An Optimisation Study on the Control of Clutch Engagement in an Automotive Vehicle

J C Matthews
Department of Engineering
University of Warwick

Thesis submitted to the University of Warwick for the degree of Doctor of Philosophy

September 1994

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank Dr. R. P. Jones for his help and guidance with my research. I would also like to thank S.E.R.C. and Ford Research and Engineering Centre for their sponsorship of the project. In particular, I would like to thank Dr. P. R. Crossley and Dr R. E. Dorey of Ford Research and Engineering Centre, for their technical advise, and for the data that they have provided. I would also like to recognise the assistance of other Ph.D. students, for their companionship, support and advise. In particular, the assistance of Simon Farrall, with his enlightening pragmatic approach to automotive research. Finally, I would like to thank my wife Mandy, for her support and patience during the last four years.

Declaration

The technical content of this thesis is entirely the work of the author. Some of this material has been presented at international conferences, as indicated in the bibliography.

Summary

This thesis contains a formal mathematical investigation of clutch engagement in automotive vehicles. This investigation is conducted by developing a model of an automotive powertrain, and investigating undesirable effects that can occur in clutch engagement. This naturally leads to the development of a multi-objective optimal control problem describing how to best to engage a clutch. An algorithm for solving this optimal control problem is then presented. Arguments for the development of a feedback control strategy are then discussed, with the construction of such a feedback strategy, along with the computations required to evaluate the feedback controls detailed. A further extension, of adapting the feedback controls, to cope with powertrain model perturbations then follows, along with a method of estimating such perturbations. Finally, the use of this research in implementing clutch engagement control is outlined. Throughout the thesis, the various control strategies designed are evaluated by carry out simulations of models representing the powertrains of two different family cars.

Chapter 1

Introduction

The ability to automatically control processes performing desired tasks to a satisfactory standard is an important problem that has been addressed in one form or another for many years. Over a century ago, steam engines were regulated using Watt flying ball governors and arc lamps were controlled using electromagnetic devices [1]. Later, at the beginning of this century, controllers which were more conventional, were manufactured for the purpose of regulating industrial processes [2]. In these early controllers it is questionable whether much theory was used in the design process.

The theory of controlling systems has been an area of active research for many years. Early work tended to concentrate on the stability of processes often by examining the characteristic equation. In the 1930's, further performance assessment techniques, analysing the frequency response of systems were developed by Bode at the Bell Laboratories. Subsequent work through the 1940's and 1950's continued to concentrate on performance assessment techniques, generating methods which are generally referred to as classical control theory. In the late 1950's and throughout the 1960's a new form of state space control or modern control theory was developed. This included the development of concepts such as observability and controllability, observer theory and optimal control theory, in

which controllers were designed using predominantly theoretical techniques, and less engineering intuition. These techniques are very much model based, requiring modelling techniques such as system identification and parameter estimation. Since the development of state space theory, a number of other control topics have been studied, notable examples of which are the work on adaptive and learning control and intelligent control. However, the most significant advances have been made in the problem of robust control, the problem of controlling uncertain systems. A more detailed history of control engineering can be found in [3].

Classical control theory includes a variety of performance assessment techniques used as design aids for linear system controllers, which are normally single input single output (SISO). Most of the techniques involve interpreting graphical representations of the system dynamics and require the intuition and experience of control engineers in controller design. For example, systems can be examined by analysing their step and impulse responses. Time response features, such as percentage overshoot, settling time, steady state error, and system gain can be used to assess controller performance. Other important quantities can be inferred from these time responses, such as dead time and the system time constant. Frequency response techniques were developed in the telecommunications field at the Bell Laboratories, and describe how the gain and phase lag of oscillating signals passing through the system vary as a function of frequency. Bode developed a method of plotting these variations, in which gain and frequency are scaled logarithmically, and can be used to ascertain system characteristics such as cut off frequencies and system order. The graph of gain against phase lag, again with the gain scaled logarithmically, is accredited to N.B.Nichols. A further graphical technique useful in stability assessment was developed by Nyquist [4], in which the gain and phase lag of the frequency responses are plotted in polar coordinates. Subsequent analysis of these graphs by examining the open loop poles, zeros and

the number of encirclements of the point (1,-180°), often referred as the -1 point, indicates whether the closed loop system is stable or unstable. An understanding of how the controller affects the Nyquist plot can be used in their design. A further feature of these plots is that the distance between the nyquist plot and the -1 point is a measure of closed loop robustness. Nyquist plots for single input single output (SISO) systems can be generalised for multi-input multi-output (MIMO). W.R.Evans [5] developed yet another graphical technique, known as root locus plots, which depict how the poles of a feedback system move as the gain of the controller varies. All of the above graphical techniques are particularly useful in controller design, using modern computer graphics, allowing the plots to be quickly generated, understood and analysed. An example of a software package which performs such tasks is the MATLAB control toolbox.

Prior to the introduction of state space control theory, other algebraic assessment techniques were available, such as stability assessment by examining the roots of the characteristic equation. These roots are not always easy to calculate, however, a procedure referred to as the Routh-Hurwitz method gives an indication of system stability, and can be extended to investigate the conditions which would cause the onset of instability. Lyapunov [6] developed a further general non-linear stability analysis method in which stability is guaranteed if a Lyapunov function exists. Unfortunately, finding such a Lyapunov function is often difficult and no conclusion can be drawn if such a function is not located.

State space control theory allows the analysis and control of general multi-input multi-output (MIMO) systems. This theory has particular relevance to linear systems because the notation allows the model to be conveniently expressed, allowing straight forward algebraic techniques for analysing such systems. Examples of the theory that has been developed for linear systems include: notions such as controllability and observability; the placement of closed loop poles; the esti-

mation of system states from system measurements; and stability assessment by calculating eigenvalues, which are all performed using simple matrix operations. One of the most important contributions to state space control theory was the development of optimal control theory, such as the maximum principle of Pontryagin [7] and the dynamic programming of Bellman [8]. In general, the solution of the resulting equations is non-trivial. However, for some linear problems with quadratic costs, an optimal control problem referred to as the Linear Quadratic Regulator (LQR) problem, can be solved by solving a matrix Riccati equation. An extension to the LQR problem, known as the Linear Quadratic Guassian (LQG) problem, uses a Kalman filter [9], to estimate the system states in the presence of noise. This filter is designed using optimal stochastic control and forms a dual problem to the LQR problem. The concepts and theory of optimal control are well explained by Athans and Falb [10].

Since the development of modern state space control theory a number of control topics have been investigated. Three of the most prevalent control topics are now outlined.

Adaptive and learning control aims to improve upon traditional techniques by the use of on-line measurements in addition to a priori knowledge. Such control strategies are necessary for control problems with uncertainties and/or properties that vary in time. A universal definition of adaptive control is lacking, but broadly speaking an adaptive controller is a control strategy where a controller is automatically adjusted by dynamics which are slower than the dynamics of the controller. These adjustments react to internal or environmental variations which may occur. The problem of adaptive control is that the resulting control systems are inherently non-linear, resulting in slow progress in the development of adaptive control theory, with much effort addressed to analysing the stability of the adaptive processes. An introduction to adaptive control which describes

many of the different approaches and techniques which exist can be found in [11]. Learning control embraces adaptive control but tends to make greater use of past and present information in order to predict 'good' control strategies from measured values rather than react to them. An introduction to learning control was recently presented at a symposium [12], with some of the theory used in such control strategies found in [13].

Intelligent control is the term referring to control techniques inspired by observations of biological systems, and tend to use methods borrowed from the field of artificial intelligence. This area of control has particular importance to the problem of controlling systems requiring autonomy, the ability of self government as described in [14]. A common way of constructing such an autonomous system is by using a hierarchical structure. For instance a standard PID controller can be automatically tuned by some higher level devise, in order to cope with system variations. This devise can then be supervised by an additional level in order to prevent the system encroaching into undesirable operating conditions, which can be in turn supervised by yet another level which might be designed to cope with system failures. Intelligent control in its present form includes the three areas: fuzzy logic; neural networks; and expert systems.

Fuzzy logic provides a deterministic environment for vague linguistic concepts. In this environment, operations which are extensions to set operations can be performed, and are used in the operation of fuzzy controllers. The strength of this fuzzy technique is that it formalises the knowledge and experience of control engineers, transferring this expertise into the control strategy, in essence transferring the intelligence. An introduction to fuzzy control can be found in [15] with a recent application of such techniques to automotive control found in [16].

Neural networks, have a parallel architecture, inspired by the interconnections of neurons in biological nervous systems. They consist of nodes which are simple

non-linear functions connected via weights to form a network, allowing the network to be trained to perform a desired task by adjusting the weights. Described in [17] and [18] are applications of neural networks to control.

An expert systems is a term referring to a computer system imitating a human expert, an expert in performing a particular task. A common example are systems that are devised to imitate a medical doctor, which try to diagnose a patient's illness. Expert systems consist of two separate units, a knowledge base supplied by human experts (perhaps containing a list of all illnesses and their symptoms) and an inference engine, which is a program which uses this knowledge to make decisions (for instance by asking questions about the patient the program decides the most probable illness). For control purposes these techniques have particular importance to controlling systems too complex to model, designing controllers and fault detection as explained in [19].

In recent years, considerable progress has been made in robust control, the problem of designing controllers to achieve performance objectives for uncertain systems. In most robust control theories, the uncertainty of the system is expressed by treating the model as an element of a set of models. Many different methods of expressing this set have been used and as a result many different branches of robust control have been developed. For instance, the system model set could be represented by a system characteristic equation, with coefficients bounded by upper and lower limits. In this case the stability of this model set can be assessed using Kharitonov polynomials [20]. Another method of expressing the model set is to use differential inclusions, in which the state space model is extended by allowing the state derivatives to be an element of a set dependent on the states, the controls and possibly time. An example of some stability theory resulting from this model set form can be found in [21]. The most abundant robust control theory is referred to as robust H_∞ control theory, in which a transfer

function representing the model is perturbed by a parameter, bounded by an H_∞ norm. Performance is expressed by a tracking condition, where a weighted sensitivity function is again bounded by an H_∞ norm. Other criteria such as control effort can also be expressed using H_∞ bounds, with weights in the problem formulation being used to trade off performance against robustness. Having defined the problem, if a solution exists, it can be found using an algorithm known as γ-iteration. The concepts and ideas of H_∞ robust control are described in [22], with details of algebraic solution techniques described in [23]. Other more recent state space solutions to the standard problem can be found in [24], where the solution is obtained by solving two Riccati equations. There are a large number of different ways in which the parameter perturbations can be included in the system transfer function. These different approaches have resulted in the existence of many different robust H_∞ control theories. One well known method for solving MIMO robust control problem was initially proposed by McFarlane and Glover [25], in which, due to the form in which the parameter perturbation is included, the complexity of the calculations is reduced. Recently the ideas of robust H_∞ control theory have been extended to cope with structured bounds of the perturbation parameters. The concepts and theory of these techniques, referred to as μ-synthesis, can be found in [26]. For a detailed bibliography on robust control refer to [27] and [28].

This thesis is concerned with applying this wealth of control theory to control problems in automotive engineering. To date most effort has been applied to the designing of engine management systems, although many other problems have been successfully examined. As a brief review of automotive control this area is subdivided into three groups: suspension control; steering control; and propulsion control. A good reference source for automotive control is [29].

Suspension control is concerned with dynamic adjustments of the suspension

system in order to improve vehicle ride and handling. In its most complicated form, referred to as active suspension, hydraulic or pneumatic devices are included in the suspension system which may be connected in parallel or series with a conventional or passive system components. These additional devices generate external power which can be used to stiffen or slacken the suspension system as required. Semi-active suspension system is the term referred to the dynamic adjustments of a passive suspension systems characteristics. For instance the damping can be varied by adjusting an aperture restricting the flow of oil in a damping strut. Two other terms, referring to suspension control systems are, slow-active suspension which refers to an active suspension system with a low bandwidth so as not to excite particular vibration modes and roll control. A further application of control theory in vehicle suspensions is the problem of levelling the vehicle under variable loading, using self-levelling devices. A recent thesis [30] has looked at suspension control and contains more details including a review and the history of suspension control.

A driver can be assisted in steering a vehicle with the use of rear wheel steering control. This is to say that the rear wheels are steered by a control system, with the rear wheel steering angle determined from the front wheel steering angle, steered by the driver, and possibly other variables such as vehicle speed. For instance, for large front wheel steering angles which normally occur at low vehicle speeds, if the rear wheels steer in the opposite direction to the front wheels then the manoeverability of the vehicle can be improved. On the other hand, for small front wheel steering angles which normally occur at high vehicle speeds, if the rear wheels steer in the same direction as the front wheels then the yawing of the vehicle can be reduced. An example of such a control strategy can be found in [31]. Another higher level steering control problem is to design a control system to track a vehicle along a path, possibly defined by a white line of a submerged

electric cable. An recent example of literature addressing this problem is [32].

Propulsion control, reviews of which can be found in [33] and [34], is the largest area of automotive control and can be subdivided into engine control, transmission control and braking control.

Engine management systems, control the complex combustion and mechanical processes that occur in the engine and arise from the need to design engines with smooth and reliable performance, and to meet emission regulations. Examples of individual control tasks in engine management systems include: the control of the air/fuel ratio prior to combustion, sometimes called lambda control; the control of engine idle speed allowing the engine to idle at lower speeds; the minimisation of fuel consumption; and the control of ignition. Details of some of these control systems can be found in [35] and [36].

Transmission control deals with the problem of scheduling the ratio between the engine flywheel and the driveshaft, in order to influence the engine speed and the torque being transmitted to the tyres. This can be accomplished using either an automatic or semi-automatic transmission with the gear ratios governed by either a conventional discrete gearbox, or a continuously variable transmission (CVT). With a conventional gearbox, gear changes must be scheduled, the 'when to change problem', and whilst the gear change is effected, the engine has to be either fully or partially disengaged from the gearbox using a clutch, fluid coupling or torque converter, the 'how to change' problem. Examples of research, investigating this problem can be found in [37] and [38]. With a CVT the disengagement is unnecessary apart from when idling at rest. However, with the greater availability of gear ratios and the non-linearity of CVTs, the problem of choosing the gear ratio is more complex, and has been addressed in a recent thesis [39]. The disconnection of the engine from the transmission using a friction clutch, requires the clutch to disengage and reengage smoothly bringing the clutch plates speeds

back together, whilst preventing large torque variations. Many torque converters also have bypass clutches which lockup when cruising so as to prevent energy loss through the torque converter, which also needs controlling.

Braking control is concerned with using the braking system, either to prevent the wheels locking up under braking, known as so called anti-lock breaking system (ABS), or to prevent the tyres spinning under acceleration, so called traction control. Such systems normally consist of rule based algorithm reacting to rapid changes in the wheel dynamics. For instance, for traction control, a sudden increase in the speed of a particular wheel could be corrected by increasing the brake pressure to that wheel.

An additional propulsion control problem which mainly involves engine control but which might also require transmission control is cruise control, in which the speed of the vehicle is maintained at a reference value by controlling the engine throttle. An example of this problem can be found in [40].

Two further areas of automotive control that have not yet been mentioned are: the control of auxiliary systems such as air conditioning systems; and the hierarchical problem of combining propulsion, steering and suspension control to manage the total motion of the vehicle [41].

This thesis addresses the task of automatically controlling a friction plate clutch. A typical friction plate clutch, consists of high friction plate attached to the gear box input shaft which is squeezed between two other plates attached to the flywheel. The force squeezing the plates together is exerted by diaphragm springs which are released by a release bearing when the clutch pedal is depressed. There are three distinct problems in controlling the dynamics of such a clutch. These problems are:

1. to engage the clutch plates,

2. to disengage the clutch plates,

3. and to control the slipping of the clutch plates allowing the vehicle to travel at very low speeds.

The first two problems are similar, with the engagement problem being more difficult as engine stalling has to be avoided and driveline oscillation excitation occurs at the point of clutch lock up, as will be seen in Chapter 2. The final problem is that of causing the vehicle speed to follow a reference value. In this thesis the first problem, the engagement problem is addressed. To date, little work seems to have been done in the field of clutch control, with most of the work that has been done using intuitive ideas which fail to consider the problem in a systematic way. For instance, one technique previously used to engage a clutch [42], is to initially engage the clutch quickly, then slowly, then quickly again, which ensures that at the point of clutch lock up the torque being transmitted through the clutch is not too large, preventing large transient oscillations. The most popular technique used to automatically engage a clutch is a technique known as centrifugal clutch control, where the clutch torque pressure is a function of engine flywheel speed. This centrifugal approach has been studied in the context of conventional clutches and torque converter bypass clutches [43] and [44]. The idea behind this technique is that, assuming that the torque being transmitted through the clutch is proportional to pressure exerted between the clutch plates, then the clutch torque can be chosen so that:

- for low engine flywheel speeds the engine torque is greater than the clutch torque,
- for high engine flywheel speeds the engine torque is smaller than the clutch torque,
- and for an acceptable engine flywheel speed operating range the clutch torque approximates the engine torque.

This ensures that: for low engine speeds the resultant torque on the flywheel is positive: for high engine flywheel speeds it is negative; and for acceptable engine speeds the resultant torque is small. This ensures that the engine flywheel speed moves into the acceptable range if it is outside this range. Hence, stalling can be prevented by preventing the engine speed from falling excessively. An example of such a piecewise linear function of clutch torque against engine flywheel speed is shown in figure 1.1, with this approach being well described by [45]. Refinements to this method can be made by including additional terms such as rate of engine flywheel speed change and throttle angle, as in [46], to cope with unconsidered aspests such as dynamics of the clutch actuation mechanism. Other clutch control problems such as clutch to clutch shift operations have also been studied [47].

In applying control strategies to automotive applications a number of hardware components are required. Sensors are needed to acquire measurements of systems variables, a processing unit is needed to carry out the necessary on-line computations demanded by the control strategy, and actuators are needed to influence the system.

There are a number of sensors available for automotive control, many of which are designed for specific automotive applications. Examples of such sensors are: lambda sensors which measure the number of free oxygen molecules in the exhaust gases; hot air mass flow flow sensors which can measure the flow of fuel through the inlet manifold; flywheel reluctance sensors which measure the rotational speed of the crankshaft by counting teeth in a reluctor ring; and gearbox sensors which detect the current gear that is engaged. Other universal sensors such as temperature sensors and pressure sensors are also available. Similarly there are a number of actuators designed specifically for automotive applications, such as fuel injection systems and ignition amplifiers. Again universal actuators such as stepper motors and hydraulic actuators are sometimes used.

The computations in most modern automotive systems are performed by digital microelectronic processing units, which consists of a central processing unit (CPU) together with a large number of additional features. Some of these additional features include: non-volatile memory usually electronically erasable read only memory (EEPROM); analogue to digital and digital to analogue converters; timers to control data flow releasing the CPU from such tasks; an internal clock; and ports allowing communication with other external devices and allowing hardware extensions to the unit. Often these automotive control units can be tailor made to individual control applications by the choice of the units modules. Examples of such control unit systems in current use are the TMS370 [48] and the 8096BH [49] families.

In recent years, sensor and actuator technology has improved, developing new devices such as smart sensors that may well enable system measurements which come with appreciation of their reliability. Furthermore, the speed of processors is continually increasing; a figure often quoted is that the speed of microelectronic processors increases by a factor of 10 every five years. An additional innovation, the development of serial communication between individual control units, allows devices to share information. Many different automotive serial communication systems exist as described in [50], whose requirements are discused in [51]. This continual improvement in automotive electronics, motivated by to use of automotive control, and more importantly the reduction in cost of outmoded components, has meant that increasingly complex control strategies will become more feasible.

With this background to the current state of control theory, automotive control applications, and the available technology for the implementation of automotive control, the aims of this thesis project are as follows:

1. to carry out a formal mathematical investigation of the control of clutch engagement,

2. to identify design requirements of clutch engagement,

3. to design a control strategies which take account of all of the design requirements,

4. and to develop techniques making the implementation of the control strategy more realistic.

The remaining chapters describe how these aims have been addressed. Chapter 2, develops a mathematical formulation of the clutch engagement, expressed as an optimal control problem. Chapter 3 attempts to solve this optimal control problem, locating control strategies which exhibit 'good' clutch engagement. Chapter 4, goes on to consider clutch engagement under perturbations, developing a method of controlling clutch engagement under such perturbation. Chapter 5, develops this approach one step further, developing a method for coping with perturbations in the powertrain if they are known, along with approaches for obtaining such variations. Finally, the use of this research in implementing automatic clutch control is discussed.

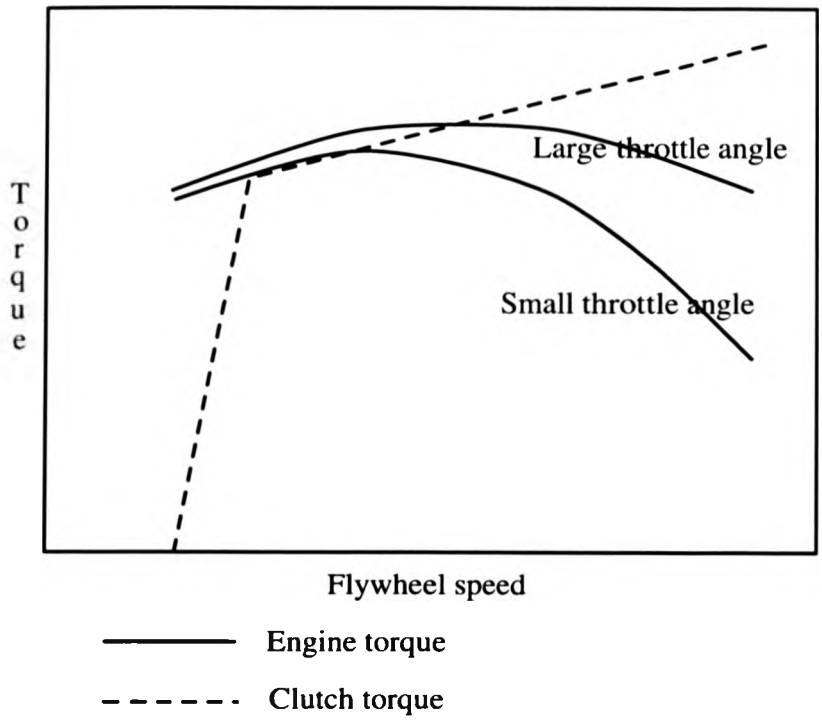


Figure 1.1: Centrifugal clutch control example

Chapter 2

Mathematical Formulation of the Clutch Engagement Problem

The first step towards carrying out a rigorous investigation of clutch engagement is to develop a mathematical formulation of the problem. This formulation begins with the development of a simple automotive powertrain model. Analysis and simulation of this model allows the notion of clutch performance to be expressed. This quantification of clutch performance, along with the powertrain model, then naturally results in the formulation of a continuous multi-objective optimal control problem.

