4 Deriving the linear form of the model

The five-state quarter vehicle model reverts to a linear form if a linear damping rate is adopted. This rate is required to represent the full range of damping that may be invoked in closed-loop control. To choose a suitable rate, a closed-loop nonlinear simulation is carried out using full state feedback, and time histories for damping force \( F_d \) and suspension velocity \( x_s \) are recorded. The linear rate is then evaluated by an ordinary least-squares fit to the overconstrained equation set:

\[
x_s(k) \dot{B}_s = F_d(k)
\]

The advantage of this approach is that the rate is implicitly weighted according to realised damper velocities and forces invoked by the control sequence, although it is conceivable that the choice of damping rate may vary with an alternative feedback control law.

11.2.5 Estimation of the output noise covariance model

The form of the measurement error covariance matrix, \( Q_s \), is taken as

\[
Q_s = \begin{bmatrix}
\alpha & 0 \\
0 & \beta
\end{bmatrix}
\]

where the choice of zero off-diagonal elements assumes no correlation between the errors in each sensor. This assumption is simplistic given that high frequency vibrations can easily travel within rigid members of the suspension structure, and these can affect both accelerometers. However, as the principal output noise statistics are difficult to estimate, it is not considered useful to attempt an estimation of interactions.

An 'order of magnitude' estimate of \( \alpha \) and \( \beta \) can be determined by examination of ASD plots of \( y \) and \( y_s \), and these are given in Figure 60. Note that to ensure a proper representation of the measurement noise levels that will affect the Kalman filter, the sampling frequency for this data is equal to the design frequency for the filter, 500Hz.

Consider first the match between measured and simulated accelerations below 20Hz. Here, the differences represent system errors that have not been modelled within \( Q_s \). They are greatest in the body accelerations at frequencies above the body bounce mode, and the effect is thought to be largely due to unmodelled pitch and roll excitations.
Figure 60: ASD of simulated and measured accelerations over the test road
The measured body acceleration signal contains broad band excitations that are not present in the simulated response. These are probably the result of direct transmission of high frequency road irregularities, or through vibrations caused by the engine and drivetrain. An estimate for $\beta$ can be made on the basis of the difference in mean square signal magnitudes above around 20Hz. Estimating the power of this difference in the signals at approximately 0.002 over the remaining bandwidth:

$$\beta = 0.002 \times (250 - 20) = 0.5(m/s^2)^2$$

250Hz being the Nyquist frequency. The difference in wheel acceleration signals is much less dramatic, with road irregularities affecting simulated and measured signals more equally. Note that the road profile used for simulation was recorded at intervals of 0.1m, so for a forward velocity of 20ms, the input signal has a Nyquist frequency of 100Hz - this explains the sharp reduction in the simulated response at that frequency. Assuming that the broad agreement between the signals below 100Hz actually extends up to the Nyquist frequency, the measurement error might be estimated from differences between 70 and 90Hz. Estimating the error level at 0.1,

$$\alpha = 0.1 \times 20 = 2(m/s^2)^2$$

The noise to signal ratios described by $\alpha$ and $\beta$ are 44% and 3% respectively. Having fulfilled the Kalman filter design requirements, the observer is designed according to equations 8.1.5.

### 11.3 Road-testing the Kalman Filter

Full validation of the Kalman filter on the road is difficult without the use of additional instrumentation. Considering the three sources of information described in Table 19, it is immediately clear that simulations and rig tests will not provide an adequate test environment. It follows that as much information as possible must be gleaned from the three vehicle sensors, with due attention given to $D_4$. With some loss of information, all of the states except tyre deflection can be reconstructed from the measured sensors (see Table 20).

The most significant limitation is the lack of low frequency absolute velocity information, which prevents the effects of 'rigid body' drift (discussed in Chapters 9 and 10) being assessed. High quality, piezo-resistive type accelerometers were used on the vehicle, and these are capable of recording low frequencies down to steady state values. However, the numerical integration of these very low frequencies is a poorly conditioned procedure, with low amplitude errors being greatly magnified. The result is a false dominance of very
Table 20: State reconstruction from the suspension deflection measurement $d_s$ and accelerations $a_w$ and $a_b$

<table>
<thead>
<tr>
<th>signal</th>
<th>source</th>
<th>limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>-</td>
<td>no suitable source of road deflection information</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$d_s$</td>
<td>transducer accuracy</td>
</tr>
<tr>
<td>$x_3$</td>
<td>$\int a_w$</td>
<td>unreliable at low frequencies</td>
</tr>
<tr>
<td>$x_4$</td>
<td>$\int a_b$</td>
<td></td>
</tr>
<tr>
<td>$(x_3-x_4)$</td>
<td>$\frac{d}{dt}(d_s)$</td>
<td>unreliable at high frequencies</td>
</tr>
</tbody>
</table>

low frequencies in the integrated signal, and in order to compare integrated accelerations and Kalman filter estimates, it was necessary to digitally filter out the lowest frequencies (below 0.25Hz) in both signals.

Figure 61 shows the measured and estimated acceleration signals for the test section of the 'Breakback' road. These illustrate the innovations (see Section 8.1) and they are equivalent to the spectral comparisons made in Figure 60. The difference in frequency content explained by the model of each acceleration is clear.

Considering the lack of road information available to the model, the wheel acceleration is estimated very well. The quality of fit, especially in respect of high amplitude wheelhop oscillations, is indicative of good tyre deflection estimation; this is the only assessment of $x_1$ prediction that is available. Body acceleration is also well estimated in the body bounce mode, although at the wheelhop frequency, the estimates seem rather exaggerated (e.g. between 10.5 and 10.7 seconds) and even in error (11.6 - 11.8 seconds). This may be due to the modelling assumption that the wheelstations are independent. The net effect of wheelhop forces from all four independent suspensions will probably be less than the quarter-car model assumes, if the vehicle body can be assumed rigid. Consider for example a torsional configuration of forces from the suspension, which would have a zero net effect on a rigid body.
Figure 62 shows the estimated suspension deflection and velocity, against reconstructed measurements. The suspension deflection is of particular interest, being a directly measured variable not included in the observer sensor set. The gain error that dominates the estimate is similar to that seen in Chapter 10, with comparable magnitude. Suspension velocities are compared below 25Hz although this filtering has a negligible effect on the result; the gain error is again significant.

The comparison of absolute wheel and body velocities, considered separately, is shown in Figure 63. Above the low frequencies, both perform extremely well; $\hat{x}_3$ tracks the integrated wheel acceleration with great accuracy although differences occur at very low frequencies. Due to the integration, it is not possible to attribute these errors to the filter or the sensor data. The body velocity appears to be consistently under-estimated and this appears to be the main contributing factor to the suspension velocity errors noted above.

In summary, the Kalman filter appears to perform reasonably well. The most significant errors that were detectable appear to follow the trends of those in the earlier chapters, and they have a similar magnitude. This is particularly impressive considering the level of sensor noise on the body acceleration signal; feedback gains for the body acceleration are quite low, in response to the expected signal to noise ratio of nearly 50%.
Figure 61: Kalman filter performance in closed-loop control of the moving vehicle; innovations
Figure 62: Kalman filter performance in closed-loop control of the moving vehicle; suspension states
Figure 63: Kalman filter performance in closed-loop control of the moving vehicle; absolute velocities
Chapter 12

Discussion

A discussion of the main points of interest arising out of this study is given here, divided into issues concerning parameter identification and those concerning state estimation. An overall summary of conclusions is then given, along with specific suggestions for further research.

12.1 Model Identification

The quarter vehicle model is recognised as a good simplification of vehicle ride dynamics. In this study identification tests have been carried out on a test vehicle, to validate this claim, and develop simple models to be used as a basis for state estimation, control and simulation.

Two time-domain methods have been presented for parameter identification by fitting measured suspension input-output data to a given model structure. The first involves direct application of an ordinary least squares (OLS) error criterion, and the second is a novel method which employs a randomised integral error criterion (RIEC). A formal investigation has been conducted into the optimality of the RIEC method.