2.1 Powertrain Model

An automotive powertrain consists of vehicle components transmitting power from the engine to the vehicle mass. Typical powertrains consist of: an engine which produces the power by burning fuel; a clutch or a fluid coupling which either fully or partially disengages the engine from the remainder of the powertrain; a discrete or continuously variable transmission which allows the torque transmitted to the tyres to be manipulated; a differential incorporating a final drive ratio which allows the vehicle wheels to rotate at different rates whilst cornering; and driveshafts which transmit the power between components and finally to the tyres.

The tyres and the suspension then convert the power into longitudinal motion. For this study the powertrain considered is as in most family cars with manual transmissions, consisting of an engine, a clutch, a discrete gearbox, a differential and driveshafts.

A simple model of this powertrain is now developed by making a number of modelling assumptions, from which model equations arise. These assumptions have been made by considering the powertrain dynamics and other more complex powertrain models including previous powertrain modelling work done at the University of Warwick [52]. Model details that are either felt to have negligible effect for the clutch engagement problem or felt to be too problematic in the later optimisation techniques for the small improvement in model accuracy, are ignored.

The first component in the powertrain is the engine. Extensive effort has been applied to engine modelling for control studies, resulting in a wealth of literature detailing engine models, such as the models described in [53], [54] and [55]. A review of automotive engine modelling has been conducted by Powell [56]. The first progression towards an engine model is normally to describe its steady state characteristics, acquired from engine test bed data. An example of such a static map, acquired from previous engine modelling work done at the University of Warwick [57], is detailed in figure 2.1. In this figure, the throttle angle is the engine input and the engine output is the engine torque which is also dependent on the engine flywheel speed fed back from the flywheel dynamics. More detailed engine maps can incorporate additional engine inputs such as spark advance and air/fuel ratio. The engine dynamics can, in its simplest form, be represented by a dynamic lag from the throttle angle to the engine torque, with this lag being dependent on quantities such as inlet manifold volume, engine temperature and flywheel speed. More detailed models represent the fluctuations in pressure of various engine chambers. As the time constants of the dynamics in the engine

are significantly faster than the time constants of dynamics in the rest of the powertrain, for the model, the engine dynamics are ignored. Typically the engine responds to a step in throttle angle in about 0.01 - 0.05 seconds as identified by [53], with the powertrain dynamics being controlled over a time interval of 0.5 - 2.0 seconds. Furthermore some engine inputs such as spark advance respond even faster. With this assumption that the engine dynamics are neglected, the engine torque is taken as a engine input, control u_1. The corresponding steady state throttle angles and other possible engine inputs such as spark advance can, if required, be obtained from the inverse of an engine map. Additional dynamics in the engine result from the movement of the engine on its mounts. This movement can be modeled by 'earthing' (connecting to a static mass) the engine block inertia via a suitable compliance. As other compliance, felt to be more dominant, is present in other powertrain components, these extra dynamics are ignored with the engine represented by the engine torque and the flywheel represented by an inertia J_1. However, this last modelling assumption will result in a lack of oscillations at the engine side of the clutch, that would be present with this compliance included.

A clutch acts as a torque limiter, limiting the torque transmitted through the clutch to a maximum value, the clutch torque capacity. This clutch torque capacity is dependent on many factors including the pressure between the clutch plates, the temperature of the clutch plates and the rate of clutch slip. The variation of clutch torque capacity with respect to clutch slip has been associated with self exciting shudder [43]. The variations with respect to temperature requires the modelling of additional dynamics, examining the dissipation of energy through the clutch. However, as the variation of clutch torque capacity with respect to clutch slip is normally small and the dynamics of the clutch temperature fluctuation are slow, both of these effects are neglected, if it is noted that the variations with respect to temperature induce some uncertainty of the clutch torque capac-

ity. The remaining correlation between the clutch torque capacity and the clutch plate pressure previously mentioned, is felt to be the key clutch characteristic. For this reason, the clutch is modelled by assuming that the torque capacity of the clutch is proportional to the pressure exerted between the clutch plates. If this clutch plate pressure is taken to be a control, then the clutch torque capacity can be taken as a control, control u_3, eliminating a proportionality constant in the model equations. Deterioration of clutch performance, due to clutch wear, can be modelled by multiplying the expected clutch torque capacity u_3 by a parameter $\mu \in [0, 1]$. This parameter can also be used to model the uncertainty of the clutch torque capacity due to temperature changes noted above. Other compliance dynamics, resulting from clutch springs, is ignored as the compliance in the remainder of the powertrain is felt to be dominant.

A typical automotive gearbox consists of oiled cogs contained in a chamber with the input shaft connected to the clutch and the output shaft connected to a driveshaft or differential. As well as exhibiting the desired gear ratio, due to the meshing of the cogs, lash is present in the gearbox, and due to friction between the cogs and their bearings, some torque loss is present. In new gearboxes the size of the lash is not too large, of the order of 4 or 5 degrees. Furthermore, if only positive torque is being transmitted through the gearbox, the lash effect is redundant, a likely situation for the clutch engagement problem. Hence for this simple model, lash is neglected. The torque drop through the gearbox is also ignored, just introducing some uncertainty on the torque being transmitted through the powertrain. Hence the gearbox is modelled by taking the gearbox input and output shafts to be lumped inertias, inertias J_2 and J_3 respectively, with the gearbox itself being modelled by setting the ratio of the output shaft speed to the input shaft speed to a constant value, control u_2.

The remaining components between the gearbox output shaft and the vehicle

mass consist of driveshafts, a differential, tyres and a suspension system. These components exhibit compliance, a drop in the torque being transmitted through the components (referred to as rolling resistance) and an effect referred to as wheel slip, where the expected vehicle speed calculated from the gearbox output shaft speed differs from the actual vehicle speed. One way of modelling these components is to model them as a spring in series with a non-linear damper, which is a tyre model. This modelling is relatively complex, a simpler model would be preferable. As identified by previous powertrain modelling work done at the University of Warwick [52], the key dynamic effect is the compliance in these components, which does not seem to be too nonlinear. Indeed, comparison of a typical tyre model linearised about a particular vehicle speed, in series with a linear spring, indicates that these components can be represented by a linear spring and damper in parallel. Hence for this model these components are represented by a linear spring (having a spring constant k) and a damper (with damping rate ν), in parallel, connected to a lumped inertia (inertia J_4) representing the vehicle mass reflected up the powertrain through the differential. It should be observed that the analysis of tyre models suggests that the compliance parameters, in particular the damping rate, may vary with vehicle speed. It should also be observed that this model fails to represent rolling resistance and wheel slip. However, for the case of rolling resistance a constant torque term representing the torque drop can be included in the retarding torque z, or the torque drop can be treated as a further uncertainty of the torque being transmitted through the powertrain. This retarding torque can also be used to represent air resistance and retarding forces due to the gradient of the terrain the vehicle is travelling over.

The above modelling assumptions result in a model diagram, detailed in figure 2.2. Model equations can be constructed from the above modelling assumptions, or the model diagram, either by considering the transfer of energy between

the rotational inertias or by considering resultant torques acting on the inertias and using Newton's second law. The limitation on the transfer of torque through the clutch induced by the clutch model (the torque being limited by the clutch torque capacity) results in the model equations being dependent on whether the clutch is slipping or locked up. The two conditions of the clutch will be referred to as the modes of the clutch.

The first clutch mode is when the clutch is slipping. For this mode, the torque being transmitted through the clutch is equal in magnitude to the clutch torque capacity, with the sign of this torque being determined by the direction of the clutch plate slip, a standard representation of friction between slipping surfaces. With this clutch characteristic, the model equation construction results in the following equations

$$\begin{aligned} J_1\dot{x_1} &= u_1 - \mu u_3 \text{sign}(x_1 - x_2) && (2.1)\\ (J_2 + J_3 u_2^2)\dot{x_2} &= \mu u_3 \text{sign}(x_1 - x_2) - u_2 p && (2.2)\\ J_4\dot{x_3} &= p - z && (2.3)\\ \dot{x_4} &= x_3 - u_2 x_2 && (2.4)\\ p &= -kx_4 - \nu(x_3 - u_2 x_2). \end{aligned}$$

In these equations: p is the torque being transmitted up the powertrain from the inertia J_4 to the gearbox output shaft; x_1 is the engine flywheel clutch plate speed; x_2 is the other clutch plate speed; x_3 is the rotational speed of the inertia J_4; and x_4 is the wind up in the compliance. All the other variables and parameters are as previously noted. Note that the dynamics of the flywheel is only dependent on the sign of the clutch plate slip. Hence, in essence, the flywheel dynamics are independent from the rest of the powertrain when the clutch is slipping. Air resistance is modelled by taking the retarding torque z to be proportional to the

square of the vehicle's velocity

$$\text{i.e. } z = bx_3^2. \tag{2.5}$$

Vehicle load resulting from a constant gradient is modelled by a constant retarding torque

$$z = Mgr_1r_2\sin(\theta) \tag{2.6}$$

where M is the vehicle mass; g is the gravitational constant; r_1 is the final drive ratio; r_2 is the effective tyre radius and θ is the angle of the constant gradient from the horizontal.

The second clutch mode is when the clutch plates are locked up. With the clutch plate speeds equal, the following equations result

$$x_1 = x_2 \tag{2.7}$$

$$(J_1 + J_2 + J_3u_2^2)\dot{x}_2 = u_1 - u_2p \tag{2.8}$$

$$J_4\dot{x}_3 = p - z \tag{2.9}$$

$$\dot{x}_4 = x_3 - u_2x_2 \tag{2.10}$$

with the variables and parameters as previously defined. This set of model equations implicitly determine the transfer of torque through the clutch, the value of which is that required to maintain the clutch plate speeds equal. This torque is

$$\tau = \frac{u_1(J_2 + J_3u_2^2) + u_2pJ_1}{J_1 + J_2 + J_3u_2^2}. \tag{2.11}$$

In order to fully describe the dynamics at any instant in time, the clutch mode which is active must be determined. This is done by describing conditions determining which mode is initially active and switching conditions describing when the mode of the clutch changes. These conditions arise from considering the model equations representing a locked up clutch or a slipping clutch and determining when they become invalid. For instance, when the clutch plates are

locked up the magnitude of the torque being transmitted through the clutch must not exceed the torque capacity of the clutch. Hence the clutch switches to being in its slipping mode when

$$|\tau| > \mu u_3. \tag{2.12}$$

The reverse switching condition occurs when the clutch plate speeds return to being equal, so long as the clutch torque capacity is not exceeded. Hence the clutch switches to its locked up mode when

$$x_1 = x_2 \text{ and } |\tau| \leq \mu u_3. \tag{2.13}$$

The switching conditions also determine the initial clutch modes as they partition, into two sets, the space of model variables. That is

$$\text{Clutch locked up when} \quad x_1 = x_2 \text{ and } |\tau| \leq \mu u_3 \tag{2.14}$$

$$\text{Clutch slipping when} \quad x_1 \neq x_2 \text{ or } |\tau| > \mu u_3. \tag{2.15}$$

Initially these conditions are used to define the initial clutch mode. These conditions could also be used to determine the clutch mode at other instants in time, but are far less convenient for simulation purposes.

A final detail of the powertrain model is the constraints on the controls. For this simple model, the engine torque and the clutch torque capacity are taken to be bounded above and below. The gear ratio value is taken to be an element of a finite set $\mathcal{U}_2$. Hence

$$u_1 \in [u_{1min}, u_{1max}] \tag{2.16}$$

$$u_2 \in \mathcal{U}_2 = \left\{u_2^1, \cdots, u_2^r\right\} \tag{2.17}$$

$$u_3 \in [u_{3min}, u_{3max}]. \tag{2.18}$$

The powertrain model now created is a state space model. The four states of the model are: x_1 - the engine flywheel clutch plate speed; x_2 - the other clutch

plate speed; x_3 - the rotational speed of the inertia J_4 (which is proportional to the vehicle's velocity); and x_4 - the wind up in the compliance. The three controls of the model are: u_1 - the engine torque; u_2 - the gear ratio; and u_3 - the clutch torque capacity. This state space model is not of a standard form as the state equations are dependent on discrete events, the switching conditions, equations (2.14) and (2.15), between the clutch modes.

Examination of the model equations establishes a reduction in the order of the model when the clutch is locked up. In this clutch mode the number of states have reduced by one, as the state equations are no longer dependent on the engine flywheel speed. Furthermore, the number of active controls has also reduced by one as the clutch torque capacity control u_3 no longer affects the state equations. However, this control does still effect the switching conditions, effecting whether the clutch remains in this clutch mode. For this reason in order to clarify and summarise this variable state order model, a narrow class of state space models which the powertrain model belongs to is introduced. Rigorous conditions can then be applied to this model class, conditions which allow the optimisation techniques of later chapters to be applied.

The powertrain state space model for the first clutch mode, representing the clutch when the clutch plates are slipping, is of the form

$$\dot{\underline{x}} = f(\underline{x}, \underline{u}) \tag{2.19}$$

where $\underline{x}$ and $\underline{u}$ are vectors representing the n states and the m controls respectively and $f : \Re^n \times \Re^m \mapsto \Re^n$. The second clutch mode is a reduced model form, with some of the states lacking state equations. The states that lack states equations are dependent on the states that still have state equations. Furthermore, the state equation still remaining are not dependent on all of the controls. Algebraically, this second form can be represented by partitioning the states into states still having state equations and states that do not, and partioning the controls into

controls still active and controls that are not. With this partitioning the model form for the second mode is

$$\underline{x}_1 = g(\underline{x}_2) \tag{2.20}$$

$$\dot{\underline{x}}_2 = h(\underline{x}_2, \underline{u}_1) \tag{2.21}$$

where $\underline{x}_1$ are the n_1 states lacking state equations, $\underline{x}_2$ are the n_2 states having state equations, $\underline{u}_1$ are the m_1 active controls for the second mode, $\underline{u}_2$ are the m_2 redundant controls for the second mode, $g : \Re^{n_2} \mapsto \Re^{n_1}$ and $h : \Re^{n_2} \times \Re^{m_1} \mapsto \Re^{n_2}$. The discrete events are of the form

$$\text{mode 1} \longrightarrow \text{mode 2 if } \underline{x}_1 = g(\underline{x}_2) \text{ and } (\underline{x}, \underline{u}) \in \mathcal{S} \tag{2.22}$$

$$\text{mode 2} \longrightarrow \text{mode 1 if } (\underline{x}, \underline{u}) \notin \mathcal{S} \tag{2.23}$$

where $\mathcal{S} = \{(\underline{x}, \underline{u}) || \tau | \leq \mu u_3\}$ for the clutch engagement problem, is a set in $\Re^n \times \Re^m$. The initial conditions and initial modes are given by

$$\underline{x}(0) = \underline{x}_0 \tag{2.24}$$

$$\text{initially mode 1 if} \quad \underline{x}_1(0) \neq g(\underline{x}_2(0)) \text{ or } (\underline{x}(0), \underline{u}(0)) \notin \mathcal{S} \tag{2.25}$$

$$\text{initially mode 2 if} \quad \underline{x}_1(0) = g(\underline{x}_2(0)) \text{ and } (\underline{x}(0), \underline{u}(0)) \in \mathcal{S}. \tag{2.26}$$

In order to enable the optimisation techniques of later chapters to be applied to the clutch engagement problem, conditions on the model form are introduced. These conditions are that the functions

$$f, g \text{ and } h \text{ are all continuously differentiable (or smooth)} \tag{2.27}$$

and that

$$\mathcal{S} \text{ is a smooth manifold.} \tag{2.28}$$

These conditions are sufficient for all of the optimisation techniques, but can be relaxed for individual techniques. For instance, for the open loop study it

is sufficient for f, g and h to only be continuous, differentiable and to have a continuous derivative.

Inspection of the above powertrain model equations verifies that for the clutch engagement problem they would be in the desired form if it were not for the sign term of equations (2.1) and (2.2), and for the fact that $\mathcal{S}$ is not a smooth manifold. If during clutch slip, the sign of $x_1 - x_2$ remains constant then this sign term can be replaced by a constant, equal to the sign of $x_1 - x_2$ at an initial starting point or at a point in time just after the clutch has switched to its slipping mode, thereby reducing the non-linearity of the state equations. This simplification is not totally unjustified. From equation (2.13), as the states are continuous in time, for the sign of $x_1 - x_2$ to change whilst the clutch remains slipping, $|\tau| > \mu u_3$ at the point of sign change. However, just prior to clutch plate speed equality it can be shown that $\tau < \mu u_3$. Both of these conditions can only be satisfied if $\tau < -\mu u_3$. For large clutch torque capacities, or for large engine torques, this final condition is extremely unlikely. Hence, it is assumed that, when the clutch plates come together with large clutch torque capacities,

$$\tau > -\mu u_3. \tag{2.29}$$

With this assumption, not only is the sign simplification valid but the second term of the switching condition when switching to a locked up clutch, equation (2.13), is automatically satisfied (i.e. $|\tau| \leq \mu u_3$). This removes the difficulties with $\mathcal{S}$ not being a smooth manifold, as in the optimal control problem formulation $\mathcal{S}$ can be taken to be $\Re^n \times \Re^m$.

With the above simplifications, the model is now in the required form. Any state space model of this form would be adequate for the optimisation techniques. Indeed many of the dynamic effects in the powertrain mentioned, but not modelled, could be incorporated without violating the model conditions. However, additional modelling which includes pure time delay or pure lash, used to model

the engine or gearbox dynamics respectively, is prohibited, as this would violate condition (2.27). The penalty of additional modelling is an increase in complexity of the computation required for the later optimisation techniques. This is the reason that the powertrain model is kept as simple as possible, only representing the key dynamics of the powertrain.

Two sets of model data, both representing a typical family car, have been applied to the model. The first set of data, referred to as vehicle A, is obtained from previous powertrain modelling work done at the University of Warwick [52]. This modelling work details measurements of inertias, gear ratios, final drive ratio, vehicle mass, air resistance coefficient and compliance coefficients. These measurements either provide the required model parameters directly or provide data enabling the parameters to be calculated. For instance, the inertia J_4 is calculated from the vehicle mass, the final drive ratio and the effective tyre radius. Note that, for this data set, the compliance coefficients, the spring constant and damping rate, are chosen so that the model approximates the key dynamics of the components that the compliance represents in the powertrain, when excited, with their values being obtained by comparisons with experimental data. Bounds on the controls are also needed. For the engine torque, the upper and lower bounds are taken so that they are obtainable for a large set of engine speeds. These values are obtained by studying some previous engine modelling work done at the University of Warwick [57]. The lower bound for the clutch torque capacity, is naturally chosen to be 0 Nm, representing a fully disengaged clutch, with the upper bound taken to be 2.25 times the maximum engine torque. This value was decided upon after consultation with P.R.Crossley of Ford Research and Engineering Centre [58]. The values of this first set of data are detailed in table 2.1.

The second set of data, referred to as vehicle B, contains similar information of a different, but realistic, powertrain. This data provides values of the inertias, final

drive ratio, gear ratios and vehicle mass, enabling calculation of the vast majority of the model parameters as before. Details of steady state engine characteristics, again, are used to choose the engine torque bounds, again choosing bounds that are obtainable for a large set of engine speeds. However, this time, additional details of clutch actuation characteristics are used to calculate the clutch torque capacity upper bound. For this data set, the compliance coefficients are calculated by inspecting tip in (a sudden change in engine torque) simulations of a complex powertrain model representing the same powertrain, and choosing the values which give the best comparison. The only parameters not provided, or calculated, is the air resistance coefficient which is taken as in the first set. The values of this second set of data are detailed in table 2.2.

Comparing the two sets of data, all of the parameter values are similar in order, apart from the gearbox inertias. The data representing vehicle B has much smaller gearbox inertias, sufficiently small that the extra dynamics generated by the gearbox modelling is negligible. This fact has been ignored for later work as they are significant for the other set. For vehicle B, this observation could be utilised in order to simplify later work. In particular, when the clutch is slipping, if the rate of change of the controls is not too great, then the clutch dynamics can be represented by the following state equations

$$J_1\dot{x_1} = u_1 - \mu u_3 \text{sign}(x_1 - x_2) \tag{2.30}$$

$$J_4\dot{x_3} = \frac{\mu u_3 \text{sign}(x_1 - x_2)}{u_2} - z \tag{2.31}$$

$$x_2 = \frac{x_3}{u_2} \tag{2.32}$$

$$x_4 = -\frac{\mu u_3 \text{sign}(x_1 - x_2)}{u_2 k} \tag{2.33}$$

with the torque being transmitted through the clutch in order to maintain the

clutch plates speeds equal now being

$$\tau = \frac{u_1 J_4 u_2^2 + u_2 z J_1}{J_1 + J_4 u_2^2}. \tag{2.34}$$

This simplification halves the number of states in the model when the clutch plates are slipping, leaving the number of controls unchanged. Elimination of x_2 from the state equations using equations (2.7) and (2.32) results in a model of a similar form to the form identified by equations (2.19) - (2.23). Similar simplifications of the sign term can be made to ensure that this new simplified model is in the desired form, allowing the optimal control techniques to be applied.

There are also some smaller differences between the two data sets. For vehicle B, the compliance is twice as stiff and less damped, the engine torque upper bound is higher and the gear ratio steps are smaller with an additional gear ratio included. These differences are not large enough to alter the modelling but are large enough to affect the powertrain characteristics, as will be seen in the next section.

2.2 Model Validation

The powertrain model now developed has been validated by simulating it using the computer package ACSL. This computer package can simulate state space equation models and contains a state event finder. This state event finder is used to effect the discrete events in the form of the switching conditions required to simulate the powertrain model. For these simulations the control values of the engine torque and the clutch torque capacity have been ramped up and down as appropriate, with the gear ratio being stepped up and down as required. A simulation which indicates many characteristics of the powertrain is an acceleration from rest simulation, incorporating a gear change from first gear to second gear. This simulation has been carried out for both sets of vehicle data, with graphs

detailing the controls, clutch plate speeds, vehicle speeds and driveline torques being detailed in figures 2.3 - 2.6.