A significant finding, which motivated development of the RIEC method, is the improved rejection of errors caused by higher unmodelled modes that results from integration of the model differential equations. The RIEC method considers net momentum transfer within the system, over given intervals, and the summation of constrained vibrations caused by higher modes within, say the vehicle body, is negligible compared with the transfer from wheel to body that may occur over the same period. This is a result which is predictable for high frequency modes, from the frequency attenuation property of integration, but the method also seems to offer good rejection if the undesirable mode appears at a frequency close to that of the modelled dynamics.

The example considered in simulation is that of vertical powertrain vibration, the resonance for which was assumed to occur within 1Hz of the wheel-hop resonance frequency of the suspension. The integrated error method was found to give significantly better results than the direct OLS approach, provided the powertrain resonance frequency is not specifically excited. The conclusion is that the new method is better able to
distinguish relatively low power, bounded disturbance oscillations from those of the principal system dynamics.

In the course of defining the RIEC algorithm for suspension identification, several free parameters arose. These allow significant freedom in the method; the parameters were chosen to minimise the potential for parameter error, given the known structure of the model to be identified. It therefore follows that RIEC could be tuned to the specific identification requirements of other systems; this flexibility is unique among system identification methods, and it may prove critical in maximising parameter accuracy, particularly if some information is available about likely errors in the measurements. The method also maximises the information available from a given data batch; bias across data batches is therefore easily detected, so faulty batches can be detected with some degree of certainty.

Following successful results in simulation, RIEC was used to identify parameters on the test vehicle. The case study presented illustrates the advantage of an incremental approach to model design based on dynamic measurements, in that the most efficient solution is derived for a given complexity of model. Equation fitting errors and parameter bias motivate the 'next most relevant' modification to the model that will improve explanation. In the case study, a series compliance element in the nonlinear damper model appeared to offer improvements, and this was verified in component tests.

The damper modelling exercise exposed one notable weakness in the identification approach used here, in that the new model structure could not be described by a direct regression of linear parameters; a recursive estimation algorithm was required. Even so, the damper model remains significantly simplified, as no attempt has been made to model force transients to the control input. Development of a comprehensive damper model for control represents a significant additional effort and the problem has not been considered further here.

A further complication of the modelling approach arose when the revised damper model structure was included in the quarter vehicle identification (Chapter 7); spring stiffnesses which were physically interpreted as constant across damping, were found to vary considerably. The allowed flexibility of parameters within the model is an interesting issue. One may legitimately choose to sacrifice physical interpretation to gain additional accuracy, though this can present problems in application of the model; for example if suspension spring stiffness is assumed to vary with damping, a fully nonlinear control strategy would be required where a bilinear method would otherwise be suitable. Also a non-physical structure may limit future freedom to extend the model using further physical justification (e.g. to include suspension spring nonlinearities). Here, strict physical consistency has been maintained.
12.2 State Estimation

The design of a real-time suspension state observer has centred on development of a Kalman filter. The filter design assumption of zero auto-correlation in both measurement and system errors was accepted as a potential weakness in the observer, given that time correlation of the errors occurs. The observer design has been altered to account for non-zero cross-correlation between measurement and system errors however, which is also known to have a significant effect where a dynamic model is required for the output.

To assess sensor requirements, a factorial experiment was carried out, based on simulation and rig-based data. This suggested that wheel hub and body mounted accelerometers alone represent an adequate sensor set for the reconstruction of accurate state estimations; the findings were particularly encouraging in that results for simulated data closely match those based on rig data.

Up to this point the observer had been designed using linear models directly derived from dynamic measurements. This was not a suitable approach for the moving vehicle, so a new approach was proposed, based on a linear model with identified parameters. The design of this observer was constrained however by the poor availability of information about expected error levels; in the design, the measurement noise estimate is based on a comparison of vehicle sensor records with simulated outputs over a known road profile. Nevertheless, within the accuracy of state estimates derived from the available transducers, the observer performance was considered good, and it reflected in trend the findings of earlier observer design studies.

The problem which is most detrimental to observer accuracy is that of low frequency errors. These are manifest in two ways, both as a drift of the state estimation induced by the presence of low frequency information in the sensors, and as an insensitivity to large amplitude, very low frequency events in the road. In Chapter 10 it is intimated that the latter may have some advantage in active control, though this is only true because the active control strategy that was considered, is incapable of assigning a suitable control action to low frequency events; the failure of the Kalman filter thus helps to negate a corresponding failure in the controller.

It seems that an ideal solution would be to remove very low frequencies from the sensor measurements before they are input to the observer. Without low frequency inputs, the rigid body mode of the filter would not be excited, so the resulting closed-loop system should only benefit, by removal of dynamically insignificant geographical features. There are practical problems of phase lag associated with applying a high pass filter to achieve this however, and in any case a more appropriate solution is to attempt to supply more information to the observer and controller, rather than less.
In Chapter 9 the suggestion was made that the noise matrices might be chosen to achieve a faster pole location for the rigid-body drift mode. This could only be achieved by redesigning the filter assuming auto-correlated noise, or more practically, by a multi-parameter optimisation of the noise model, both of which may be difficult to achieve. An approach equivalent to this, which has previously been investigated in [20] and [46] is to recognise the non-white nature of the road input process, and compensate for it in the observer by assuming a pre-filter of white noise, included in the system model.

Another approach might be to include additional states which track the low frequency content of the road input, possibly from correlated, low frequency information in both wheel and body accelerometers. This would mean accurate transducers would be required, rather than simply desirable in a vehicle application, and there may conceivably be problems with poor observer response to sudden events, such as potholes in the road, if the states were related to slow-acting filter poles.

The study of observer errors in closed-loop control, in Chapter 10, suggests that low frequency errors do not significantly affect suspension performance under semi-active control. Note however that the semi-active system used here represents the current state of the art in semi-active actuation, and we should assume that as better actuation allows more effective control, there will be consequential increases in the effect of state errors on performance. Another assumption, that it is not necessary to feed back control inputs to the observer, may not hold under better actuation. Formulation of the observer with a control input is not a difficult problem, and the nonlinear damping force can be properly represented if the observer input is evaluated according to estimated damper velocity applied to the known damper map.

Another interesting point, discussed in Chapter 11, is the difficulty of calculating noise matrices for the moving vehicle. The reliance on order-of-magnitude estimates for the sensor noise casts the optimality of the designed filter into question, and in any case, accurate noise covariance estimates may not provide the best observer; Section 8.3.3 shows that there may be some benefit in deliberately inflating the noise model to compensate auto-correlations. For this reason, on-line modification of the noise estimates may be worth considering.

To this end, a simple vehicle-based experiment was carried out; the magnitude of observer innovations was used to estimate measurement noise levels, and the filter system was replaced at discrete time intervals from a set of pre-designed observers. Although this experiment proved inconclusive for a number of reasons which will not be discussed here, there remains scope for further investigations of on-line refinement of the observer.
For any practical control application, robustness of the observer is likely to be an issue. The Kalman filter has been based on an identified quarter vehicle model with fixed parameters, but in practice body mass will certainly vary, and tyre pressures (and hence dynamic tyre stiffness) may also change appreciably. As an estimator of suspension state, the observer may be well suited to self regulation in this respect, and one area of further research which the author would most like to pursue, is on-line estimation of parameters. One possible approach is via an extended Kalman filter, which employs additional slowly varying states; the filter might then be redesigned at discrete intervals if the parameter estimates change significantly.

Extending the concept of on-line identification further, it may be possible to detect failure modes of the suspension system, such as breakdown of the damper, or slow puncture of the tyres.

12.3 Conclusions

The main conclusions of this study can be summarised thus:

- Given reasonably accurate dynamic measurements, suspension parameters can be estimated using a simple time-domain identification method, based on an ordinary least squares error criterion. This method is unsatisfactory in the presence of unmodelled modes close to system dynamic frequencies.

- For the case considered, identification using a randomised integral error criterion provides improved parameter estimation in the presence of an unmodelled system mode.

- Simple model structures can be used to provide accurate estimation of states describing vehicle ride.

- In spite of time-correlated errors, suspension states can be accurately observed in real-time from body and wheel-hub mounted accelerometers, using a Kalman filter based on a linear quarter-car model. Where estimation errors occur, they are predominantly at low frequencies.