For vehicle A, the controls initially remain at zero, with the vehicle at rest and first gear engaged. After 0.5 seconds the engine torque is quickly ramped up to 60 Nm, after which the clutch torque capacity is ramped up, more slowly, to its maximum value of 225 Nm. The controls then remain constant until a time of 4 seconds has elapsed, at which point the clutch is disengaged and the engine torque is ramped down to zero. While the clutch is fully disengaged the gear ratio is stepped up to second gear. The re-engagement of the clutch is similar to the engagement from rest, with the clutch torque capacity being ramped up a little faster.

The details of the clutch plate speeds, the vehicle speed and the driveline torque, in figure 2.4, describe the resulting dynamics. From these graphs it can be observed that as the clutch is engaged and the engine torque is increased, the vehicle starts to accelerate away from rest, but with some small drive line oscillations being present. Initially the engine flywheel speed increases, but as the clutch torque capacity continues to increase the engine flywheel speed eventually decreases. The increase in clutch torque capacity also causes the vehicle to accelerate at a faster rate. At the point of clutch lock up, the rate of acceleration steps down, with large oscillations in the drive line being generated, which then die away to leave a steady rate of acceleration. The occurrence of these oscillations is an effect experienced by learner drivers when engaging a clutch poorly, causing the car to 'kangaroo' forward in extreme cases. When the clutch is disengaged and the gear ratio is changed, oscillations are generated in the powertrain. The re-engagement dynamics are similar to the previous engagement dynamics, with oscillations being generated at the point of lock up. The resulting steady state acceleration in second gear is smaller than that for first gear.

For vehicle B, a similar simulation is carried out as detailed in figures 2.5 and 2.6. The main differences between the two simulation results are that the oscillations when the clutch is locked up are faster and less damped, the magnitude of the gear change is smaller, and that the oscillation frequency when the clutch is slipping is much higher, for vehicle B. The first two differences are explained from noted differences between the two data sets. An appreciation of the high frequency oscillations when the clutch is slipping, can be obtained from analysing the model equations. It can be shown that when the clutch is slipping, the natural frequency and the damping coefficient are

$$\omega_s = \sqrt{k\left(\frac{1}{J_4} + \frac{u_2^2}{J_2 + J_3 u_2^2}\right)} \tag{2.35}$$

$$\xi_s = \frac{\nu}{2}\sqrt{\frac{1}{k}\left(\frac{1}{J_4} + \frac{u_2^2}{J_2 + J_3 u_2^2}\right)}. \tag{2.36}$$

For vehicle A, it can be calculated that $\omega_s = 19.6 rads^{-1}$ and $\xi_s = 0.33$, and for vehicle B $\omega_s = 108.6 rads^{-1}$ and $\xi_s = 0.31$, in first gear. As well as confirming the observed high frequency oscillations, the equations (2.35) and (2.36), demonstrate that small gearbox inertias equate to high frequency oscillations when the clutch is slipping, which are heavily damped. These highly damped, high frequency oscillations are the dynamics that the reduced model equations (2.30) - (2.33) representing a slipping clutch, ignores. An identical simulation carried out using the reduced order model for vehicle B, detailed in figure 2.7, results in nearly identical simulation results. The only noticeable difference is the loss of the fast dynamics at the point of clutch disengagement and gear change. The above simulation results are consistent with the observed dynamics of an automotive powertrain in the frequency range up to approximately 10Hz.

2.3 Performance Quantification

A natural question, when examining clutch performance, is what is 'good' performance. In order to analyse this notion clutch performance is quantified. However, as there are many different aspects to clutch performance, the initial quantification results in several 'cost' values.

In order to carry out this performance quantification for the clutch engagement problem, a number of engagement for rest simulations have been conducted, for both vehicles. The simulations for the engagement from rest problem, the worst case of the clutch engagement problem, identify a number of undesirable effects that can occur. For each vehicle, two simulations are detailed, a slow and a fast engagement, from rest, in first gear (see figures 2.8 - 2.15). For the slow engagement from rest simulation, the engine torque is ramped up to 60 Nm in 0.1 seconds, with the clutch torque capacity being ramped up to its maximum value in 2.4 seconds (see figures 2.8 and 2.12). For the fast engagement from rest simulation, the engine torque is ramped up to its maximum value in 0.1 seconds, with the clutch torque capacity being ramped up to its maximum value in 0.4 seconds (see figures 2.10 and 2.14). The resulting dynamics for these engagement simulations are detailed in figures 2.9, 2.11, 2.13 and 2.15. Each figure contains four graphs, a graph of the clutch plate speeds, a graph showing the rate of clutch energy dissipation, a graph of the vehicle speed (proportional to the state x_3) and a graph showing the driveline torque. The rate of clutch energy dissipation is calculated from the states and controls, by examining power loss through the clutch, which is equal to $\mu u_3 |x_1 - x_2|$. The first two graphs detail the performance characteristics of the clutch. The last two graphs detail the effect on the vehicle dynamics and the driver, with the driveline torque being proportional to the force acting on the driver.

One undesirable effect present in the simulations is the occurrence of driveline

torque oscillations. The driveline torque directly affects the force felt by the driver, hence any severe transient fluctuations in this torque, such as the oscillations, noted in all of the performance quantification simulations, are undesirable, as they cause driver dissatisfaction. The oscillations are also undesirable as they can cause driveability problems and can cause the wear, fatigue and eventual failure of powertrain components. Large fluctuations in driveline torque can occur while the clutch is slipping, with the driveline torque oscillating about a steady rate of increase on this torque, as in the fast engagement from rest simulation for vehicle A (figure 2.11). For vehicle B, the oscillation about the steady rate of increase is negligible, due to the light gearbox inertia, as previously discussed. However, the driveline oscillations are more prevalent after clutch lock up, with the oscillations appearing to be excited at the point of clutch lock up, for all of the simulations. In general, the oscillations tend to be worse for fast engagements.

Another undesirable effect is the dissipation of energy through the clutch. This energy dissipation is undesirable, not only for the resulting power loss, but for the clutch wear and the clutch temperature changes that can result. Excessive temperature variations can change the characteristics of the clutch, causing further difficulties including exacerbating clutch wear. Referring to the simulations, it can be seen that this energy dissipation only occurs when the clutch is slipping. It can also be noted that the total energy dissipated, for one completed engagement, is greatest for slow engagements, the total energy dissipated being the area under the rate of clutch energy dissipation graph.

Stalling is another problem that must be prevented. Stalling can be prevented by maintaining the engine flywheel speed above a minimum permissible value. For vehicle A, an engine flywheel speed below 80 rad/s can cause erratic engine performance and possible stalling problems. Referring to the fast engagement graph (figure 2.11), it can be noticed that the engine flywheel speed drops below

this minimum value, indicating that stalling is a serious and realistic problem. Stalling problems tend to occur while the clutch plates are slipping and when the clutch plates are being engaged quickly. This is due to large torques which are transmitted through the clutch when the clutch is slipping, together with the large clutch torque capacity. These torques are larger than would normally be experienced with a locked up clutch. This large torque transmits the kinetic energy of the flywheel inertia down the powertrain.

These undesirable effects, mentioned above, suggest that in designing a clutch engagement strategy a compromise must be sought, as the various undesirable effects seem to conflict. For fast engagements, the oscillations and the problem of stalling tend to be large, whilst for the slow engagements the clutch energy dissipation is large. This finding has been separately obtained in parallel research [59]. This dilemma must be addressed when designing the clutch performance measure.

Now that these undesirable effects have been observed, they can be quantified by designing several cost functionals for the clutch engagement problem. These cost functionals measure the magnitude of the individual undesirable effects by equating the observed effects with the states and controls of the powertrain model, high values of the cost functionals equating to poor performance. These cost functionals are constructed so they take one of two forms. The first form, the integral form is

$$F_i = \int_0^{t_f} f_i(\underline{x}, \underline{u}) dt \tag{2.37}$$

where t_f is the time of clutch lock up, $\underline{x}$ and $\underline{u}$ are the states and controls of the powertrain model and $f_i : \Re^n \times \Re^m \mapsto \Re$ is a continuously differentiable function. The second form, the terminal form, is a functional dependent on the controls and states of the model at the time of clutch lock up. Hence

$$F_i = f_{it}(\underline{x}(t_f), \underline{u}(t_f)) \tag{2.38}$$

where t_f, $\underline{x}$ and $\underline{u}$ are as previously defined, and $f_{it} : \Re^n \times \Re^m \mapsto \Re$ is again a continuously differentiable function. This is to say that, for the clutch engagement problem, the performance is dependent on the controls and states up until the clutch plates lock up, and the states and control at the point of lock up. The terminal cost form can, and will, be used to predict the clutch performance after clutch lock up. The restrictions on the form of the costs enable optimisation used in later chapters, to be used in order to solve the resulting optimisation problems. In many cases the cost functions naturally falls into one of the desired forms. However, this is not always the case, as will be seen with the construction of a cost measuring the likelihood of stalling.

For the three bad effects previously noted, six cost functions are now constructed. The oscillations in the driveline torque are measured using four costs. One cost measures the oscillation when the clutch is slipping. A reasonable measure of these oscillations is

$$F_1 = \int_0^{t_f} (x_3 - u_2 x_2)^2 dt. \tag{2.39}$$

This is just the square of the difference of the rotational speeds across the compliance integrated over the interval whilst the clutch is slipping. Analysis of the powertrain model determines that a steady state offset of $x_3 - u_2 x_2$ can be present during clutch slip. This offset is proportional to the rate of change of the clutch torque capacity. Hence this cost also measures the rate of increase of the clutch torque capacity and minimising it will try to slow down engagement, as well as minimising the oscillations. The oscillations after engagement are measured by two costs. These costs are just the displacement of the states and controls from their steady state values at the point of clutch lock up, estimating the oscillations after clutch lock up, and are

$$F_2 = \dot{x}_4^2(t_f) = (x_3(t_f) - u_2(t_f)x_2(t_f))^2 \tag{2.40}$$

$$F_3 = \ddot{x}_4^2(t_f) = \left[p\left(\frac{1}{J_4} + \frac{u_2^2}{J_1 + J_2 + J_3 u_2^2}\right) - \frac{z}{J_4} - \frac{u_2 u_1}{J_1 + J_2 + J_3 u_2^2}\right]^2\Bigg|_{t_f} \quad (2.41)$$

where the state equations representing a locked up clutch, equations (2.7) - (2.10), are used to evaluate the expressions. It can be shown that the square of the resulting oscillation magnitude is proportional to

$$\omega^2(1-\xi^2)\dot{x}_4^2(t_f) + (\ddot{x}_4(t_f) + \xi\omega\dot{x}_4(t_f))^2 \quad (2.42)$$

where ω and ξ are the natural frequency and damping coefficient of the driveline oscillations when the clutch is locked up. Hence, if both these costs are small, small oscillations after clutch lock up will result. The last equation could be used to replace the two costs F_2 and F_3 but it is particularly messy when expanded out into the desired form. In the later optimisation techniques, the required differentiation of this cost would be laborious, hence the reason for using the two costs. The final cost functional, measuring driveline oscillations, measures the excitation of the compliance at the point of clutch lock up. Examination of the two state equation sets, determines that at the point of clutch lock up the torque being transmitted through the clutch jumps down. This torque jump excites the compliance as observed. It can be shown that this torque jump is proportional to

$$F_4 = (\dot{x}_2 - \dot{x}_1)|_{t_f} = \left[\mu u_3 \text{sign}(x_1 - x_2)\left(\frac{1}{J_1} + \frac{1}{J_2 + J_3 u_2^2}\right) - \frac{u_1}{J_1} - \frac{u_2 p}{J_2 + J_3 u_2^2}\right]\Bigg|_{t_f} \quad (2.43)$$

where the state equations representing a slipping clutch, equations (2.1) - (2.4), are used to evaluate the expression. It should be noted that not all the above cost functionals are needed to ensure small oscillations. For instance, small oscillations while the clutch is slipping, and a small amount of compliance excitation, will suffice. Similarly, the excitation cost can be omitted, replacing it with the two costs measuring the oscillations after engagement.

The cost functional measuring the dissipation of energy through the clutch, naturally, arises from the calculation of the rate of clutch energy dissipation from the powertrain states and controls. As the energy dissipation only occurs when the clutch is slipping, the total energy dissipation is

$$F_5 = \int_0^{t_f} \mu u_3 |x_1 - x_2| dt. \tag{2.44}$$

A measure for stalling is a more difficult task, than for previous undesirable effects, as the natural functional measuring how low the engine flywheel speed gets is

$$-\min\{x_1\} \tag{2.45}$$

which is not of the required form. However, a functional which has high values for low engine flywheel speeds is

$$F_6 = \int_0^{t_f} r(t) dt \tag{2.46}$$

$$\text{where } r(t) = \begin{cases} (x_1 - x_{1min})^2 & \text{if } x_1 < x_{1min} \\ 0 & \text{otherwise} \end{cases}$$

where x_{1min} is chosen to be sufficiently high in order to prevent the flywheel speed dropping below its minimum permissible level, a level smaller than x_{1min}. This cost measures the distance of x_1 below x_{1min} while the clutch is slipping and, hence, measures the likelihood of stalling. As low engine flywheel speeds are less of a problem when the clutch is locked up, a cost measuring the likelihood of stalling for a locked up clutch is omitted.

Other cost functionals have been examined and explored using the following optimisation techniques. For instance a cost functional is included which is a cost equal to the time of clutch engagement.

$$F_7 = \int_0^{t_f} 1 dt \tag{2.47}$$

This cost can be used to ensure that the clutch engagement process is completed as quickly as is possible. Most of the remaining costs measure subjective characteristics such as vehicle acceleration or vehicle speed, which can be used to ensure that the optimisation techniques result in acceptable control strategies. A control strategy for the engagement from rest problem resulting in excessive acceleration and possible wheel spin is probably undesirable. Again a cost functional measuring excessive acceleration does not naturally arise in the required form. However, the vehicle speed at the time of clutch lock up is closely related to large accelerations. So the cost

$$F_8 = x_3(t_f) \tag{2.48}$$

can be used to prevent excessive vehicle accelerations. Another way to limit this acceleration is to artificially limit the engine torque upper bound. Other physically based costs ensuring practical solutions are left to automotive engineers more perceptive to these demands.

2.4 Multi-Objective Optimal Control Problem

The form of the powertrain model, together with the form of the individual cost functionals, means that when they are combined a multi-objective optimal control problem of the form

$$\min_{\underline{u} \in \mathcal{U}} [F_i(\underline{u}) = \int_0^{t_f} f_i(\underline{x}, \underline{u})dt + f_{it}(\underline{x}(t_f), \underline{u}(t_f))\Big] \text{ for all } i \tag{2.49}$$

subject to

$$\frac{d\underline{x}}{dt} = \underline{f}(\underline{x}, \underline{u}), \ \underline{x}(0) = \underline{x}_0 \tag{2.50}$$

$$\underline{R}(\underline{u}) = \int_0^{t_f} \underline{r}(\underline{x}, \underline{u})dt + \underline{r}_t(\underline{x}(t_f), \underline{u}(t_f)) = \underline{0} \tag{2.51}$$

$$\underline{Q}(\underline{u}) = \int_0^{t_f} \underline{q}(\underline{x}, \underline{u})dt + \underline{q}_t(\underline{x}(t_f), \underline{u}(t_f)) \leq \underline{0} \tag{2.52}$$

results, where $\underline{R}$ and $\underline{Q}$ represent p_1 equality constraints and p_2 inequality constraints respectively, t_f is the time of clutch lock up, and $f : \Re^n \times \Re^m \mapsto \Re^n$, $f_i, f_{it} : \Re^n \times \Re^m \mapsto \Re$, $\underline{r}, \underline{r}_t : \Re^n \times \Re^m \mapsto \Re^{p_1}$, $\underline{q}, \underline{q}_t : \Re^n \times \Re^m \mapsto \Re^{p_2}$ are all continuously differentiable. This optimal control problem allows for the inclusion of additional equality and inequality constraints, but in order to ensure that the terminal time t_f is indeed the time of clutch lock up, at least one equality constraint

$$x_1(t_f) - x_2(t_f) = 0 \tag{2.53}$$

must be included.

The theory relating to such optimal control problems assumes that the controls can vary instantaneously. For the clutch engagement problem, due to the noted fast dynamics of the engine, instantaneous variation of the engine torque is not that unrealistic. The gear ratio during clutch engagement can not change, and is thus taken as a constant. So the only problem is with the clutch torque capacity, for which an instantaneous variation assumption is unrealistic. The solution to this difficulty is to include a clutch actuation mechanism in the state space powertrain model, the input to this actuator being the new control. The actuator model used, arises from some modelling work [60] of a hydraulic clutch actuation mechanism. The inputs to this actuator are electrical and so now instantaneous variations are reasonable. The actuator model is a second order linear model, which is critically damped ($\xi = 1$), having a natural frequency of 10 rad/s (ω). This results in the inclusion of the two state equations

$$\dot{x}_5 = x_6 \tag{2.54}$$

$$\dot{x}_6 = -2\xi\omega x_6 - \omega^2 x_5 \tag{2.55}$$

in the powertrain model, where x_5 is the actual clutch torque capacity, u_3 is the required clutch torque capacity. As u_3 is no longer the clutch torque capacity, in

order to reconstruct the optimal control problem x_5 must replace u_3 in the previous state equations, equations (2.1) - (2.4), and in the individual cost functionals, in particular in cost (2.44). The result is a multi-objective optimal control problem of the form previously noted which satisfies the instantaneous variation assumptions.

For a good solution to the clutch engagement problem, ideally, all of the cost functionals F_i should be small. However, as the undesirable effects seem to conflict, this might not be possible. As previously mentioned, a compromise must be sought when solving the optimal control problem. How to express this compromise is far from trivial.

Multi-objective optimisation theory describes how to express such a compromise. The effects of this theory can be best understood by examining the multi-objective techniques in the cost space, the space of individual cost functionals (i.e. $\Re^p$, where there are p cost functionals). This cost space contains a subset $\mathcal{F}$, the set of all feasible costs. For the clutch engagement problem, this set is defined by the constrained control values mapped into the cost space by the powertrain model and the cost definitions. Note that all of the cost functionals previously defined are bounded above and below, due to the restrictions on the controls. Hence $\mathcal{F}$ is bounded for the problem in question. Furthermore due to the continuity assumptions on the model and cost forms, the mapping from the controls into the cost space is continuous. This ensures that $\mathcal{F}$ is both compact and connected, if the set of feasible controls is both compact and connected. This observation justifies the graphical representation of $\mathcal{F}$ which is used later in this section. With this idea of a feasible cost space, the notion of the best, or optimal, solution can be defined by designing an ordering on the set of feasible costs $\mathcal{F}$. This ordering does not have to be a total ordering. For instance, one technique detailed in optimisation theory literature, such as [61], is to define an ordering by

$$x < y \text{ if there exists } z \in \mathcal{C} \text{ such that } x + z = y \tag{2.56}$$

where $\mathcal{C}$ is a convex cone in the cost space. An example of such an ordering is Pareto optimisation, where the convex cone $\mathcal{C} = \{z \in \Re^p | z_i \geq 0 \ \forall i\}$. Such techniques do not guarantee a unique optimal solution. As there is no particular reason for the clutch engagement problem to adopt this approach, it is felt that expressing the compromise using a total ordering is better. For total orderings, points in the feasible cost space can be indexed by a real number. That is

$$x < y \text{ if } F(x) < F(y) \tag{2.57}$$

where $F(.)$ is the index. Hence the compromise reduces to defining a map $F : \mathcal{F} \mapsto \Re$.

Three such multi-objective optimisation techniques are reviewed and summarised by Garbett [39]. The first method, the most intuitively obvious technique, is referred to as the weighted sum method. This involves defining F, the performance measure, as

$$F = \rho_1(\underline{w}, \underline{F}) = \sum_{i=1}^{N} w_i F_i \tag{2.58}$$

where F_i are the individual cost functionals, w_i are their corresponding weights, and $\underline{F}$ and $\underline{w}$ are vectors of the costs and weights. Referring to this method in the cost space, the optimal solution is the initial intersection of a hyperplane with the feasible cost space, as pictured in figure 2.16. In this figure, the hyperplanes (represented as lines) are contours of constant performance measure value. The direction of the hyperplane progression is determined by the weights.

The second method described by Garbett is referred to as the ϵ - constraint method. This involves minimising one cost functional with respect to upper bounds on the others. Hence

$$F = F_i \text{ subject to } F_j \leq F_{jmax} \text{ for all } j \neq i \tag{2.59}$$

Graphically, in the cost space (figure 2.17), the optimal solution is the initial intersection of a hyperplane, whose direction of progression is in the direction of

F_i, with the intersection of the feasible cost space $\mathcal{F}$ and a set defined by the inequality constraints.

The final method described by Garbett is referred to as the goal attainment method. This method is expressed by defining

$$F = \min\{\epsilon_i\} \text{ subject to } F_i - \frac{\epsilon_i}{w_i} \leq c_i \text{ for all } i \tag{2.60}$$

where w_i and c_i are the weights and goals respectively, corresponding to the cost functional F_i. However this is equivalent to defining F by

$$F = \rho_\infty(\underline{w}, \underline{F} - \underline{c}) = \max_i \{w_i(F_i - c_i)\} \tag{2.61}$$

where $\underline{F}$, $\underline{w}$ and $\underline{c}$ are vectors of the costs, weights and goals, respectively, a more meaningful way to express this method. Graphically (figure 2.18), this can be interpreted as locating an optimal solution where a 'corner' initially intersects the feasible cost space, again with the corners representing contours of constant performance measure value. These corners progress from a fixed point $\underline{c}$, the goals, with the direction of the progression again being determined by the weights. The form of this progressing shape allows for goals to affect the multi-objective solution. The addition of the goals, an additional multi-objective design parameter, allows a better compromise to be expressed.

For the goal attainment method, for $F > 0$, F is equal to the maximum norm of the weighted difference between the individual cost functions and their goals. i.e. $F = ||w_i(F_i - c_i)||_\infty$. With this analogy to standard norms, a new smooth goal attainment method is proposed, smooth in the sense that for $F \geq 0$ the map from $F_i \mapsto F$ is continuous, differentiable and has a continuous derivative. This is achieved by defining

$$F = \rho_2(\underline{w}, \underline{F} - \underline{c}) = \begin{cases} \sqrt{\sum_{i=1}^{N} (\max\{0, w_i(F_i - c_i)\})^2} & \text{if } \exists i \text{ s.t. } F_i > c_i \\ \max_i \{w_i(F_i - c_i)\} & \text{otherwise} \end{cases} \tag{2.62}$$

Graphically, the method is interpreted as in figure 2.19, where the optimal solution is the initial intersection of a 'blunted corner' with the feasible cost set. This new method still preserves the significance of the goals and has additional smoothness properties, as mentioned above.

Finally, all of the above methods can be embraced in one form by defining

$$\begin{aligned} F &= \rho_k(\underline{w}, \underline{F} - \underline{c}) \text{ for } k \in \{1, 2, \infty\} \\ \text{subject to} \quad & F_j < F_{jmax} \text{ for all } j \in J \end{aligned} \tag{2.63}$$

This general form allows for the addition of extra inequality constraints, which can be used to prevent undesirable effects. For instance an upper bound on F_6, the cost measuring the likelihood of stalling can be used to prevent a solution that might cause stalling problems. The definition of this general form maps the multi-objective optimal control problem into a standard optimal control problem of the form

$$\min_{\underline{u} \in \mathcal{U}} [F(\underline{u}) = \rho_k(\underline{w}, \underline{F} - \underline{c})] \text{ for } k \in \{1, 2, \infty\} \tag{2.64}$$

subject to

$$\frac{d\underline{x}}{dt} = \underline{f}(\underline{x}, \underline{u}), \ \underline{x}(0) = \underline{x}_0 \tag{2.65}$$

$$\underline{R}(\underline{u}) = \underline{0} \tag{2.66}$$

$$\underline{Q}(\underline{u}) \leq \underline{0} \tag{2.67}$$

where $\underline{R}$, $\underline{Q}$ and $\underline{f}$ are as previously defined, $\underline{F}$ is a vector of the individual cost functionals and $\underline{w}$ and $\underline{c}$ are the weights and goals of the multi-objective compromise. This final general form is the form of the optimal control problem that will be examined in the following chapter.

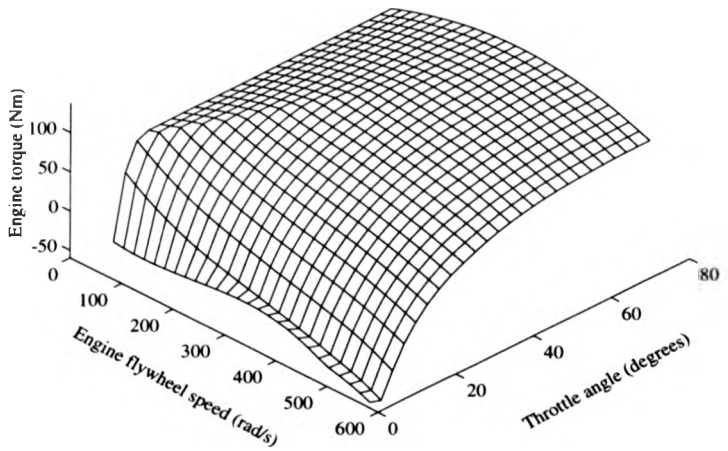
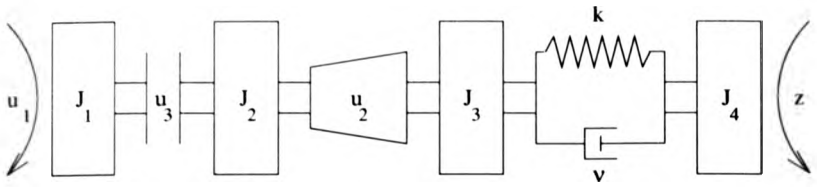


Figure 2.1: Engine map example



J_1 - Flywheel inertia

J_2 - Gearbox input shaft inertia

J_3 - Gearbox output shaft inertia

J_4 - Inertia representing vehicle mass

u_1 - Engine torque

u_2 - Gear ratio

u_3 - Clutch torque capacity

k - Spring constant

v - Damping rate

z - Retarding torque

Figure 2.2: Schematic diagram of powertrain model

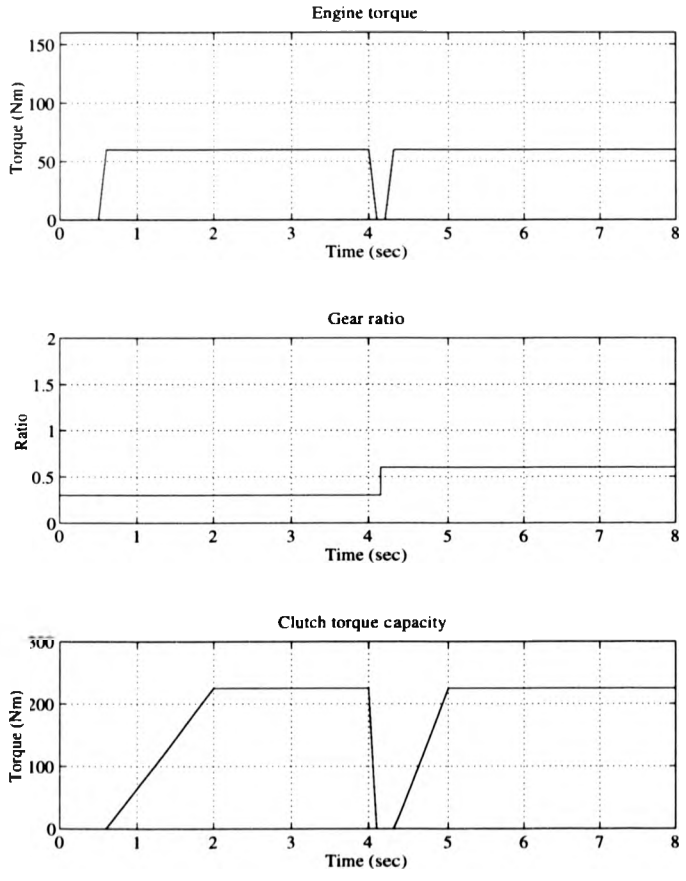


Figure 2.3: Vehicle A: acceleration from rest controls

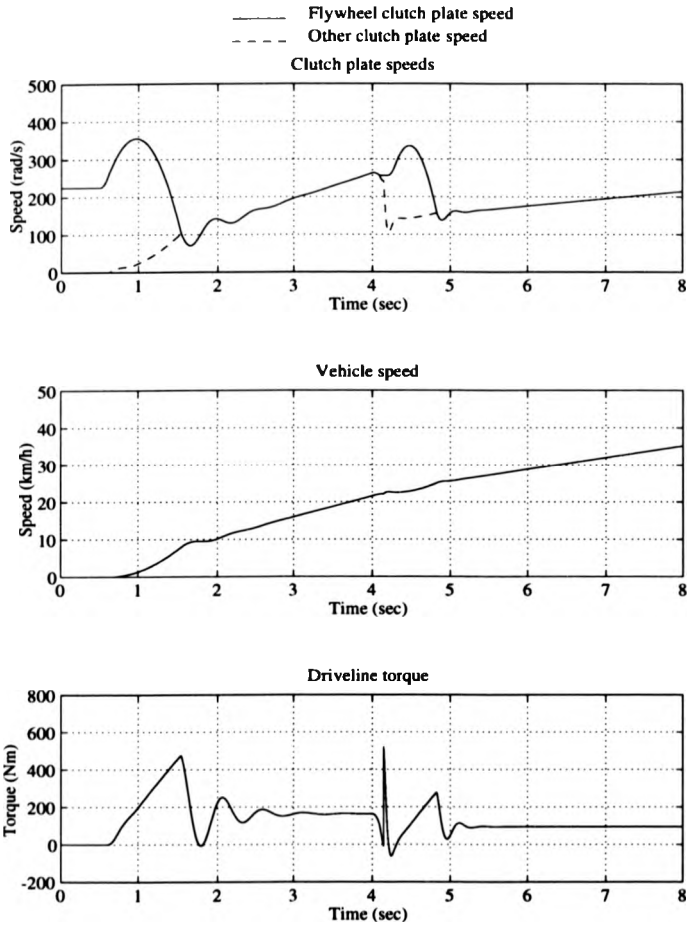


Figure 2.4: Vehicle A: acceleration from rest states

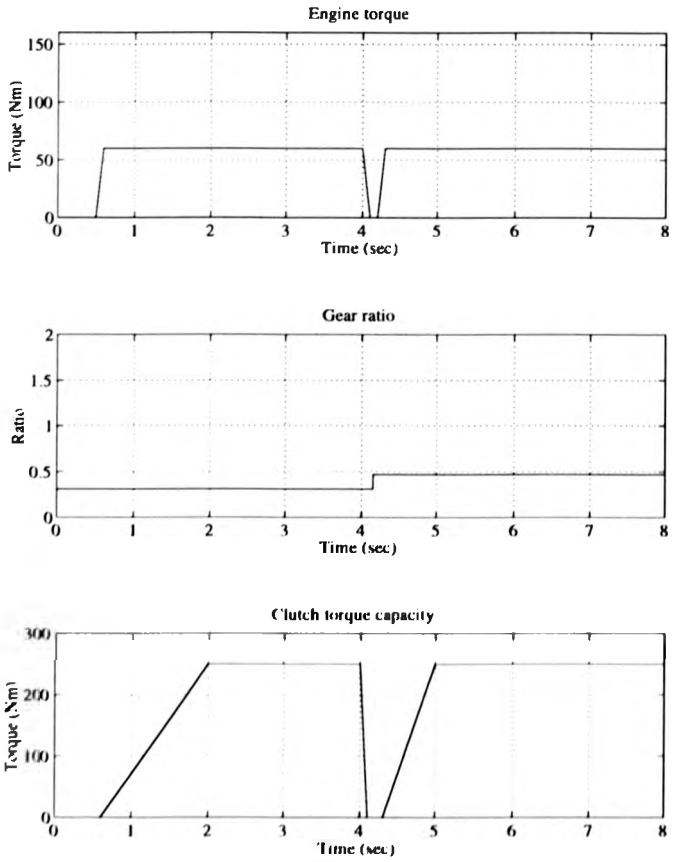


Figure 2.5: Vehicle B: acceleration from rest controls

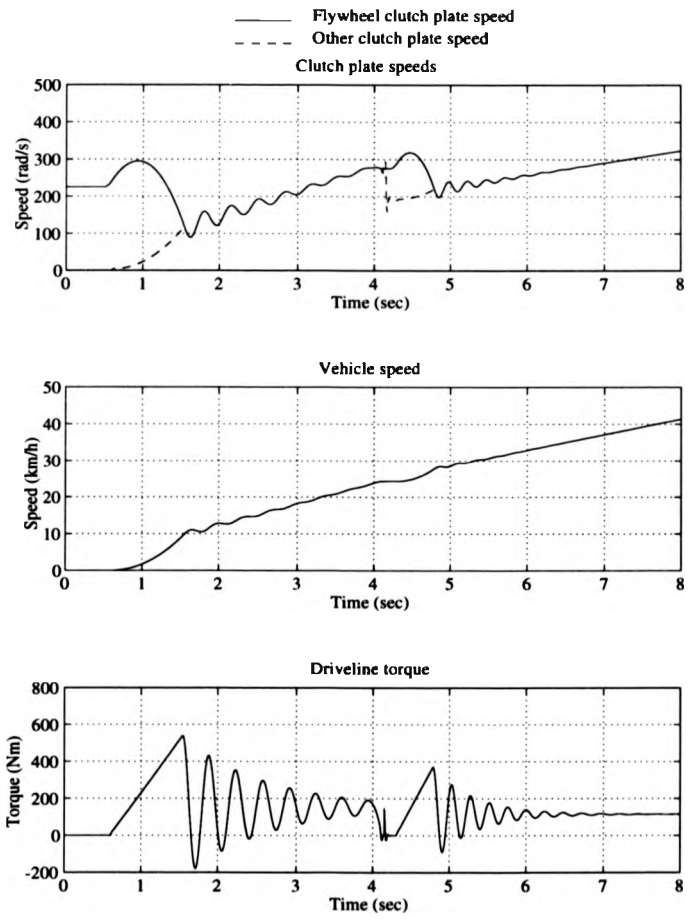


Figure 2.6: Vehicle B: acceleration from rest states

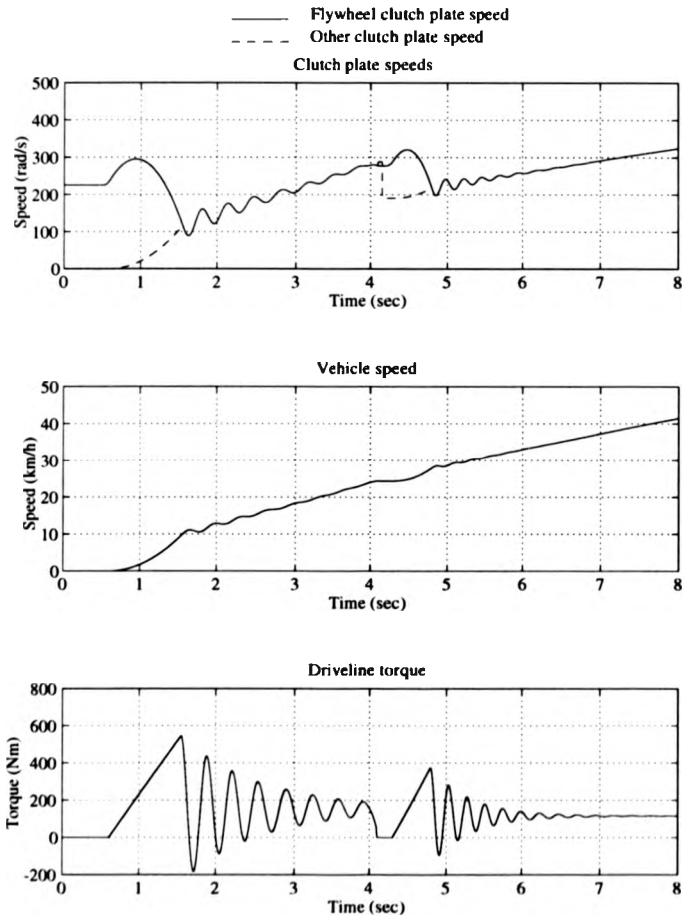


Figure 2.7: Vehicle B: acceleration from rest states (for reduced model)