- The inclusion of a measurement / system error cross correlation model is significant in Kalman filter performance where acceleration or force sensors are employed.

- Although observer errors occur at low frequencies, for the semi-active actuation considered, a simple linear Kalman filter is adequate for implementation of closed-loop control.
12.4 Recommendations for Further Research

The RIEC method of identification may prove to be more generally efficient in the development of models, particularly for mechanical vibrating systems. The technique should be investigated more widely and compared with other methods of parameter identification to properly assess its potential.

The following issues, which have arisen during development of the state observer, need to be addressed:

- the problem of low frequency errors in state estimation should be investigated further. Extensions to the Kalman filter that enable a more successful response to low frequency inputs may be considered.

- to maximise observer accuracy, techniques for improving estimation of noise, and robustness to system variations should be investigated. On-line parameter and noise model identification seem attractive propositions.
References

Optimal Control: Linear Quadratic Methods
Prentice-Hall International Inc, 1989

Nonlinear Control of a Quarter Car Active Suspension

Random data: Analysis and Measurement Procedures
Wiley and Sons, New York, 1971

A Factorial Analysis of Kalman Filtering for Semi-Active Vehicle Suspension Control

[5] Best, M.C., Gordon, T.J.
Real Time Estimation of Vehicle Suspension Characteristics Using Kalman Filtering

Statistics for Experimenters
Wiley and Sons, New York, 1987

[7] Eykhoff, P.
System identification. Parameter and state estimation
Wiley and Sons, London / New York, 1974
Bayesian Estimation of Transit Rail Vehicle Parameters

The Application of Learning Automata to Advanced Vehicle Suspension Control
IUTAM Symposium on the active control of vibration, Sept 1994, Bath UK, pp 153-159

Applied Optimal Estimation
MIT Press, 1974

Fundamentals of Vehicle Dynamics
SAE Inc, 1992

[12] Gordon, T.J., Best, M.C.
Dynamic Optimisation of Nonlinear Semi-active Suspension Controllers

Engineering Quality and Experimental Design
Longmann, Harlow UK, 1992

[14] Haç, A.
Design of Disturbance Decoupled Observer for Bilinear Systems

Minimal Order State Observers for Bilinear Systems
References

[16] Hedrick, J.K., Rajamani, R., Yi, K.
*Observer Design for Electronic Suspension Applications*

[17] Hemingway, N.G.
*Immittance Identification: An Application to the Dynamic Modelling of Vehicle Components*

*A Nonlinear Sub-optimal Observer-based Control for Semi-Active Suspension*

[19] Hrovat, D., Hubbard, M.
*Optimum Vehicle Suspensions Minimizing RMS Rattlespace, Sprung-Mass Acceleration and Jerk*

*State Estimation of a Road Surface and Vehicle System using a Kalman Filter*

*Identification Methods for Vehicle System Dynamics*

[22] Kalman, R.E.
*A New Approach to Linear Filtering and Prediction Problems*
Journal of Basic Engineering, March 1960, pp 35-46

[23] Kalman, R.E., Bucy, R.S.
*New Results in Linear Filtering and Prediction Theory*

*Vibration Control Using Semi-active Force Generators*
References

[25] Kropáč, O., Sprinc, J.
Identification of the System Vehicle Road Parameters

[26] Lin, Y., Kortüm, W.
Identification of System Physical Parameters for Vehicle Systems with Nonlinear Components

[27] Mahajan, S.K., Krishnan, S.
Method of State Observer Design in Bilinear Suspension System

[28] Margolis, D.L.
Semi-active Control of Wheel-hop in Ground Vehicles

Determination of Mass, Damping and Stiffness Matrices using Structural and Parametric Identification of Linear Vehicle Frame Models

An Introduction to Identification

[31] Ogata, K.
Discrete-time Control Systems
Prentice-Hall Inc, New Jersey USA, 1987

Demonstration of an Automotive Semi-Active Suspension Using Electrorheological Fluid
Proceedings of the SAE International Congress and Exposition, Detroit, USA, March 1995, pp 237-242
Numerical Recipes : The Art of Scientific Computing