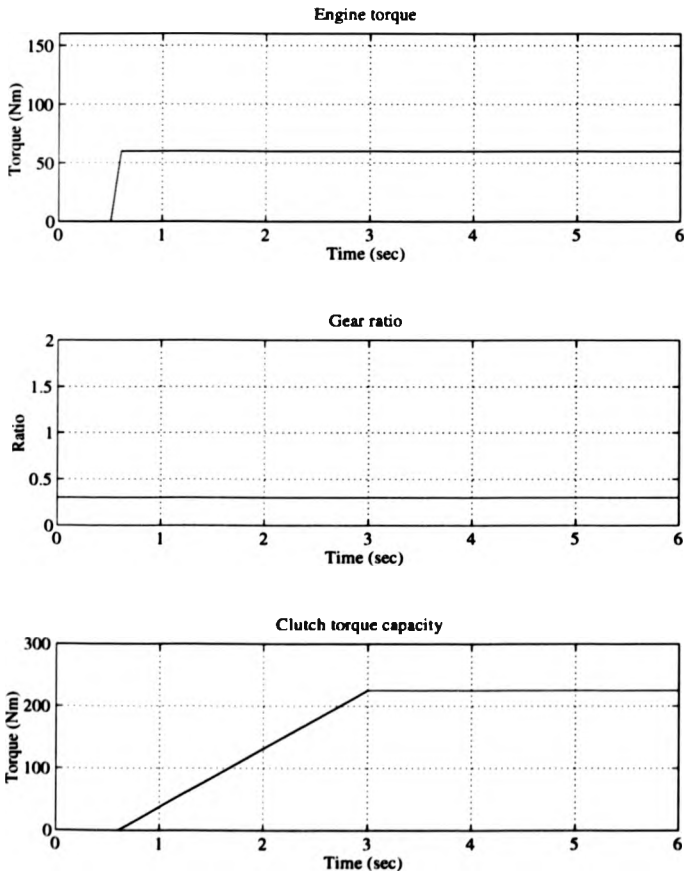


Figure 2.8: Vehicle A: slow engagement from rest controls

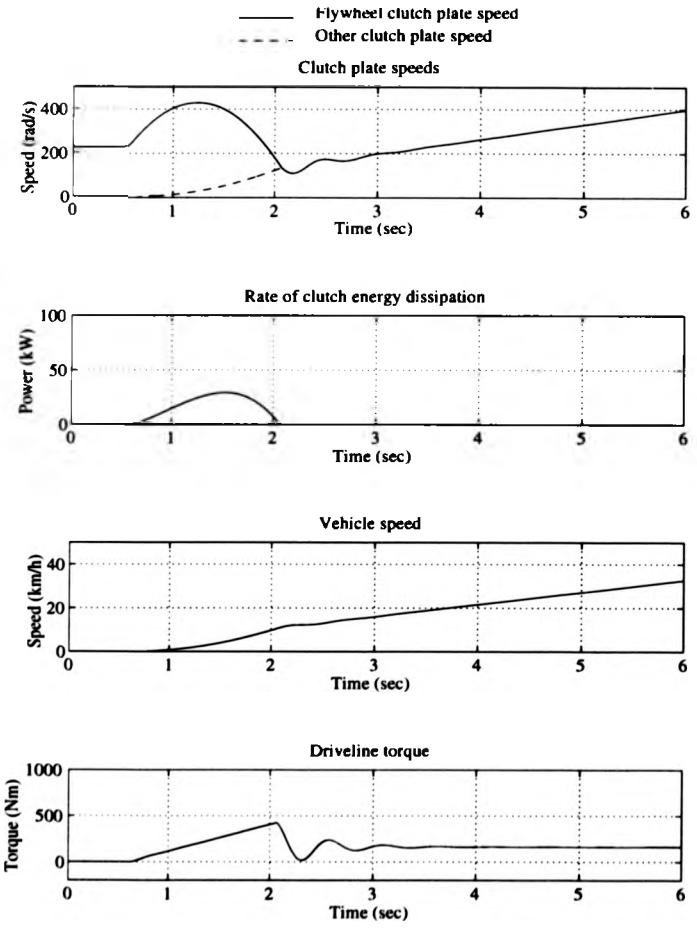


Figure 2.9: Vehicle A: slow engagement from rest states

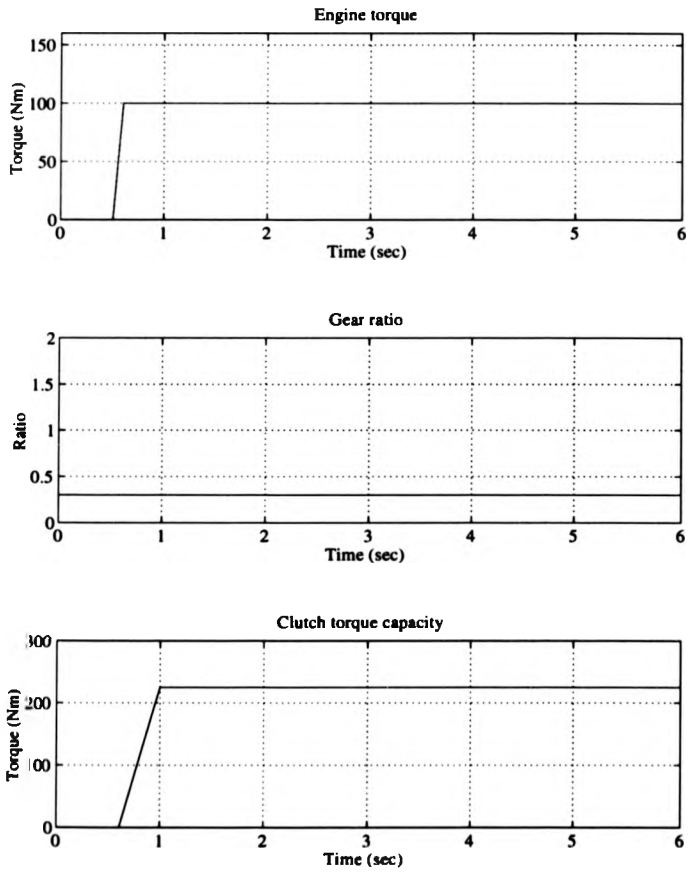


Figure 2.10: Vehicle A: fast engagement from rest controls

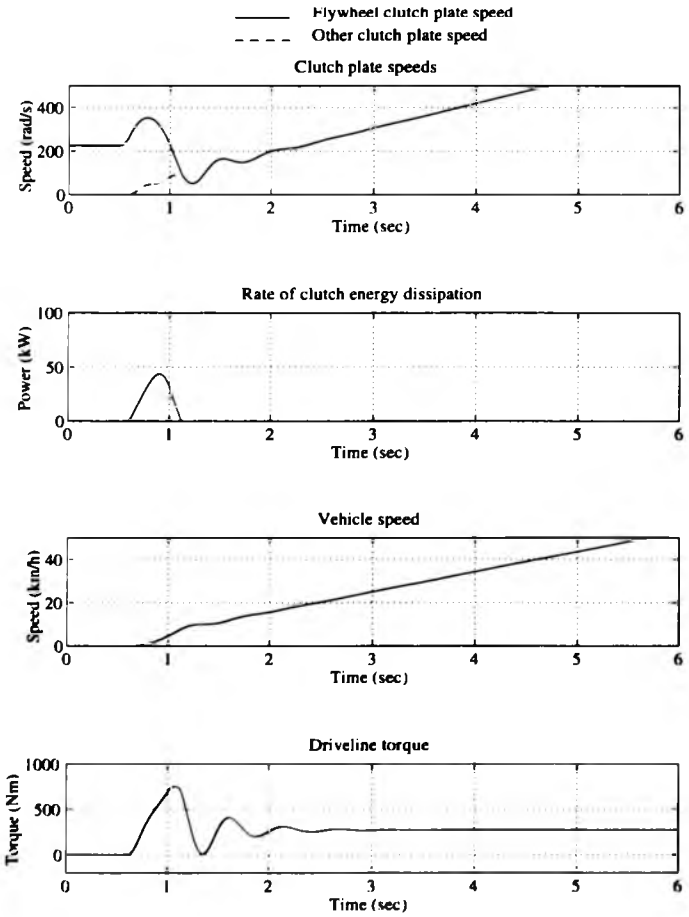


Figure 2.11: Vehicle A: fast engagement from rest states

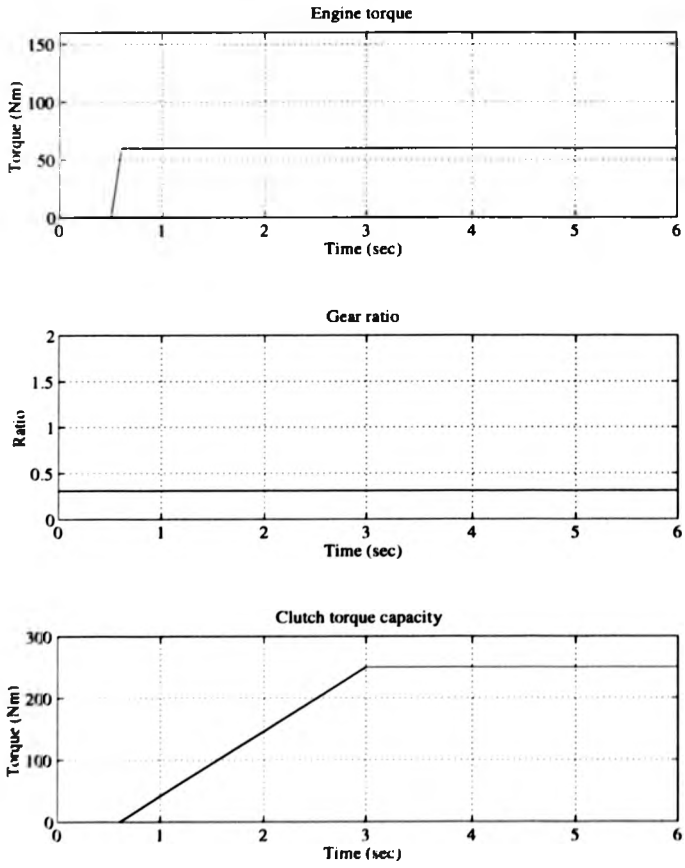


Figure 2.12: Vehicle B: slow engagement from rest controls

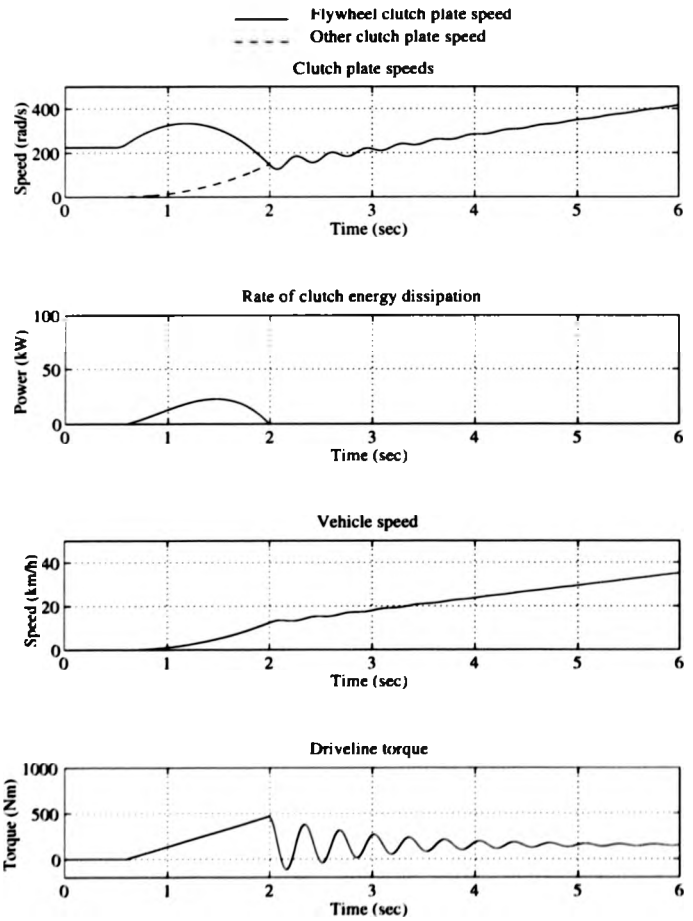


Figure 2.13: Vehicle B: slow engagement from rest states

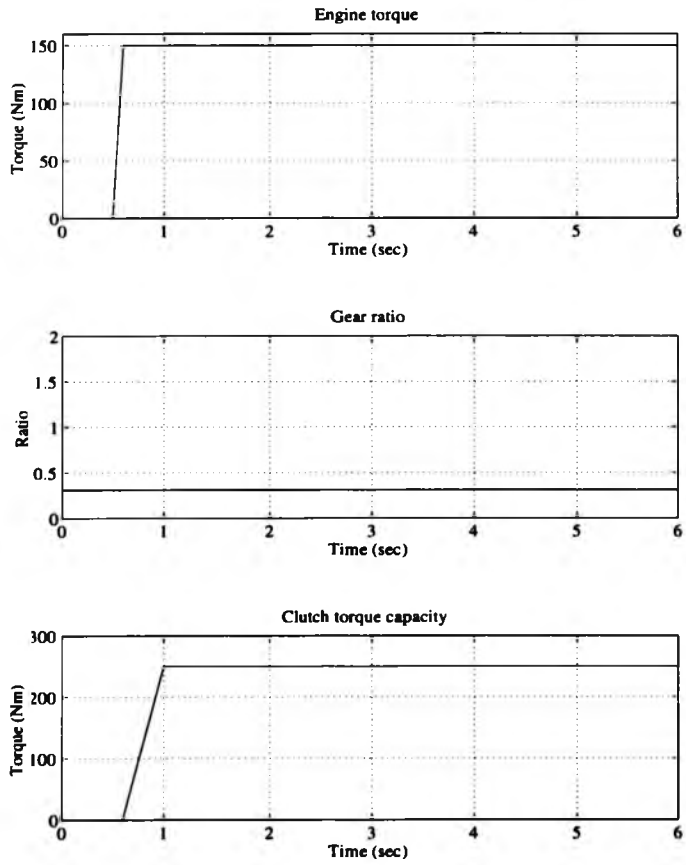


Figure 2.14: Vehicle B: fast engagement from rest controls

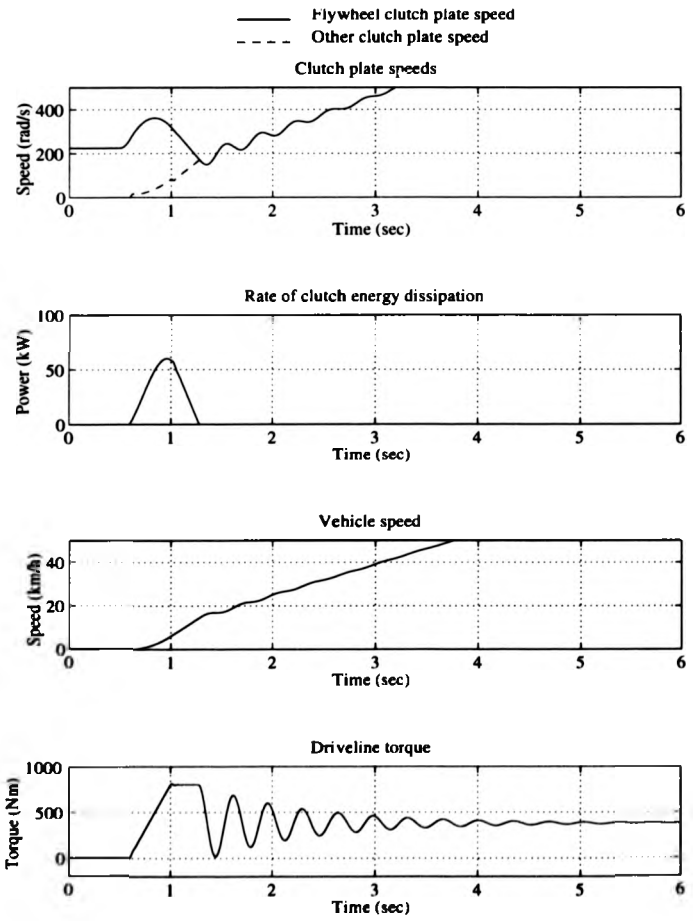


Figure 2.15: Vehicle B: fast engagement from rest states

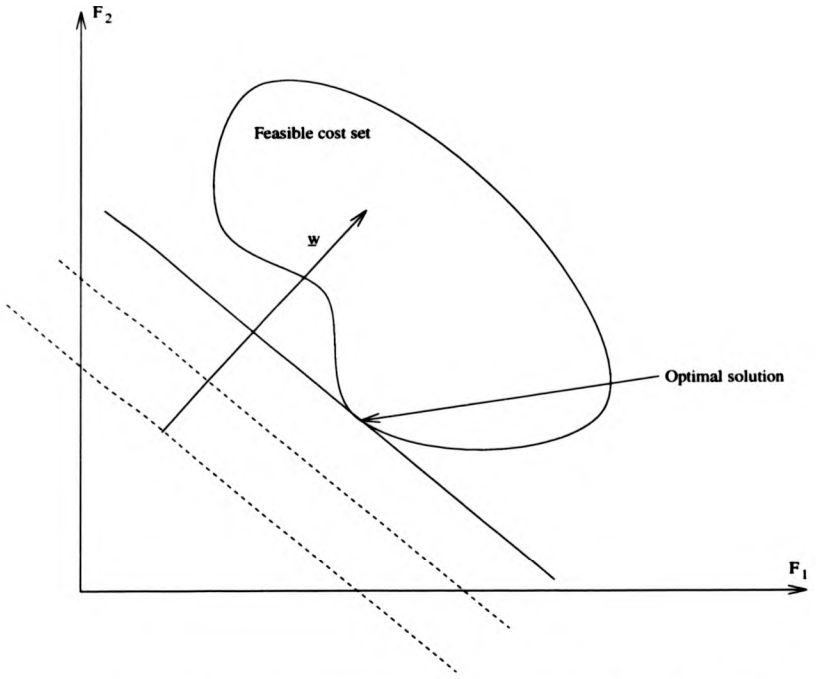


Figure 2.16: Multi-objective compromise in cost space: weighted sum method

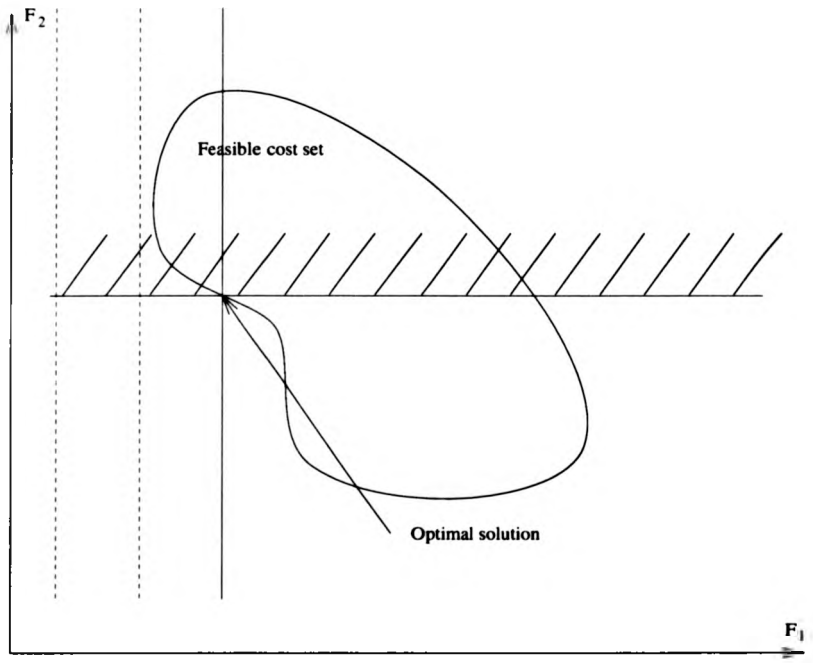


Figure 2.17: Multi-objective compromise in cost space: ϵ constraint

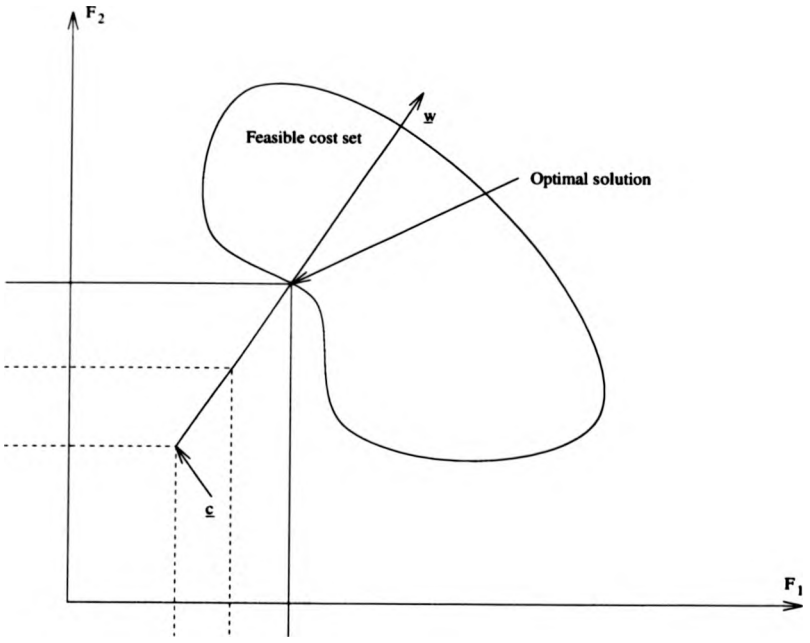


Figure 2.18: Multi-objective compromise in cost space: goal attainment

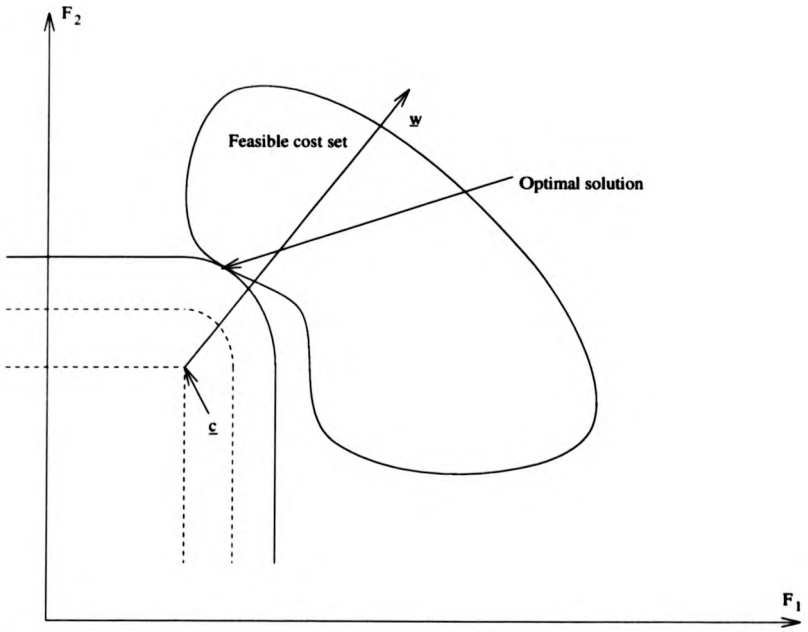


Figure 2.19: Multi-objective compromise in cost space: smooth goal attainment

Parameter	Symbol	Value	Units
Flywheel inertia	J_1	0.109	kgm^2
Gearbox input shaft inertia	J_2	0.05	kgm^2
Gearbox output shaft inertia	J_3	0.05	kgm^2
Inertia representing car mass	J_4	7.945	kgm^2
Compliance spring constant	k	217.0	$Nmrad^{-1}$
Compliance damping rate	ν	7.3	$Nmsrad^{-1}$
Car mass	M	1340.0	kg
Final drive ratio	r_1	0.275	-
Effective tyre radius	r_2	0.28	m
Air resistance coefficient	b	0.0001689	Nms^2rad^{-2}
Lower bound of engine torque	u_{1min}	0.0	Nm
Upper bound of engine torque	u_{1max}	100.0	Nm
Lower bound of clutch torque capacity	u_{3min}	0.0	Nm
Upper bound of clutch torque capacity	u_{3max}	225.0	Nm
Gear ratios	$\mathcal{U}_2$	$\{0.3, 0.6, 0.8, 1.0\}$	-

Table 2.1: Vehicle A: model parameters

Parameter	Symbol	Value	Units
Flywheel inertia	J_1	0.187	kgm^2
Gearbox input shaft inertia	J_2	0.0037	kgm^2
Gearbox output shaft inertia	J_3	0.0065	kgm^2
Inertia representing car mass	J_4	7.23	kgm^2
Compliance spring constant	k	529.22	$Nmrad^{-1}$
Compliance damping rate	ν	3.0	$Nmsrad^{-1}$
Car mass	M	1255.0	kg
Final drive ratio	r_1	0.262	-
Effective tyre radius	r_2	0.29	m
Air resistance coefficient	b	0.0001689	Nms^2rad^{-2}
Lower bound of engine torque	u_{1min}	0.0	Nm
Upper bound of engine torque	u_{1max}	150.0	Nm
Lower bound of clutch torque capacity	u_{3min}	0.0	Nm
Upper bound of clutch torque capacity	u_{3max}	250.0	Nm
Gear ratios	$\mathcal{U}_2$	$\{0.31, 0.47, 0.67...$ $..., 0.9, 1.17\}$	-

Table 2.2: Vehicle B: model parameters

Chapter 3

Optimal control problem solution

In the previous chapter, an optimal control problem was formulated. In this chapter, the optimal control problem is analysed with the aim of obtaining a numerical solution. Initial attempts at obtaining a solution for a simplified optimal control problem, motivates the creation of a related higher order problem. For this new problem, the location of such a solution is attempted by adapting existing gradient optimisation techniques to this constrained higher order problem. This involves designing a series of algorithms, tailored to the specific requirements of the clutch engagement problem, which carry out the individual tasks of the optimisation procedure calculating the numerical solution. The performance of this procedure is then analysed, for a variety of clutch engagement circumstances, assessing the ability of the algorithm to obtain a solution. Finally, the characteristics of the control strategies located by the algorithm are detailed. Some of the details in the chapter have been presented at a recent conference on modelling and control [62].

3.1 Optimal control problem analysis

For clarity, the previously defined optimal control problem under consideration is

$$\min_{\underline{u} \in \mathcal{U}} [F(\underline{u}) = \rho_k(\underline{w}, \underline{F} - \underline{c})] \text{ for } k \in \{1, 2, \infty\} \tag{3.1}$$

subject to

$$\frac{d\underline{x}}{dt} = f(\underline{x}, \underline{u}),\ \underline{x}(0) = \underline{x}_0 \tag{3.2}$$

$$\underline{R}(\underline{u}) = \underline{0} \tag{3.3}$$

$$\underline{Q}(\underline{u}) \leq \underline{0} \tag{3.4}$$

where $\underline{w}$ and $\underline{c}$ are vectors of weights and goals, ρ_k are as defined in equations (2.58), (2.61) and (2.62), $\mathcal{U}$ is defined by control upper and lower bounds, and f is continuously differentiable. In these equations, the functions $\underline{F}$, $\underline{R}$ and $\underline{Q}$ are vectors of functionals of the form

$$\int_0^{t_f} f_0(\underline{x}, \underline{u})dt + f_t(\underline{x}, \underline{u})|_{t=t_f} \tag{3.5}$$

with f_0 and f_t continuously differentiable. This optimal control problem does have a solution. This is due to the feasible cost space $\mathcal{F}$ being compact, as noted in section 2.4, together with the continuity of each ρ_k, which ensures that the image of $\mathcal{F}$, which is $\{\rho_k(\underline{w}, \underline{F} - \underline{c}) | \underline{F} \in \mathcal{F}\} \subset \Re$, is compact. This is enough to insure a solution, as this image is closed and bounded, insuring that an element of the image exists, which is also a lower bound.

The location of such a solution is a different matter. For optimal control problems defined using the weighted sum method (using ρ_1), Pontryagin's maximum principle provides some necessary conditions which must be satisfied for locally optimal control strategies. Solution techniques utilising these conditions for the general non-linear optimal control problem form described in equations (3.1) - (3.4), could be designed, but as a solution might be easier to obtain if particular characteristics of an optimal control problem representing the clutch engagement

problem is made use of. this simpler approach is first explored. For instance, after the simplification previously discussed, as a result of assumption (2.29), the state equations are almost linear, linear apart from the air resistance coefficient term which is negligible for low vehicle speeds. Another characteristic of the clutch engagement problem is that the components of the individual cost functionals are quadratic. That is, the cost integrands and terminal functions can be expressed in the form

$$\left[\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} + \begin{pmatrix} \underline{x} & \underline{u} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\right] \begin{pmatrix} \underline{x} \\ \underline{u} \end{pmatrix}. \tag{3.6}$$

These two observations on the optimal control problem form suggests that the problem might be a standard LQR problem. However, this is not the case, even when the weighted sum method is used to express the multi-objective compromise as: their are mandatory constraints; the linear components q_1 and q_2 are not necessarily zero and can not always be nullified by linear transformations; not all of the matrices $\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$ in the cost components are positive semi-definite; and not all of the matrices Q_{22} in the cost components are positive definite. This last point arises from the observation that the controls are linear in most of the individual cost components. Even if this LQR approach is not applicable, use might be made of the linearity of the controls. From Pontryagin's maximum principle, the linearity of the controls in the problem ensures that the optimal solution to an optimal control problem, designed using the weighted sum method, is bang-bang. That is, the controls flip between their upper and lower bounds, as between their bounds the necessary conditions of the maximum principle can not be satisfied. With the assumption that the optimal solution is bang-bang, then the optimal control problem is reduced to a standard optimisation problem

$$J : \Re^n \mapsto \Re \tag{3.7}$$

where n is the number of banging points and J is defined by the original optimal

control problem. The only difficulty is that the number of banging points is not known.

The solution of the optimal control problem is not trivial, so as an initial investigation of solution procedures, a simple case is taken. This case is an optimal control problem, designed using the weighted sum method, with only one constraint, the mandatory constraint equation (2.53). A further simplification is that the engine torque is taken as a constant, ensuring that the only remaining control is the required clutch torque capacity, and that the terminal terms of the costs are no longer directly dependent on the controls. In the design of this simplified optimal control problem, the weights of the compromise are chosen by trial and error, with only a handful of costs having non-zero weights. The costs chosen to have non-zero weights are the costs measuring the clutch energy dissipation, the clutch lock up time and the excitation of the oscillations at the point of lock up, which do not have quadratic control terms. These three costs do not measure all the undesirable effects, for instance they do not measure the transient oscillations, only the excitation of the oscillation at the point of lock up. However, they do conflict with the costs measuring the clutch energy dissipation and lock up time, small, when the clutch engages quickly, whilst the oscillation excitation cost is small when the clutch engages slowly. A reasonable set of weights for these costs is detailed in table 3.1. In solving this simplified problem, two different solution procedures were tried, one using the assumption that the optimal solution is bang-bang, and the other by iterating on the necessary conditions of the maximum principle.

For the first solution attempt, it is assumed that the optimal solution is bang-bang with just the one banging point, banging from its maximum value to its minimum value at a stipulated banging time. Hence an optimisation problem

$$J : \Re \mapsto \Re \tag{3.8}$$

results. With this assumption the solution procedure proceeds as follows:

1. Given an initial banging time, calculate the control and hence, calculate the model states, the terminal time and the performance measure by simulating the powertrain model, with the terminal time taken to be the time when the clutch plate speeds first become equal.

2. For two variations about either side of the initial banging time, calculate the performance measure values for these points using the same procedure as in step (1).

3. (a) If one of the variations yields a performance measure value smaller than the initial value, then take steps in that direction until the performance measure begins to increase again.

   (b) If both of the variations produce higher performance measure values, then take smaller variations and continue from step (2)

4. Obtain an interval $[a, b]$, which must contain a locally optimal solution, one interval bound being the last banging time reached (whose corresponding performance measure value has just increased from the last value), and the other bound being the banging time reached two steps ago.

5. Obtain a locally optimal solution, using a standard interval halving procedure.

A similar algorithm was also experimented with which tries to cope with more banging points. This operated as before with very small variations being taken about the initial banging times in order to numerically estimate the derivative of the performance measure with respect to the banging times. The step by step search was then conducted along the direction of steepest descent using a one dimensional search algorithm similar to the algorithm just discussed. On location

of a reasonable estimate for the one dimensional local minimum, the procedure is then repeated. This procedure was never very successful, as computational requirements rapidly increased with the number of banging times, and convergence of the procedure was poor.

The second solution technique uses the necessary conditions of the maximum principle. For the special case under consideration, the optimal control problem is of the form

$$\min_{\underline{u} \in \mathcal{U}} [F(\underline{u}) = \int_0^{t_f} f_0(\underline{x}, \underline{u})dt + f_t(\underline{x})|_{t=t_f}] \tag{3.9}$$

subject to

$$\frac{d\underline{x}}{dt} = f(\underline{x}, \underline{u}), \underline{x}(0) = \underline{x}_0 \tag{3.10}$$

$$r(\underline{x})|_{t_f} = \underline{0} \tag{3.11}$$

where f, f_0 and f_t are continuously differentiable. For this special case, the corresponding necessary conditions for a locally optimal solution u^*, with x^* the resulting optimal states are:

$$(1) \quad \dot{p}^* = \left.\frac{\partial H}{\partial x}\right|_* \tag{3.12}$$

$$(2) \quad H(x^*, p^*, u^*) \leq H(x^*, p^*, u) \text{ for all } u \in \mathcal{U} \tag{3.13}$$

$$(3) \quad H^*(t) = H^*(t_f) = 0 \tag{3.14}$$

$$(4) \quad p^*(t_f) = \alpha \left.\frac{\partial r}{\partial x}\right|_{*t_f} + \left.\frac{\partial f_t}{\partial x}\right|_{*t_f} \tag{3.15}$$

where $H = f_0(\underline{x}, \underline{u}) + p^T f(\underline{x}, \underline{u})$

as detailed in [63]. The solution attempt then proceeded as follows:

1. Given an initial control strategy, again calculate the states and terminal time t_f by simulating the model.

2. Calculate the costates by simulating backwards in time, which are defined by conditions (1) and (4) of the necessary conditions, where the constant

α is calculated using the third conditions $H^*(t_f) = 0$, which results in $\alpha = -\left.\frac{f_0 + \frac{\partial f_x}{\partial x}^T f}{\frac{\partial f_x}{\partial x}^T f}\right|_{*t_f}$

3. Calculate H_u from the states and costates.

4. Choose as a new control strategy $u_1 = u_0 + \lambda H_u$, where λ is a relaxation parameter, apply the control bounds, and repeat from step (1).

From examining the necessary conditions and the solution procedure, it can be noted that a locally optimal solution which satisfies the necessary conditions is a fixed point in the iterative solution procedure. However, as the controls are linear in the optimal control problem, the necessary conditions are never satisfied, with the controls moving towards their bounds as the controls converge to a solution.

Both of these algorithms run into problems due to the free terminal time t_f. This is partly due to the cost $F_4 = (\dot{x}_2 - \dot{x}_1)|_{t_f}$, which is desired to be small, which when small, increases the sensitivity of the terminal time. Indeed it can be shown that

$$\delta t_f = \left.\frac{\delta x_2 - \delta x_1}{\dot{x}_2 - \dot{x}_1}\right|_{t_f} \tag{3.16}$$

where dt_f is the variation of the terminal time and δx_1 and δx_2 are variations of two states at the terminal time. This sensitivity can, and often did, cause the terminal time to become undefined as the clutch plates speeds fail to become equal. As an example of this problem, using the first solution procedure applied to vehicle A, the data detailed in chapter 2, the algorithm might proceed as in figures 3.1 and 3.2. In these figures, simulations are shown for four different banging times, as the banging time is stepped down. For each of the banging times, a graph of the control is detailed, with the control banging down to zero at the predefined banging time, and returning to the control maximum value when the clutch plates lock up. This final increase in the required clutch torque capacity ensures that the clutch remains locked up after engagement, but does not affect

the dynamics of the powertrain as the clutch actuation mechanism is decoupled from the remainder of the powertrain after engagement. A graph of the clutch plate speeds is also shown. The values of the banging times, along with the corresponding performance measures and individual costs, are detailed in tables 3.2. From these figures and table, it can be seen that as the banging times decrease, the cost measuring clutch energy dissipation and the cost measuring lock up time both increase, whilst the cost measuring the excitation of the compliance at lock up decreases. The result is that the performance measure decreases. However, when the banging time steps down to 0.38, the terminal time becomes undefined, although the performance measure has yet to start increasing again. This problem can, of course, be prevented by taking small enough steps (in the example steps sizes of 0.1 seconds would suffice), but this increases the computational requirements of the solution procedure. Furthermore, if the weights in the optimal control problem are poorly chosen, the free terminal time problems are exacerbated.

For the more general solution procedure using the necessary conditions, the problems are if anything worse, with the relaxation parameter λ required to be reduced to prevent the terminal time becoming undefined. The adjustments to λ can be automated, by decreasing λ if the terminal time becomes undefined. The free terminal time also causes problems in this solution approach with difficulties in the recording of the control on the interval $[0, t_f]$. For a control strategy previously calculated using step (4) of the solution procedure, when returning to step (1), it is quite possible that the interval $[0, t_f]$ is exceeded before the clutch locks up. This requires the determination of suitable control values beyond t_f, perhaps by linear extrapolation, or even just by using the control value at t_f. This problem introduces further complications, that have to be managed, problems which are again exaggerated when the cost $F_4 = (\dot{x}_2 - \dot{x}_1)|_{t_f}$ is small.

The previous two solution attempts are far from state of the art, but they

demonstrate an important problem with the free terminal time t_f. This motivates a trick which removes this free terminal time, a trick that can be used to relate the necessary conditions for a time invariant fixed terminal time optimal control problem, to the necessary conditions for a time dependent free terminal time optimal control problem [63]. The trick proceeds as follows:

1. Introduce a new independent variable s.
2. Take t (time) to be a state with state equation $\frac{dt}{ds} = T^2(s)$, where T is a new 'dummy' control.
3. Let $t_f = t(s)|_{s=1}$.
4. Transform the original optimal control problem defined over the time interval $[0, t_f]$ to a new optimal control problem defined over a new interval $[0, 1]$, using the maps:

$$\int_0^{t_f} f_0(\underline{x}, \underline{u})dt \quad \mapsto \quad \int_0^1 T^2(s) f_0(\underline{x}, \underline{u})ds \tag{3.17}$$

$$\frac{d\underline{x}}{dt} = f(\underline{x}, \underline{u}) \quad \mapsto \quad \frac{d\underline{x}}{ds} = T^2(s) f(\underline{x}, \underline{u}) \tag{3.18}$$

This trick embeds the original problem into a higher order fixed optimal control problem, having an extra state and an extra control. The optimal control problem can now be envisaged as a map

$$\begin{aligned} F : \mathcal{D} \subset L^{\infty}_{m+1}([0,1]) \times \Re^m \quad &\mapsto \quad \Re \\ (\underline{u}(s), T(s), \underline{u}_f) \quad &\mapsto \quad F \end{aligned} \tag{3.19}$$

where $L^{\infty}_{m+1}([0,1])$ represents the set of bounded functions from the interval $[0,1]$ to $\Re^{m+1}$ (the controls and the 'dummy' control on the interval $[0,1]$), $\underline{u}_f$ represents $\underline{u}|_{s=1}$, and $\mathcal{D}$ is the domain of the map, defined by the set $\mathcal{U}$ and the constraints

$\underline{R}$ and $\underline{Q}$. It should be noted that the controls when $s = 1$ are independent from the other control values, the reason for the map being dependent on $\underline{u}_f$. The solution to the optimal control problem is now to locate an element of the domain $\mathcal{D}$ which minimises the value of the mapping, the value of F. For convenience, when referring to this higher order problem, u with no underscore represents an element of the domain $\mathcal{D}$, namely the column vector of $\underline{u}$, T and $\underline{u}_f$. The great advantage of this embedding is that the space, $L^{\infty}_{m+1}([0,1]) \times \Re^m$, in which the domain lies, is a Hilbert space, a complete normed vector space, with the norm arising from an inner product. Inner products are maps from two elements of the space to the real line with certain properties, and are extensions of the dot product in an n-dimensional Euclidean space. In the case of $L^{\infty}_{m+1}([0,1]) \times \Re^m$ an inner product can be defined by

$$\langle u_1, u_2 \rangle = \left\langle \begin{pmatrix} \underline{u}_1 \\ T_1 \\ \underline{u}_{1f} \end{pmatrix}, \begin{pmatrix} \underline{u}_2 \\ T_2 \\ \underline{u}_{2f} \end{pmatrix} \right\rangle = \int_0^1 \begin{pmatrix} \underline{u}_1 \\ T_1 \end{pmatrix}^T \begin{pmatrix} \underline{u}_2 \\ T_2 \end{pmatrix} ds + \underline{u}_{1f}^T \underline{u}_{2f} \tag{3.20}$$

with the norm arising from the inner product in the normal way by $||u||^2 = \langle u, u \rangle$, a measure of the distance of u from the zero. Hilbert space inner products, as well as defining the notion of distance of elements in the space, also allow notions such as orthogonality to be extended to such spaces. This in turn allows algebraic projection operations, allowing the optimal control problem to be treated as a conventional optimisation problem. The completeness of the normed vector space, allows differentiation to be extended to such higher order spaces, under certain conditions. Indeed, it can be shown, a trivial extension to the derivation found in [64], that for a functional of the form

$$F(u) = \int_0^1 T^2(s) f_0(\underline{x}, \underline{u}) ds + f_t(\underline{x}, \underline{u})|_{s=1} \tag{3.21}$$

$$\text{subject to} \qquad \frac{d\underline{x}}{ds} = T^2(s) f(\underline{x}, \underline{u}),\ \underline{x}(0) = \underline{x}_0 \tag{3.22}$$

where f, f_0 and f_t are continuously differentiable, then

$$\lim_{||\delta u|| \to 0} \frac{F(\hat{u}+\delta u) - F(\hat{u}) - \langle \nabla F_{\hat{u}}, \delta u \rangle}{||\delta u||} = 0 \qquad (3.23)$$

$$\text{with } \nabla F_{\hat{u}} = \begin{pmatrix} T^2 H_u \\ 2TH \\ f_{tu} \end{pmatrix}_{\hat{u},\hat{x},\hat{p}} \qquad (3.24)$$

where $H = (f_0 + p^T f)$ is the Hamiltonian which is dependent on the states, $\hat{x}$, corresponding to $\hat{u}$, and the costates, $\hat{p}$, which are calculated as usual from $\dot{p} = -T^2 H_x$ and $p(1) = f_{tx}$. Hence ∇F is the gradient of F with respect to the controls at $\hat{u}$. These calculations allow all of the individual cost functional and constraint gradients to be calculated. With this gradient information, the obvious solution technique is to use gradient methods, normally used in optimisation problems with the domains subsets of n-dimensional Euclidean spaces, but just as applicable here due to the properties of the Hilbert space in which the domain lies.

3.2 Gradient method algorithm

As discussed, this algorithm is an extension of a family of standard optimisation solution techniques for locating the local minima of maps of the form $\mathcal{F} : \Re^n \mapsto \Re$ to maps of the form $\mathcal{F} : \mathcal{D} \subset \mathcal{H} \mapsto \Re$, where $\mathcal{H}$ is a Hilbert space, which result from the embedding of the optimal control problem into a higher order problem. This extension is well known and is discused in [64] and [65]. As mentioned, with such problems two notions can be extended, differentiability and orthogonality used to project elements of the Hilbert space to vector subspaces. As the techniques are mearly an extension, the understanding of the algorithm is aided by representing the problem as a map from a 3-D Euclidean space to the real line and describing the algorithm in terms of the notions that can be extended. In this representation the equality constraints $\underline{R} = \underline{0}$ are represented by a surface in

$\Re^3$, and the inequality constraints $\underline{Q} \leq \underline{0}$ and $u \in \mathcal{U}$ are represented by another surface partitioning $\Re^3$. These constraints create the set of all admissible controls, or alternatively, the domain $\mathcal{D}$ of the mapping, with each point in the domain, and indeed other points outside the domain having a unique performance measure value.

With this representation, the algorithm, a standard gradient method algorithm, is as follows with a single iteration of the algorithm being demonstrated in figure 3.3, using the 3-D representation.

1. Calculate an initial control u_0, which is an element of the domain $\mathcal{D}$, ie a control that satisfies $u \in \mathcal{U}$, $\underline{R} = \underline{0}$ and $\underline{Q} \leq \underline{0}$, by projecting a point onto $\mathcal{D}$.

2. Calculate the gradient of the performance measure and the constraints at the current control $u_i \in \mathcal{D}$.

3. Calculate a search direction h from gradient information, which is a tangent to the domain at u_i and along which the performance measure value will initially decrease.

4. Carry out a one dimensional optimisation procedure along the projection of h onto the domain.

5. Let the solution of this 1-D optimisation procedure be u_{i+1} and repeat from step (2).

This algorithm produces a sequence of controls in the domain with monotonically decreasing performance measure values, a control being added to the sequence every time this main iteration, steps (2) - (5), are completed. The details of how each of these steps are carried out for the higher order optimal control problem

representing the clutch engagement problem now follows, with again u representing an element of $L^{\infty}_{m+1}([0,1]) \times \Re^m$.

3.2.1 Calculation of an initial control strategy

An initial control strategy $u_0 \in \mathcal{D} = \{u \in L^{\infty}_{m+1}([0,1]) \times \Re^m | u \in \mathcal{U}, \underline{R} = \underline{0}, \underline{Q} = \underline{0}\}$, is obtained by taking any control u_{-1} in the Hilbert space and projecting it to an estimate of the tangent space of the domain $\mathcal{D}$ obtained from gradient information at u_{-1}. This projection to $\mathcal{D}$ is very similar to the Newton-Raphson, or quasi-Newton, iterative algorithms for solving non-linear equations, with the gradients being obtained at the current estimate of the solution in order to estimate a better solution. For the projection to the domain, the equations to be solved are

$$\underline{R}(u) = \underline{0} \tag{3.25}$$

$$Q_{k_2}(u) = 0 \text{ for all } k_2 \text{ such that } Q_{k_2}(u_{-1}) > 0 \tag{3.26}$$

$$u_l(s) = \begin{cases} u_{lmin} & \text{if } u < u_{lmin} \\ u_{lmax} & \text{if } u > u_{lmax} \end{cases} \tag{3.27}$$

where u_l represents an element of the column vector u, and $u(s) = u_{-1}(s) + \sum_{k_1} \alpha_{k_1} \nabla R_{k_1}(s) + \sum_{k_2} \beta_{k_2} \nabla Q_{k_2}(s) + \gamma(s)$ with the added restrictions that

$$\beta_{k_2} = 0 \text{ if } Q_{k_2} < 0 \tag{3.28}$$

$$\gamma_l(s) = 0 \text{ if } u_l(s) \in (u_{lmin}, u_{lmax}), \tag{3.29}$$

ensuring that only projection to constraints that are not satisfied occurs. This equation is difficult to solve in one step, due to the difference in order of the finite dimensional constraints $\underline{R}$ and $\underline{Q}$, and the infinite dimensional constraints resulting from the control bounds. For this reason, the projection to the subspace defined by $\underline{R}$ and $\underline{Q}$, and the projection to the subspace defined by the control bounds, are performed separately. In the case of the projection to the subspace defined by $\underline{R}$ and $\underline{Q}$, the projection is obtained by solving equations (3.25) and

(3.26), where $u(s) = u_{-1}(s) + \sum_{k_1} \alpha_{k_1} \nabla R_{k_1}(s) + \sum_{k_2} \beta_{k_2} \nabla Q_{k_2}(s)$. The solution of these equations is obtained by using a first order approximation of the constraint functions and solving for the α's and the β's. For the projection in question, this results in the formula

$$u(s) = u_{-1}(s) - \begin{pmatrix} \nabla \underline{R}(s) \\ \nabla \underline{Q}_{K_2}(s) \end{pmatrix}^T \left\langle \begin{pmatrix} \nabla \underline{R}(s) \\ \nabla \underline{Q}_{K_2}(s) \end{pmatrix}^T, \begin{pmatrix} \nabla \underline{R}(s) \\ \nabla \underline{Q}_{K_2}(s) \end{pmatrix}^T \right\rangle_{vec} \begin{pmatrix} \underline{R}(s) \\ \underline{Q}^{-1}_{K_2}(s) \end{pmatrix}_{u_{-1}} \tag{3.30}$$

assuming that the gradients of the constraints are linearly independent, where $Q_{K_2} = \{Q_{k_2} | Q_{k_2}(u_{-1}) > 0\}$, the gradients are calculated at u_{-1} using the procedure described in the following section, and $\langle .,. \rangle_{vec}$ is an extension to the Hilbert space inner product with $\langle (u_1, \cdots, u_N), (v_1, \cdots, v_N) \rangle_{vec} = (\langle u_1, v_1 \rangle, \cdots, \langle u_N, v_N \rangle)$. The second projection to $\mathcal{U}$, solving equation (3.27), where $u(s) = u_{-1}(s) + \gamma(s)$, is equivalent to applying the bounds by truncating the values of u_{-1}.

The problem with taking the projections separately is that they interfere with each other, with one projection moving the control away from the subspace that the other projection is trying to project the control to. For the second projection, it is very difficult to carry out the applying of the bounds without moving the controls away from the subspace defined by the constraints $\underline{R}$ and $\underline{Q}$, as the dimensions of the constraints introduced by the bounds exceeds the dimensions of the finite dimensional constraints $\underline{R}$ and $\underline{Q}$. However the projection onto $\underline{R} = \underline{0}$ and $\underline{Q} \leq \underline{0}$ can be performed so that the bounds are not exceeded, or at least are not likely to be exceeded by a large amount. This is done by restricting the freedom of the controls, to controls that are not likely to exceed their bounds. In practice, this is done by introducing two m by m matrices $G_1(s)$ and G_2, with m the number of controls in the powertrain model, which indicate the controls which are likely to exceed their bounds, where the components of the matrices are

defined by

$$g_{1ij}(s) = \begin{cases} 1 & \text{if } i = j \text{ and } \underline{u}_i(s) \text{ is unlikely to exceed it's bounds} \\ 0 & \text{otherwise} \end{cases} \tag{3.31}$$

$$g_{2ij} = \begin{cases} 1 & \text{if } i = j \text{ and } \underline{u}_i(1) \text{ is unlikely to exceed it's bounds} \\ 0 & \text{otherwise} \end{cases} . \tag{3.32}$$

with $\underline{u}_i$ representing and element of the original set of controls in the optimal control problem. These matrices can be used to filter an element, or elements, of the Hilbert space to its components, which are unlikely to exceed the control bounds by

$$G(u_1, \cdots, u_n)^T = G \begin{pmatrix} \underline{u}_1(s) & \cdots & \underline{u}_n(s) \\ T_1 & \cdots & T_n \\ \underline{u}_{1f} & \cdots & \underline{u}_{nf} \end{pmatrix}^T = \begin{pmatrix} G_1(s)\underline{u}_1(s) & \cdots & G_1(s)\underline{u}_n(s) \\ T_1 & \cdots & T_n \\ G_2\underline{u}_{1f} & \cdots & G_2\underline{u}_{nf} \end{pmatrix}^T \tag{3.33}$$

This filter can be used to solve equations (3.25) and (3.26), but this time with

$$u(s) = u_{-1}(s) + G \begin{pmatrix} \nabla \underline{R} \\ \nabla \underline{Q} \end{pmatrix}^T \begin{pmatrix} \underline{\alpha} \\ \underline{\beta} \end{pmatrix}, \tag{3.34}$$

resulting in the new formula

$$u(s) = u_{-1}(s) - G \begin{pmatrix} \nabla \underline{R}(s) \\ \nabla \underline{Q}_{K_2}(s) \end{pmatrix}^T \left\langle \begin{pmatrix} \nabla \underline{R}(s) \\ \nabla \underline{Q}_{K_2}(s) \end{pmatrix}^T, G \begin{pmatrix} \nabla \underline{R}(s) \\ \nabla \underline{Q}_{K_2}(s) \end{pmatrix}^T \right\rangle_{vec}^{-1} \begin{pmatrix} \underline{R}(s) \\ \underline{Q}_{K_2}(s) \end{pmatrix}_{u_{-1}} \tag{3.35}$$

for the projection of the controls to the subspace defined by the constraints. In practice the likelihood of whether the controls will invalidate their constraints is determined by whether they did on the previous iteration of the projection algorithm. Hence, the above analysis results in an algorithm for calculating an initial control $u_0 \in \mathcal{D}$, which is as follows

1. Given any initial control u_{-1}, define $Q_{K_2} = \{Q_{k_2} | Q_{k_2}(u_{-1}) > 0\}$ and $G_1(s) = G_2 = I_m$, an m by m identity matrix, equating to no control restrictions.