[34] Raju, G.V., Narayanan, S.
Optimal Estimation and Control of Non-stationary Response of a Two-degree-of-freedom Vehicle Model

[35] Ray, L.R.
Nonlinear Estimation of Vehicle State and Tire Forces

[36] Resh, W.F.
Some Results Concerning the Effect of Stochastic Parameters on Engine Mount System Behaviour
SAE Transactions, 1991, Vol 100, Sect 6, pp 1444-1451

[37] Robson, J.D.
Road Surface Description and Vehicle Response

[38] Roether, F., Müller, P.C.
Instrumental Variable Identification in Vehicle Dynamics

[39] Sharp, R.S., Crolla, D.A.
Road Vehicle Suspension System Design - A Review

[40] Sharp, R.S., Hassan, S.A.
The Relative Performance Capabilities of Passive, Active and Semi-Active Car Suspension Systems
[41] Shibo, X., Hang, C., Dihuan, Z.
Identification and Structural Dynamics Modification of a Nonlinear System
Vol 1923, Pt 1, pp 274-277

[42] Söderström, T., Stoica, P.G.
Instrumental Variable Methods for System Identification

[43] Stepinski, T.
Comparative Study of Robust Methods of Vehicle State Estimation
Identification and System Parameter Estimation 1982, Proceedings of the 6th
IFAC Symposium, New York, U.S.A., pp 829-834

Semi-Active Control Laws - Optimal and Sub-Optimal
Vehicle System Dynamics, 1994, Vol 23, pp 545-569

Stability Robustness of LQ and LQG Active Suspensions
Journal of Dynamic Systems, Measurement and Control, Transactions of the
ASME, Mar 1994, Vol 116, No 1, pp 123-131

[46] Venhovens, P.J.T.
Optimal Control of Vehicle Suspensions
Doctoral Thesis, Delft University of Technology, Netherlands, 1993

[47] Williams, R.A., Best, A.
Control of a Low Frequency Active Suspension
IEE Conference Publication, 1994, Vol 1, No 389, pp 338-343

[48] Würtelenberger, M., Germann, S., Isermann, R.
Modelling and Parameter Estimation of Nonlinear Vehicle Dynamics
ASME Dynamics Systems and Control Division DSC, Transportation Systems
*Advanced Engineering Mathematics*

[50] Yoshimura, T., Ananthanarayana, N.
*Stochastic Optimal Control of Vehicle Suspension with Preview on an Irregular Surface*

[51] Young, P.C.
*Parameter Estimation for Continuous-time Models - A Survey*

[52] Zhang, H.Y., Chen, J.
*Identification of Physical Parameters of Ground Vehicle Using Block-pulse Function Method.*
Appendix

The following has largely been compiled from theory presented in Applied Optimal Control [10]. The general form of the equations has been used throughout; the input transition matrices $B$ and $D$ are thus both included. For a system which has correlated system and output errors, the continuous-time model equations can generally be written

$$\dot{x} = Ax + Bu + (I)w$$
$$y = Cx + Du + Lw + v$$

where $w$ represents the vector of modelling errors and $v$ represents the sensor measurement error. When the system equation is discretised, the unit matrix multiplier of $w$ becomes the (non unit matrix) $I_d$:

$$x_{k+1} = A_d x_k + B_d u_k + I_d w_k$$
$$y_k = C x_k + D u_k + L w_k + v_k$$

At this point it is useful to describe some assumptions that are made about the noise vectors. Both noise vectors are assumed white (ie. have no autocorrelation across time), eg.

$$E[w_j w_k^T] = 0, j \neq k$$

where $E[ ]$ denotes the expectation operator. The covariance matrices for the vectors are assumed known, and the vectors are assumed to be uncorrelated with each other:

$$E[w_k w_k^T] = Q_C$$
$$E[v_k v_k^T] = Q_{ss}$$
$$E[w_k v_k^T] = 0$$

From these definitions, the discrete-time system and output noise covariance matrices can be defined,

$$E[(I_d w_k) (I_d w_k)^T] = I_d Q_C I_d^T = Q$$
$$E[(L w_k + v_k) (L w_k + v_k)^T] = L Q_C L^T + Q_{ss} = R$$
and also the discrete-time cross covariance matrix:

\[ E \left[ (I_d w_k) (L w_k + v_k)^T \right] = I_d Q_c L^T = S \]