2. Calculate a new control u using formula (3.35)

3. Calculate a new filter using formulae (3.31) and (3.32), with the likelihood of u exceeding it's bounds being determined by whether the new u from step (2), has exceeded its bounds.

4. Redefine $Q_{K_2} = \{Q_{k_2} | Q_{k_2}(u) > 0\}$ for the new control u.

5. Let $u_{-1} = B(u)$, where B(.) is a function that applies the control bounds, and repeat from step (2).

It can be seen that any control $u \in \mathcal{D}$, will be a stationary point of the iteration, steps (2) - (4). In practice as the algorithm, converges to such a point, a tolerance, governing when a control is sufficiently close to $\mathcal{D}$ to be said to be in $\mathcal{D}$, must be found. One way of measuring this closeness is to measure the distance the control moves under each projection. For the projection carried out in step (2), the square of this projection distance is

$$E_1 = ||u - u_{-1}||^2 \tag{3.36}$$

which is evaluated after step (2) of the algorithm. For the second projection, the square of the projection distance is

$$E_2 = ||B(u) - u||^2 \tag{3.37}$$

evaluated at step (4). Another measure of how close a control is to the domain is to look at the individual error of the constraints $\underline{R}$, $\underline{Q}$ and, possibly, of individual controls at each point in s, if outside their bounds. The later measure was not felt to be too important, as the measurement, E_2, in equation 3.37, provides a

meaningful error assessment of the control excursion outside its bounds. A final measurement, measuring how close the controls are to the domain, $\mathcal{D}$, in terms of how the projection affects the performance measure value is

$$E_3 = \frac{\langle \nabla F, B(u) - u_{-1} \rangle}{F}. \tag{3.38}$$

Hence, the termination condition indicating a satisfactory control is, if

$$\sqrt{E_1 + E_2} < d_{max1} \tag{3.39}$$

$$E_3 < d_{max2} \tag{3.40}$$

$$-\epsilon_{k_1} < R_{k_1} < \epsilon_{k_1} \tag{3.41}$$

$$Q_{k_2} < \delta_{k_2} \tag{3.42}$$

are all satisfied, where d_{max1}, d_{max2}, ϵ_{k_1} and δ_{k_2} are small positive tolerances, with the constraints values R_{k_1} and Q_{k_2} evaluated at step (5). Conditions (3.41) and (3.42) can be interpreted as determining whether the constraints are satisfied. In practice, no convergence problems have been encountered with this algorithm, with the algorithm seeming to inherit the quadratic convergence of the Newton-Raphson algorithm.

3.2.2 Calculation of the cost and constraint gradients

As mentioned, the gradients of the performance measure components and the individual constraint functions, which are of the form (3.21) and (3.22), can be calculated using formula (3.24). For clarity, the gradients at $\hat{u}$ are evaluated using this formula with the following procedure:

1. For a control $\hat{u}$, by simulating the transformed model equations, calculate the corresponding states $\hat{x}$ on the interval $[0, 1]$.

2. Using the differential equations $\dot{p}^T = T^2 H_x = T^2(f_{0x} + p_T f_x)$ and $p^T(1) = f_{tx}$, by simulating backward in time, calculate the corresponding costates,

$\hat{p}$. from the controls and states.

3. Evaluate $\nabla F = \begin{pmatrix} T^2 H_u \\ 2TH \\ F_{tu} \end{pmatrix} = \begin{pmatrix} T^2(f_{0u} + p^T f_u) \\ 2T(f_0 + p^T f) \\ F_{tu} \end{pmatrix}\Bigg|_{\hat{u},\hat{x},\hat{p}}$.

This procedure directly provides the required gradient information for the constraints and the cost components. The cost component gradients can, in certain cases provide the gradient of the performance measure. In particular, it can be shown that for the weighted sum multi-objective method (ρ_1) then

$$\lim_{||\delta u|| \to 0} \frac{\rho_1(\underline{w}, \underline{F}(\hat{u} + \delta u)) - \rho_1(\underline{w}, \underline{F}(\hat{u})) - \langle \rho_1(\underline{w}, \nabla \underline{F}_{\hat{u}}), \delta u \rangle}{||\delta u||} = 0 \qquad (3.43)$$

implying that

$$\nabla \rho_1(\hat{u}) = \rho_1(\underline{w}, \nabla \underline{F}_{\hat{u}}) = \sum_{k=1}^{n} w_k \nabla F_k \qquad (3.44)$$

is the gradient of the performance measure at $\hat{u}$, where F_k are the individual cost functionals and w_k are their corresponding weights. Similarly, for the quadratic goal attainment method, when $F = \rho_2(\underline{w}, \underline{F} - \underline{c}) > 0$, it can be shown that the gradient of the performance measure is

$$\nabla \rho_2(\hat{u}) = \frac{\sum_{k=1}^{n} w_k^2 \nabla F_k \max_k \{F_k - c_k\}}{F} \qquad (3.45)$$

where F_k and w_k are as previously noted and c_k are the goals corresponding to F_k. In the region where $F \leq 0$, F is equal to the standard goal attainment method. Unfortunately, for this method, where $F = \rho_\infty(\underline{w}, \underline{F} - \underline{c})$, no unique value $\nabla \rho_\infty$ can be found, so that

$$\lim_{||\delta u|| \to 0} \frac{\rho_\infty(\underline{w}, \underline{F}(\hat{u} + \delta u) - \underline{c}) - \rho_\infty(\underline{w}, \underline{F}(\hat{u}) - \underline{c}) - \langle \nabla \rho_\infty, \delta u \rangle}{||\delta u||} = 0 \qquad (3.46)$$

as the limit is dependent on the direction that δu tends to zero. This lack of Frechet differentiability can be overcome by introducing a weaker form of directional, or Gateaux, differentiation [66], where the derivative is now dependent on

the direction that δu tends to zero. Indeed, it can be shown that, for the goal attainment method,

$$\lim_{\epsilon \to 0} \frac{\rho_\infty(\underline{w}, \underline{F}(\hat{u} + \epsilon h) - \underline{c}) - \rho_\infty(\underline{w}, \underline{F}(\hat{u}) - \underline{c}) - \langle \nabla \rho_\infty(\hat{u}, h), \epsilon h \rangle}{\epsilon} = 0 \quad (3.47)$$

with the directional gradient defined by

$$\langle \nabla \rho_\infty(\hat{u}, h), \lambda h \rangle = \lambda \max_{F_k = F} \{\langle \nabla F_k, h \rangle\} \quad (3.48)$$

where $h \in \mathcal{H}$ such that $||h|| = 1$, defines the direction of the limit to zero.

3.2.3 Calculation of a search direction

A search direction $h \in \mathcal{H} = L^\infty_{m+1}([0,1]) \times \Re^m$ is calculated which is a tangent to the domain $\mathcal{D}$ at u_i, and along which the performance measure initially decreases. The term tangent to $\mathcal{D}$ is used loosely, and refers to a search direction along which the linear approximations of the constraints defining $\mathcal{D}$ at u_i are satisfied. Algebraically, these requirements can be expressed by the conditions

$$\langle \nabla F_{u_i}, h \rangle < 0 \quad (3.49)$$

$$h \in \mathcal{T}_i \quad (3.50)$$

where ∇F_{u_i} is the performance measure gradient at u_i, and $\mathcal{T}_i$ is the 'tangent' subspace of $\mathcal{D}$ at u_i, which is defined by $h \in \mathcal{T}_i$ if

$$\langle \nabla R_{k_1}, h \rangle = 0 \text{ for all } k_1 \quad (3.51)$$

$$\langle \nabla Q_{k_2}, h \rangle \leq 0 \text{ for all } k_2 \text{ such that } Q_{k_2} = 0 \quad (3.52)$$

$$h_l(s) \geq 0 \text{ for all } l, s \text{ such that } u_l(s) = u_{l_{min}} \quad (3.53)$$

$$h_l(s) \leq 0 \text{ for all } l, s \text{ such that } u_l(s) = u_{l_{max}} \quad (3.54)$$

where the gradients are evaluated at u_i, with h_l and u_l representing elements of the column vectors h and u.

One way to obtain such a search direction is to choose an element of $\mathcal{T}_i$ along which the rate of decrease of the performance measure is steepest, the so called 'steepest decent method'. For performance measures that are Frechet differentiable, the direction of steepest decent in $\mathcal{H}$ is just the negative of the gradient. For the case when the goal attainment method has been used (ρ_∞), this steepest decent direction in $\mathcal{H}$ is a linear combination of the cost component gradients which have cost values equal to the performance measure value (i.e. $F_k = F$). With this knowledge, the steepest descent direction in $\mathcal{T}_i$ is taken to be the projection of the steepest decent direction form in $\mathcal{H}$, onto $\mathcal{T}_i$. As the projection of the negative gradient onto $\mathcal{T}_i$ is obviously a subcase of the projection of a linear combination of gradients, only the later is detailed. In this general case, not only have the projection parameters to be calculated, but also the linear combination of the cost components. This linear combination is evaluated by taking $\langle \nabla F_{l_1}, h \rangle = \langle \nabla F_{l_2}, h \rangle$ for all $l_1 \neq l_2$ such that $F_{l_1} = F_{l_2} = F$, with h the final projected search direction. If this is not the case, then a new steeper decent direction can be found along which the smaller values of $\langle \nabla F_l, h \rangle$ have increased and the larger values decreased. These equations, along with the projection equations, result in the problem of locating an h, satisfying

$$\langle \nabla R_{k_1}, h \rangle = 0 \text{ for all } k_1 \tag{3.55}$$

$$\langle \nabla Q_{k_2}, h \rangle \leq 0 \text{ for all } k_2 \text{ such that } Q_{k_2} = 0 \tag{3.56}$$

$$\langle \nabla F_{k_3}, h \rangle = \kappa \text{ for all } k_3 \tag{3.57}$$

$$h_l(s) \leq 0 \text{ for all } l, s \text{ such that } u_l(s) = u_{lmax} \tag{3.58}$$

$$h_l(s) \geq 0 \text{ for all } l, s \text{ such that } u_l(s) = u_{lmin} \tag{3.59}$$

where κ, an arbitrary negative number, determines the magnitude of the search direction $h(s) = \sum_{k_1} \alpha_{k_1} \nabla R_{k_1} + \sum_{k_2} \beta_{k_2} \nabla Q_{k_2} + \sum_{k_3} \zeta_{k_3} \nabla F_{k_3} + r(s)$ with the pa-

rameters restricted by

$$\beta_{k_2} = 0 \quad \text{if } \langle \nabla Q_{k_2}, h \rangle < 0 \text{ or } Q_{k_2} < 0 \tag{3.60}$$

$$\zeta_{k_3} = 0 \quad \text{if } F_{k_3} < F \tag{3.61}$$

$$r_l(s) = 0 \quad \text{if } u_l(s) \in (u_{lmin}, u_{lmax}) \tag{3.62}$$
$$\text{or } u_l(s) = u_{lmin} \text{ and } h_l(s) > 0$$
$$\text{or } u_l(s) = u_{lmax} \text{ and } h_l(s) < 0$$

These restrictions ensure that the projection to the inequality constraint boundaries only occurs if the search direction invalidates the constraints. These equations are very similar to the equations for projecting a control onto the admissible control space, only this time a search direction is being projected onto a 'tangent' space. As a result, a similar algorithm to the algorithm described in section 3.2.1 is used to effect the calculation of the search direction, similarly assuming that the gradients of the constraints and the cost components are linearly independent. This algorithm is as follows:

1. Calculate $\underline{F}_{K_3} = \{F_{k_3} | F_{k_3} = F\}$, and a target set for $\left\langle \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \end{pmatrix}^T, h \right\rangle_{vec}$ $Z = (\overbrace{\kappa, \cdots, \kappa}^{\text{re.}\underline{F}_{K_3}}, \overbrace{0, \cdots, 0}^{\text{re.}\underline{R}})^T$ where the κ's are the targets for the cost components and the zeros are the targets of the equality constraints.

2. Calculate an initial estimate of h assuming that the inequality constraints and bounds are not exceeded, using

$$h = \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \end{pmatrix}^T \left\langle \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \end{pmatrix}^T, \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \end{pmatrix}^T \right\rangle_{vec}^{-1} Z \tag{3.63}$$

3. Calculate a filter, as in equation (3.33), but with

$$g_{1ij}(s) = \begin{cases} 1 & \text{if } i = j \text{ and } u_l(s) + \epsilon_{step} h_l(s) \in (u_{lmin}, u_{lmax}) \\ 0 & \text{otherwise} \end{cases} \tag{3.64}$$

$$g_{2ij} = \begin{cases} 1 & \text{if } i = j \text{ and } u_l(1) + \epsilon_{step} h_l(1) \in (u_{lmin}, u_{lmax}) \\ 0 & \text{otherwise} \end{cases} \tag{3.65}$$

where ϵ_{step} is a small step size, comparable to the initial step size of the following 1-D optimisation procedure

4. Calculate $Q_{K_2} = \{Q_{k_2} | Q_{k_2} = 0 \text{ and } \langle \nabla Q_{k_2}, h \rangle \geq 0\}$ and calculate a new target $Z = (\overbrace{\kappa, \cdots, \kappa}^{\text{re.}\underline{F}_{K_3}}, \overbrace{0, \cdots, 0}^{\text{re.}\underline{R}}, \overbrace{0, \cdots, 0}^{\text{re.}\underline{Q}_{K_2}})^T$

5. Calculate a new search direction using

$$h_{new} = h_{old} - G \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \\ \nabla \underline{Q}_{K_2} \end{pmatrix}^T \left\langle \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \\ \nabla \underline{Q}_{K_2} \end{pmatrix}^T, G \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \\ \nabla \underline{Q}_{K_2} \end{pmatrix}^{T-1} \right\rangle_{vec} \left[\left\langle \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \\ \nabla \underline{Q}_{K_2} \end{pmatrix}^T, h_{old} \right\rangle_{vec} - Z \right] \tag{3.66}$$

6. Test for convergence and repeat from step (3).

Note that, in practice, in this algorithm the equality conditions are relaxed, by taking them as satisfied, if they are satisfied within a given tolerance bound. In particular

$$Q_{k_2} = 0 \qquad \text{if } -\delta_{k_2} < Q_{k_2} < \delta_{k_2} \tag{3.67}$$

$$F_{k_3} = F \qquad \text{if } F_{k_3} > F(1 - \epsilon_{cost}) \tag{3.68}$$

as in equations (3.41) and (3.42). The value κ is chosen so as to normalise the calculated search direction, to prevent its magnitude affecting the 1-D optimisation search. The convergence criteria of the algorithm is when the maximum norm of the difference between $\left\langle \begin{pmatrix} \nabla \underline{F}_{K_3} \\ \nabla \underline{R} \\ \nabla \underline{Q}_{K_2} \end{pmatrix}, h \right\rangle$ and its target Z is less than a certain tolerance d_{max3}. Some convergence problems have been encountered with this algorithm, especially when F_{K_3} is large and its components are nearly linearly

dependent. Generally, this only occurs when a 'good' solution has been located by the algorithm, or the problem has been poorly posed.

The steepest decent method is a first order gradient technique and, hence, may be slow to obtain a solution. As a result, for cases when the gradient is Frechet differentiable and second derivatives exist, which is when the multi-objective methods using ρ_1 or ρ_2 have been used, second order techniques, such as conjugate gradient techniques have been investigated. In particular, one such technique which does not require the direct evaluation of the second derivative, referred to as the Fletcher-Reeves algorithm, can be modified slightly in order to apply it to the problem under consideration. With this technique, the search direction is dependent on an iterative algorithm and is calculated as follows.

1. Let $h_0 = g_0 = P(-\nabla F_{u_0})$
2. Carry out the 1-D search direction along the projection of h_i onto $\mathcal{D}$, in order to calculate the local minimum u_{i+1}
3. Let $h_{i+1} = g_{i+1} + \lambda(h_i + p_i)$ where $g_i = P(-\nabla F_{u_i})$ and $\lambda = \frac{||g_{i+1}||^2}{||g_i||^2}$.
4. Repeat from step (2)

with $P(.)$ the projection onto the tangent space operation just described and $p_i = \sum_{k_1} \alpha_{k_1} \nabla R_{k_1} + \sum_{k_2} \beta_{k_2} \nabla Q_{k_2} + r(s)$ is a projection element, with similar restrictions to equations (3.60) and (3.62), calculated during the 1-D optimisation procedure, so that $h_i + p_i \in \mathcal{T}_{i+1}$. Now, as g_{i+1} and $h_i + p_i$ are both elements of the 'tangent' space and as λ is positive, h_{i+1} is also an element of this 'tangent' space. Furthermore, it can be shown that $\left\langle \nabla F_{u_{i+1}}, h_{i+1} \right\rangle = -\alpha ||g_{i+1}||^2$, where α is a positive constant, ensuring that conditions (3.49) - (3.50) are satisfied, that is unless a local minima has been located.

3.2.4 1-D optimisation along the search direction projection

Now a search direction has been calculated, a local one dimensional search along the projection of this search direction onto the admissible control space $\mathcal{D}$ is conducted. This is done by taking a step along this search direction and projecting the new control values to $\mathcal{D}$ in order to calculate a segment of the projected search direction on which the local optimal solution must lie. An interval halving routine is then used to refine the segment on which a local minima lies, in order to calculate this local minima.

Steps along the projection of the search direction onto $\mathcal{D}$ are calculated from an initial control $u_i^{(0)}$ using

$$u_i^{(1)} = P(u_i^{(0)} + \epsilon_{step} h_i) \tag{3.69}$$

$$u_i^{(j+1)} = P(u_i^{(j)} + \lambda_{up}(u_i^{(j)} - u_i^{(j-1)})) \tag{3.70}$$

where h_i is the search direction calculated in the previous section, ϵ_{step} is a small positive constant governing the size of the steps, λ_{up} is a positive constant slightly bigger than unity causing the step size to increase along the search direction and $P(.)$ is the projection mapping described in section 3.2.1, but only projecting onto the inequality constraints of which the initial control $u_i^{(0)}$ is on the boundary (i.e. inequality constraints satisfying $Q_{k_2}(u_i^{(0)}) = 0$). In practice, one iteration of the algorithm was normally sufficient to achieve a satisfactory projection. From iteration to iteration the magnitude of ϵ_{step} is adjusted by multiplying it by λ_{up}^j, where j is the number of steps of the 1-D optimisation procedure on the previous iteration.

The above procedure creates a sequence of controls along the projection of h_i, for which corresponding performance measures can be calculated by simulation of the model equations and evaluation of the cost functionals. With this calculated

sequence of performance measures, there are a number of possible occurrences requiring different responses. These are as follows:

1. If the performance measure initially increases ($F(u_i^{(0)}) < F(u_i^{(1)})$), then decrease ϵ_{step} by dividing it by $\lambda_{down} > 1$ and restart the calculation of the sequence. Note that, from equation (3.49), if ϵ is made sufficiently small, the performance measure will decrease.

2. If at any point in the sequence the performance measure begins to rise again with the controls remaining in $\mathcal{D}$, then locate a segment of the projection of h_i on which a local 1-D minima must lie. In particular, with $F(u_i^{(j+1)}) > F(u_i^{(j)})$ then the local minima must be an element of $\mathcal{V} = \{P((1-\mu)u_i^{(j-1)} + \mu u_i^{(j+1)} | \mu \in [0,1]\}$.

3. If at a point in the sequence any inequality constraint which is not projected to, and hence prior to the 1-D optimisation procedure satisfies $Q_{k_2}(u_i^{(0)}) < 0$, is no longer satisfied, stop the sequence and calculate the point along the projection of h_i at which the constraints are initially invalidated and replace the last step which does not satisfy the constraints by this new control. If the performance measure is still decreasing then stop, taking the control that just satisfies the constraints to be the local minima and continue with a new iteration of the algorithm, otherwise obtain a segment as in step (2) on which the local minima must lie.

This later calculation, of a point along the the projection of h_i, which just satisfies the inequality constraints, is obtained using the procedure

1. Given $u_i^{(j)} \in \mathcal{D}$ and $u_i^{(j+1)} \notin \mathcal{D}$, as it invalidates one or more inequality constraints, let $u' = P((1-\mu)u_i^{(j)} + \mu u_i^{(j+1)})$ with $\mu = \min_{k_2} \left\{ -\frac{Q_{k_2}(u_i^{(j)})}{Q_{k_2}(u_i^{(j+1)}) - Q_{k_2}(u_i^{(j)})} \right\}$. This formula results from approximating the inequality constraint values by

linear interpolation between the two given values and calculating the interpolation parameter μ which just satisfies the constraints. This parameter is then related back to a linear interpolation between the controls.

2. (a) If $u' \in \mathcal{D}$ then replace $u_i^{(j)}$ by u' and repeat from step (1).

   (b) If $u' \notin \mathcal{D}$, as it still does not satisfy all the inequality constraints, then replace $u_i^{(j+1)}$ by u' and repeat from step (1).

This procedure is repeated until the all the inequality constraints are satisfied within a given tolerance. That is

$$Q_{k_2} < \delta_{k_2} \qquad \text{for all } k_2 \tag{3.71}$$

$$-\delta_k < Q_k < \delta_k \qquad \text{for a particular } k. \tag{3.72}$$

Once a segment, on the projection of h, on which a local minima must lie, is located, the local minima is evaluated using an interval halving procedure. This procedure is as follows:

1. Given a segment $\mathcal{V} = \{P((1-\mu)u_i^{(j-1)} + \mu u_i^{(j+1)}) | \mu \in [0,1]\}$ on which the minima lies, $u_i^{(j)} \in \mathcal{V} \setminus \{u_i^{(j-1)}, u_i^{(j+1)}\}$, and performance measure values for all three points, $F(u_i^{(j-1)})$, $F(u_i^{(j)})$, and $F(u_i^{(j+1)})$, calculate

$$u' = P\left(\frac{u_i^{(j-1)} + u_i^{(j)}}{2}\right) \tag{3.73}$$

$$u'' = P\left(\frac{u_i^{(j)} + u_i^{(j+1)}}{2}\right) \tag{3.74}$$

   and their performance measure values $F(u')$, $F(u'')$.