Now, the theory we will use to design an optimal Kalman filter observer relies on there being no correlation between the system noise and the output noise. Clearly \( S \) is non-zero, so we must create an equivalent system with no such correlation. Such a system can be obtained by a ‘Lagrange’ multiplier addition to the state equation A.1:

\[
x_{k+1} = A_d x_k + B_d u_k + I_d w_k + F (y_k - C x_k - D u_k - L w_k - v_k)
\]

where

\[ F = S R^{-1} \]

giving a new system:

\[
x_{k+1} = G x_k + H u_k + F y_k + \epsilon_k \quad \text{(A.4)}
\]

\[ y_k = C x_k + D u_k + \eta_k \quad \text{(A.5)}
\]

with

\[ G = A - F C \]

\[ H = B - F D \]

\[ \epsilon_k = I_d w_k - F (L w_k + v_k) \]

\[ \eta_k = L w_k + v_k \]

It can easily be verified that, for this system description, the system and output noise vectors are uncorrelated:

\[ E [\epsilon_k \eta_k^T] = 0 \]

The filter is designed in two stages: prediction and correction. The diagram below illustrates these two parts of the process, showing the estimated states before and after correction at each discrete time step.
Associated with the predicted and corrected state estimates are error vectors, defined by
\[
\hat{x}_k^{(+)} = x_k + \tilde{x}_k^{(+)} \quad /A.6/
\]
\[
\hat{x}_k^{(-)} = x_k + \tilde{x}_k^{(-)} \quad /A.7/
\]
We desire the correction step for the filter to take the form
\[
\hat{x}_k^{(+)} = K' \hat{x}_k^{(-)} + Ky_k + K'' u_k
\]
which can be written in terms of errors, and using equation A.5 as
\[
x_k + \tilde{x}_k^{(+)} = K' (x_k + \tilde{x}_k^{(-)}) + K (Cx_k + Du_k + \eta_k) + K'' u_k
\]
giving
\[
\tilde{x}_k^{(+)} = (K' + KC - I) x_k + K' \tilde{x}_k^{(-)} + K (Du_k + \eta_k) + K'' u_k
\]
which, for an unbiased case, where
\[
E[\tilde{x}_k^{(+)}] = E[\tilde{x}_k^{(-)}] = E[v_k] = 0
\]
implies that
\[
K' = I - KC
\]
\[
K'' = -KD
\]
The form of equation A.8 can now be rewritten
\[
\hat{x}_k^{(+)} = (I - KC) \hat{x}_k^{(-)} + Ky_k - KD u_k
\]
\[
/A.9/
\]
and using equations A.5, A.6 and A.7 a similar expression can be derived describing the errors across the correction stage:

$$\dot{x}_k^{(+)} = (I - KC) \dot{x}_k^{(-)} + K \eta_k$$  /A.10/

We wish to identify the matrix $K_k$ that will minimise this error, so we define

$$P_k^{(+)} = E[\dot{x}_k^{(+)} \dot{x}_k^{(+)^T}]$$

which from equation A.10 becomes

$$P_k^{(+)} = E[\{(I - KC) \dot{x}_k^{(-)} + K \eta_k\} \{(I - KC) \dot{x}_k^{(-)} + K \eta_k\}^T]$$

and checking the definition of $\eta_k$,

$$P_k^{(+)} = (I - KC) P_k^{(-)} (I - KC)^T + KRK^T$$  /A.11/

which uses a natural definition of $P_k^{(-)}$, and also relies on the further assumption that the state errors are not correlated with output error.

To minimise the mean square correction errors of A.10, we minimise the cost function $J_k$:

$$J_k = trace[P_k^{(+)}]$$

therefore

$$\frac{\partial J_k}{\partial K} = 0$$

A useful rule can be used here:

$$\frac{\partial}{\partial \alpha} [trace(\alpha \beta \alpha^T)] = 2 \alpha \beta$$

which gives, using the chain rule,

$$\frac{\partial J_k}{\partial K} = -2 (I - KC) P_k^{(-)} C^T + 2KR = 0$$

and we have derived the expression for $K$, as

$$K = P_k^{(-)} C^T [CPC^T + R]^{-1}$$  /A.12/
Now we need a general expression for $P_k^{(-)}$ (which turns out to be a recursive form, the Riccati equation). Some manipulation of equations A.11 and A.12 gives an expression for covariance of state errors on the correction stage:

$$P_k^{(+)} = (I - KC) P_k^{(-)}$$  \hspace{1cm} (/A.13/)

and a further expression can be found for the predictive stage state errors; it can be shown that standard propagation through the $G$ matrix is acceptable in representing $\hat{x}_k$ as well as $x_k$. Equation A.4 can thus be written for the propagation of state estimates:

$$\hat{x}_{k+1}^{(-)} = G\hat{x}_k^{(+)} + Hu_k + Fy_k$$  \hspace{1cm} (/A.14/)

and subtracting equation A.4 from equation A.14 we get

$$\hat{x}_{k+1}^{(-)} - x_{k+1}^{(-)} = G (\hat{x}_k^{(+)} - x_k) - \varepsilon_k$$

therefore,

$$\tilde{x}_{k+1}^{(-)} = G\tilde{x}_k^{(+)} - \varepsilon_k$$

and using the definition for $P_k^{(-)}$ at $k+1$,

$$P_{k+1}^{(-)} = E \left[ \{G\tilde{x}_k^{(+)} - \varepsilon_k\} \{G\tilde{x}_k^{(+)} - \varepsilon_k\}^T \right]$$

if we can assume that there is no correlation between $\tilde{x}_k^{(+)}$ and $\varepsilon_k$ we can write the predictive stage state errors as

$$P_{k+1}^{(-)} = GP_k^{(+)} G^T + \widetilde{Q}$$  \hspace{1cm} (/A.15/)

where $\widetilde{Q}$ can be evaluated, from the definition of $\varepsilon_k$, as

$$\widetilde{Q} = Q - SR^{-1}S^T$$

We propose a recursive solution for $P^{(-)}$ by combining equations A.13 and A.15:

$$P_{k+1}^{(-)} = G (I - KC) P_k^{(-)} G^T + \widetilde{Q}$$

and for a stable, time invariant Kalman filter design, we require a time invariant form for $P$, which is the solution of the Riccati equation, obtained by equation A.12 for $K$ in the above:

$$P = GP^T - GPC^T [CPC^T + R]^{-1} CP^T + \widetilde{Q}$$  \hspace{1cm} (/A.16/)
and

\[ K = PC^T (CPC^T + R)^{-1} \]

To generate a predictive Kalman filter system, we require an equation predicting \( \hat{x}_{k+1}^{(-)} \)
from \( \hat{x}_k^{(-)} \) and sensor and input information. This is obtained by substituting \( \hat{x}_k^{(+)} \) from equation A.9 in equation A.14:

\[
\hat{x}_{k+1}^{(-)} = G (I - KC) \hat{x}_k^{(-)} + (F + GK) y_k + (H - GDK) u_k \tag{A.17}
\]

where all of the dependent matrices are now known.

Equation A.17 describes the predictive observer in terms of the alternative system description. Using the definitions for this system, and performing some further algebra, a simple form can be evaluated to describe the original system:

\[
\hat{x}_{k+1} = (A_d - K^0 C) \hat{x}_k + K^0 y_k + (B_d - K^0 D) u_k
\]

where

\[ K^0 = (A_d PC^T + S) (CPC^T + R)^{-1} \]

and to summarise, these two equations have been derived from the discretised linear model matrices, \( A_d, B_d, C \) and \( D, I_d \) and \( L \), and the noise covariance matrices \( Q_C \) and \( Q_{ss} \) using the Riccati equation A.16 above, and the formulae:

\[
G = A_d - SR^{-1}C
\]

\[
\overline{Q} = Q - SR^{-1}S^T
\]

\[
Q = I_d Q_C I_d^T
\]

\[
R = LQ_C L^T + Q_{ss}
\]

\[
S = I_d Q_C L^T
\]