2. (a) If $F(u') < F(u_i^{(j)})$, then let $u_i^{(j)} = u'$, $u_i^{(j+1)} = u_i^{(j)}$ and repeat from step (1)

   (b) If $F(u'') < F(u_i^{(j)})$, then let $u_i^{(j)} = u''$, $u_i^{(j-1)} = u_i^{(j)}$ and repeat from step (1)

(c) Otherwise, let $u_i^{(j-1)} = u'$, $u_i^{(j+1)} = u''$ and repeat from step (1)

This procedure is repeated until this difference between the largest and the smallest value of $F(u_i^{(j-1)})$, $F(u_i^{(j)})$ $F(u_i^{(j+1)})$ is less than a given tolerance. This tolerance is taken to be the minimum of an absolute value ΔF and a relative value, $\rho\times$ initial difference of performance measures.

3.2.5 Implementation of the gradient method algorithm

The gradient method algorithm just described is implemented using the computer packages Matlab and ACSL. Matlab, a matrix manipulation package, was used to perform the majority of the computational tasks, with the simulation package ACSL used to simulate the states and costates, and to evaluate the individual cost and constraint functionals.

In performing the algebraic operations of the algorithm in Matlab, elements of the Hilbert space must be stored and the Hilbert space inner product must be evaluated. The elements of the infinite dimensional Hilbert space, $L^{\infty}_{m+1}([0,1]) \times \Re^m$, which are represented by a vector $\begin{pmatrix} \underline{u} \\ \underline{T} \\ \underline{u_f} \end{pmatrix}$ with each component of the vector either in $\Re$ or $L^{\infty}([0,1])$, is approximated by representing the components in $L^{\infty}([0,1])$ by a further vector equating to values of the component at regular intervals in s. Typically 100 points are taken with such a component $u_l(s) \in L^{\infty}([0,1])$ represented by the vector $[u_l(0), u_l(0.01), u_l(0.02), \cdots, u_l(1)]$. The evaluation of the Hilbert space inner product defined by equation (3.20), is accomplished using the matrix operations provided in Matlab and evaluating the integral in the equation using the trapezium rule. All of the other computations required can be easily performed using standard matrix manipulations provided by Matlab.

The simulation of the states and costates in ACSL is performed using the Runge-Kutta 4th order integration procedure, with an integration step size of up to 0.01 for the independent variable s. The results of these simulations are

dumped in ascii form into files at intervals in s, corresponding to the interval on which the values of elements in $L^{\infty}([0,1])$ are approximated in Matlab, allowing these values to be read into Matlab with out too much difficulty.

In order to aid the performance of the algorithm, the controls, the cost functionals and the constraints are normalised in the described computation. The controls are normalised so as to ensure that the importance of one control on the solution procedure is similar to the other controls. If the controls were not normalised, then one control would converge faster than the other controls, slowing down the overall convergence of the algorithm. The costs and constraints are normalised so as to prevent ill-conditioning problems with the evaluation of equations 3.35, 3.63 and 3.66.

The resulting algorithm converts an initial control u_{-1} to a new control u_i for a given number of iterations. The algorithm is controlled using the various tolerances, with lack of convergence of iterative procedures being detected when maximum iteration limits are exceeded.

3.3 Optimal control problem solution results

The algorithm described in the previous section is used to solve a variety of optimal control problems representing the clutch engagement problem. In this final section, the solution of typical problems is analysed, with comments made on the performance of the solution technique and on the performance of the powertrain for the resulting control strategies.

3.3.1 Optimal control problem construction

The particular clutch engagement problem considered is an engagement from rest problem, with initially, the engine flywheel speed, $x_1(0)$, at 225 rad/s, the clutch fully disengaged and all the other powertrain components in steady state, resulting

in zero initial conditions for the remaining powertrain model states. This problem is felt to be a typical, but difficult clutch engagement problem. The problem is addressed, first of all assuming that both the engine torque and the required clutch torque capacity are free to be controlled, and secondly assuming that the required clutch torque capacity is the only control, with the engine torque set to a predefined value. The normalisation values for these controls, used in the solution approach are detailed in table 3.3.

For the optimal control problem examples in this section, the performance of the optimal control problem is expressed using eight individual cost functionals. These costs are F_1, F_2, F_3, F_5, F_6, F_7, defined in equations (2.39), (2.40), (2.41), (2.44), (2.46), (2.47), and two additional costs. The first additional cost functional is

$$F_9 = r(t_f) = \begin{cases} (x_1(t_f) - x_{1min})^2 & \text{if } x_1(t_f) < x_{1min} \\ 0 & \text{otherwise} \end{cases} \tag{3.75}$$

where t_f is the time of clutch lock up, which measures the distance of the engine flywheel speed below x_{1min} at the point of clutch lock up. This is included, as cost F_6, equation (2.46), fails to adequately prevent the flywheel speed dropping too low when the engine torque was taken to be a control. This is due to lower performance measure values resulting from dramatic reductions in engine torque just prior to clutch lock up, which cause the engine flywheel speed to drop sharply. This results in the clutch engaging in less time, reducing many of the individual cost functionals, without significant increases in the cost F_6. The increase in F_6 remains small as the time at which the engine flywheel speed drops below x_{1min}, prior to clutch lock up, is very short. Increases in the weighting of cost F_6 just caused the drop in engine torque to increase, without causing any significant improvements in preventing the engine flywheel speed dropping too far. The other

additional cost is

$$F_{10} = \int_0^{t_f} \dot{x}_1^2 dt \tag{3.76}$$

which measures the variation of the engine flywheel speed prior to clutch lock up. This cost is included to demonstrate how the control strategies resulting from the solution of the optimal control problem can be modified by including additional cost functionals. The costs F_4 and F_8, equations (2.43) and (2.48), are not included. Cost F_4, which measures the excitation of the compliance at the point of clutch lock up, is not included, as the oscillation are adequately measured by costs F_1, F_2 and F_3, and when the engine torque is taken as a control the value of $\dot{x}_2 - \dot{x}_1$ can be instantaneously varied, just prior to engagement, so that cost F_4 has a small value at engagement, making its value meaningless. In other words, altering its value without affecting the dynamics of the powertrain, as this step in torque prior to engagement excites the oscillations. The normalisation values for these costs used in the solution procedure are detailed in table 3.4.

Using these eight cost functionals and any of the three multi-objective compromises, a number of different optimal control problem performance measures are designed, with the weights and goals obtained from a combination of physical intuition and trial and error. The set of different performance measure values presented in this section are detailed in table 3.5, with the weights and goals referring to the normalised cost functionals. For each multi-objective method, two performance measures are designed. The first measure depends on the costs measuring the oscillations after engagement, i.e. the lock up time, the clutch energy dissipation and the stalling costs. The second measure depends on these costs, as well as the cost measuring the oscillations prior to engagement and the cost measuring the transients of the engine flywheel speed prior to engagement. The inclusion of these last two costs tries to improve, perhaps, the behaviour of the powertrain during engagement. In both cases, the weights on the costs are unity,

apart from the costs measuring the likelihood of stalling, which are set to 10.0 so as to prevent stalling, and the costs measuring lock up time and clutch energy dissipation. These last two costs are felt to be complementary, and as a result, for the weighted sum method, their weights are halved, and for the goal attainment methods the weight corresponding to clutch energy dissipation is halved with a goal introduced for the lock up time. This is done so as to reduce the influence of these costs on the solution. An additional performance measure using the quadratic goal attainment method is also designed, using all of the costs in table 3.4, but making greater use of the goals, by introducing goals for costs F_1 and F_{10}. The weights of these two costs are also increased to 10.0, so as too give these costs great importance when their goals are not met, but little, if any, importance when they are met, or nearly met. This ensures that these costs will not be excessively high, and that their values will not be too small at the expense of other costs.

For all of the above optimal control problems, in addition to the mandatory constraint $x_1 = x_2$ at t_f, an additional constraint

$$(\dot{x}_2 - \dot{x}_1)|_{t_f} \text{sign}(x_1(0) - x_2(0)) \geq 100 \tag{3.77}$$

is included, where the state derivatives are evaluated using the state equations representing a slipping clutch, and $\underline{u}(t_f)$ represents the terminal time at which point the controls are the terminal controls $\underline{u}_f$. This constraint ensures that for the clutch engagement problem with $x_1(0) > x_2(0)$, just after lock up $\mu u_3 \geq \tau + \frac{100J_1(J_2+J_3u_2^2)}{J_1+J_2+J_3u_2^2}$, where τ is as defined in equation (2.11), which ensures that the clutch will remain locked up if $\tau > -\mu u_3$, an assumption already made in section 2.1. Furthermore, if it is assumed that $\tau \geq \frac{100J_1(J_2+J_3u_2^2)}{J_1+J_2+J_3u_2^2} - \mu u_3$, which is a likely situation of the clutch engagement problem, then the clutch will remain locked up with a certain degree of robustness, as $\mu u_3 \geq |\tau| + \frac{100J_1(J_2+J_3u_2^2)}{J_1+J_2+J_3u_2^2}$. The normalisation of this extra constraint, and the mandatory constraint used in the

solution of the optimal control problems, are detailed in table 3.6.

3.3.2 Solution performance of algorithm

The optimal control problems defined by the performance measures in table 3.5 are solved using the described algorithm. In the solution results presented, the tolerances, described in section 3.2, used in the algorithm are detailed in table 3.7. These values are chosen by trial and error and seem to produce satisfactory performance for a variety of different optimal control problems and solution techniques.

As examples of how the algorithm performs, the case when all the individual cost are included (optimal control problems 2,4 and 6 in table 3.5), applied to vehicle A, with both the engine torque and the required clutch energy dissipation active as controls, is taken. This case is seemingly one of the most difficult cases for the algorithm to solve. For this case, solutions have been obtained for all three multi-objective methods using the steepest descent gradient method, and for the weighted sum method and the quadratic goal attainment method using the Fletcher-Reeves conjugate gradient method. For these five solutions, the solution procedure is run for 50 iterations with the initial control u_{-1} taken as

$$(u_1, u_3, T) = \begin{cases} (100, 125, 1) & \text{for} \quad s \in [0, 1) \\ (100, 225, 1) & \text{for} \quad s = 1 \end{cases} \tag{3.78}$$

taking the engine torque and the final required clutch torque capacity to be at their upper limits, with the required clutch torque capacity, up until lock up, at a point approximately half way between its upper and lower bounds. Two additional solutions are obtained for optimal control problems designed using the weighted sum method and the quadratic goal attainment method, again using the steepest descent method, but with a different initial control of

$$(u_1, u_3, T) = \begin{cases} (50, 100, 1) & \text{for} \quad s \in [0, 1) \\ (100, 225, 1) & \text{for} \quad s = 1 \end{cases} \tag{3.79}$$

These seven solution results are detailed in figures 3.4 - 3.10. For each of these figures, graphs of the performance measure, the individual cost functionals, and the norm of the difference between the control obtained at a current iteration, and the control at the previous iteration, are plotted against the number of iterations. A graph of the resulting clutch plate speeds, obtained from the simulation of the powertrain model, using the control obtained after 50 iterations, or when the algorithm stops, is also included in these figures. These graphs are first of all used to characterise the general performance of the algorithm, and then to compare the various attributes of the different solution techniques for the different multi-objective compromises.

From these graphs, it can be seen that the performance measure values monotonically decrease against the number of iterations, with the rate of decrease decreasing until little improvement is made in its value. The individual cost components, also generally decrease, again with their rate of change diminishing as the iterations continue. However, it is possible for their values to change significantly without affecting the performance measure, as in figure 3.6 and figure 3.4 between iterations 39 and 40. The size of the movement in the controls, also, generally decreases. However, this control movement seems to be fairly erratic, with its value changing from iteration to iteration. Furthermore, large changes in the controls occur which fail to reduce the performance value by any notable value. This ensures that convergence of the algorithm can not be guaranteed when the rate of change of the performance value is negligible, but only when the rate of change of the controls is negligible. Even then, it is possible for the change in the controls to be very small for several iterations and then increase again, as in figure 3.5, making convergence extremely difficult to detect.

With these difficulties already noted, it would be preferable to identify the solution approaches which seem to perform the best. First of all, comparing the

performance of the algorithm applied to the different performance measures designed using the different multi-objective compromises, it can be seen that the weighted sum method (figure 3.4) takes longer to obtain a very small performance measure value taking about 20 iterations before little improvement is made, than the other two methods (figures 3.5 and 3.6), which take about 5 iterations. Similar results have also been obtained for other problems investigated. For the goal attainment method (ρ_∞), although it obtains a satisfactory performance measure, its behaviour is far from ideal, with large changes in individual cost components resulting in computational difficulties and the algorithm being terminated after 30 iterations as a smaller performance measure value was unable to be located. Furthermore, the performance of the resulting clutch plates speeds could be improved, by decreasing the oscillations after clutch lock up, represented by costs F_2 and F_3. These costs can be decreased by timing a dip in the clutch torque capacity and a small step in engine torque, without causing the other costs to be dramatically altered. However, for the goal attainment value, if the values of cost F_2 and F_3 are less than the maximum of the difference between the other costs and their goals, then these costs measuring the oscillations after lock up have no influence on the performance measure, and hence are not required to be small, resulting in the method failing to reduce these oscillations. A final disadvantage of the goal attainment method is that it is computationally more expensive, as noted in table 3.8. This table, details the computational requirements of the algorithm when run as a sole user on a SUN 4/330, over the network. The times have been obtained by taking an average of typical runs, rounded to the nearest five seconds, with the real time including the ACSL simulation times, the transfer of these calculations to and from disc and idle and system time. The Matlab CPU time was calculated using appropriate Matlab commands.

Another comparison that can be made from the figures, is the relative perfor-

mance of the steepest descent method to the Fletcher-Reeves conjugate gradient method, applied to the problem designed using the weighted sum method and the problem designed using the quadratic goal attainment method. In figure 3.11, plots of the performance measure against the number of iterations are shown for these two problems with the solid line showing the performance measure of the steepest decent method and the dashed line of the conjugate gradient method. From this figure, it can be seen, especially in the case of the weighted sum method, that the 2nd order conjugate gradient method achieves a low performance measure value in less iterations, with each iteration of this second order technique having similar computational requirements to the steepest descent method. Furthermore, the 2nd order method normally achieves a smaller performance measure value. However, convergence is still not achieved, with significant control changes still occurring right up to 50 iterations, as shown in figure 3.7 and 3.8, with the rate of decrease of the performance measure similar to the rate of decrease in the steepest descent method.

One final analysis of the performance is to investigate the dependence of the algorithm on the initial control u_{-1}. Figures 3.9 and 3.10 show solutions to the same optimal control problem, as in figures 3.4 and 3.5, but starting at a different control. From these figures it can be seen that the values of the performance measure located are very similar. However, the values of the individual cost components are very different in magnitude and the resulting controls are very different, as shown in figures 3.12 and 3.13, with different performance of the clutch plates resulting. The conclusion of this is that the solution after 50 iterations is dependent on the initial control, and thus after 50 iterations the algorithm has not converged. Unfortunately, if the number of iterations is increased, the algorithm still fails to converge. The suggested reason for these convergence problems is that, for the optimal control problems being solved, there exists a large set of controls

with performance measure values close to the optimal solution, but which lie over a large region of the domain $\mathcal{D}$. With such a problem, the gradients, and the changes in the performance measure along the search direction, are small, resulting in very slow convergence, with the behaviour of the solution dependent on higher order gradients. It could be argued that such an optimal control problem is ill-defined, motivating the inclusion of additional constraints and costs, but it is questionable whether this is prudent. For optimal control performance measures, which adequately measure the performance of the system, then any 'good' solution will suffice, with the only disadvantage being the lack of uniqueness of such a 'good' but non-optimal solution. This is especially true when the difference between the optimal performance measure value and the value obtained is negligible, as in the clutch engagement case. This lack of uniqueness may cause difficulties in comparing similar but different problems, as the solutions are dependent on the initial control, u_{-1}, and the solution procedure. However, it may still be possible to compare characteristics which are necessary to achieve 'good' solutions.

To conclude, analysis of the algorithm's performance when trying to solve the clutch engagement problem identifies that the algorithm locates a 'good' control strategy, which is dependent on the initial control taken, in about 5-10 iterations. The algorithm seems to perform, at its best, when solving an optimal control problem designed using the quadratic goal attainment method, with the search direction obtained using the conjugate gradient method.

3.3.3 Characteristics of 'good' clutch engagement

As discussed in the previous section, the location of the 'good' control strategies is dependent on the initial control, u_{-1}, the algorithm variables (i.e. the order and tolerances of the algorithm), as well as the optimal control performance measure design. For many solution results obtained, it is possible to identify key character-

istics of 'good' clutch engagement. These characteristics seem to be independent of the multi-objective method used in designing the performance measure. Furthermore, for each method, the characteristics do not seem to be that dependent on how each individual cost functional is included in the compromise, only whether they are included with a significant weight. This can be demonstrated by comparing the two solution results of optimal control problems 4 and 7 in table 3.5, with the second optimal control problem performance measure that make greater use of the additional adjustability of the goals in the quadratic goal attainment method. These two optimal control problem performance measures, include the same cost functionals, in two different ways, with the control strategies located after 50 iterations of the algorithm shown in figures 3.14 and 3.15. In both of these figures, graphs of the clutch plate speeds and vehicle velocity are presented, which are obtained from simulating the powertrain model, along with the controls that produce these states. From these figures, it can be seen that the resulting control strategies are very similar, with, in both cases, the clutch engaging in about 0.8 seconds, the vehicle reaching a speed of 10 km/h in 1.0 seconds, and the controls not differing significantly in shape or size. Further analysis of the results, for a variety of problems, seem to suggest that the characteristics of the solutions, obtained using the algorithm, are only dependent on the vehicle dataset, the number of controls active, and the cost functionals included in the performance measure.

Eight permutations, on which the solution seems to be dependent are now detailed. These permutations are of whether all or some of the costs are included in the performance measure, whether one or two controls are active, and whether vehicle A's or vehicle B's dataset is used. Again, for each case, graphs of the clutch plate speeds, vehicle speeds and controls, are detailed in figure 3.14 and figures 3.16 - 3.22. These solutions have been obtained using the quadratic goal attainment method (problems 3 and 4 in table 3.5) and calculated using the

algorithm for 50 iterations. From these graphs a number of characteristics can be identified which are present in all cases. In particular, just prior to clutch engagement the clutch torque capacity dips down. This reduction in clutch torque capacity reduces the torque jump at the point of clutch lock up, reducing the excitation of the oscillations at this point. Another common characteristic when the engine torque is active as a control, is the general shape of the engine torque transients, with the engine torque initially being small allowing the engine flywheel speed to decrease, speeding up engagement, then increasing in order to prevent the engine flywheel speed dropping significantly below 100 rad/s. A final common characteristic is the small dip in engine torque prior to clutch lock up. This dip, just prior to engagement, causes the clutch plate speeds to come together more speedily, with a increase at the point of lock up causing the clutch to engage smoothly with little excitation of the powertrain compliance.

Other characteristics for individual cases are also present. In particular:

1. when vehicle A's dataset is used in the optimal control problem, oscillations in the required torque capacity prior to engagement and oscillations in the powertrain dynamics a short time after starting to engage the clutch result,

2. when all the cost functionals are included in the performance measure, the clutch lock up time is longer, with the rate of increase of clutch torque capacity smaller than when the costs measuring the oscillations prior to lock up and flywheel transients are omitted,

3. for vehicle B, clutch engagement appears to be easier, engaging in less time.

The first effect is attributed to the significance of the oscillations in the powertrain compliance prior to engagement, with the oscillation in clutch torque capacity being of a similar frequency to the frequency of the oscillations prior to engagement.

With such oscillations being significant, the oscillations just after the start of engagement are unavoidable without significant increase in lock up time. The third observation is put down to the insignificance of the oscillations prior to engagement for vehicle B and to the greater range of the controls. An interesting case, is when the costs F_1 and F_{10} are omitted, with only the clutch torque capacity active as a control, applied to vehicle B (see figure 3.22). In this case, a converged solution has almost been obtained.

To conclude, the optimal control problem solution fails, in general, to obtain the theoretical optimal solution that exists. However, the formality of the approach has enabled 'good' clutch engagement solutions to be obtained. These 'good' solutions can be used to obtain characteristics inherent in 'good' clutch engagement, or directly as open loop control strategies applied to the problems used to generate them.

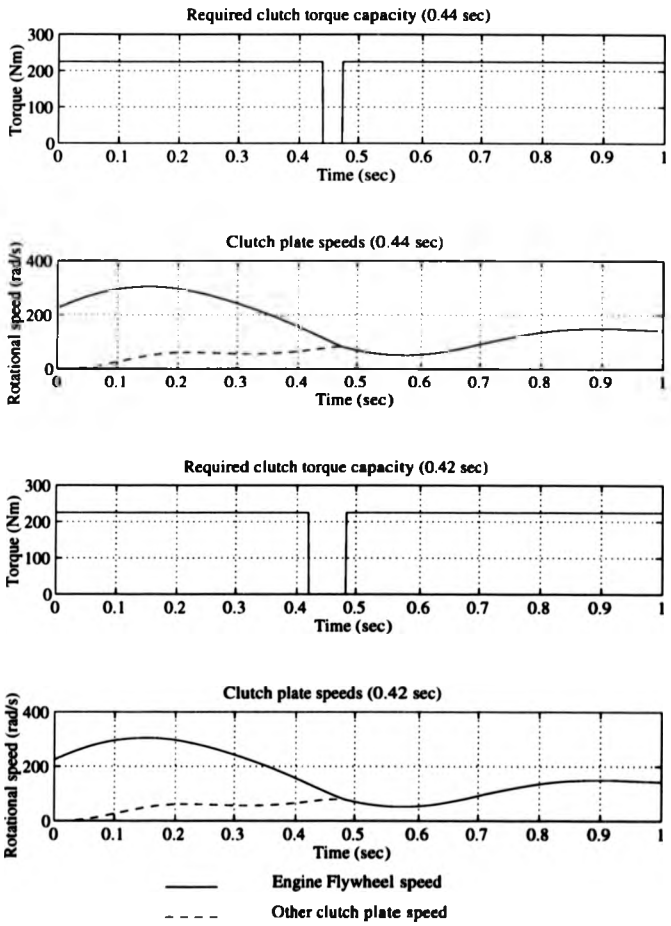


Figure 3.1: Bang-bang control simulations with banging times of 0.44 and 0.42 seconds

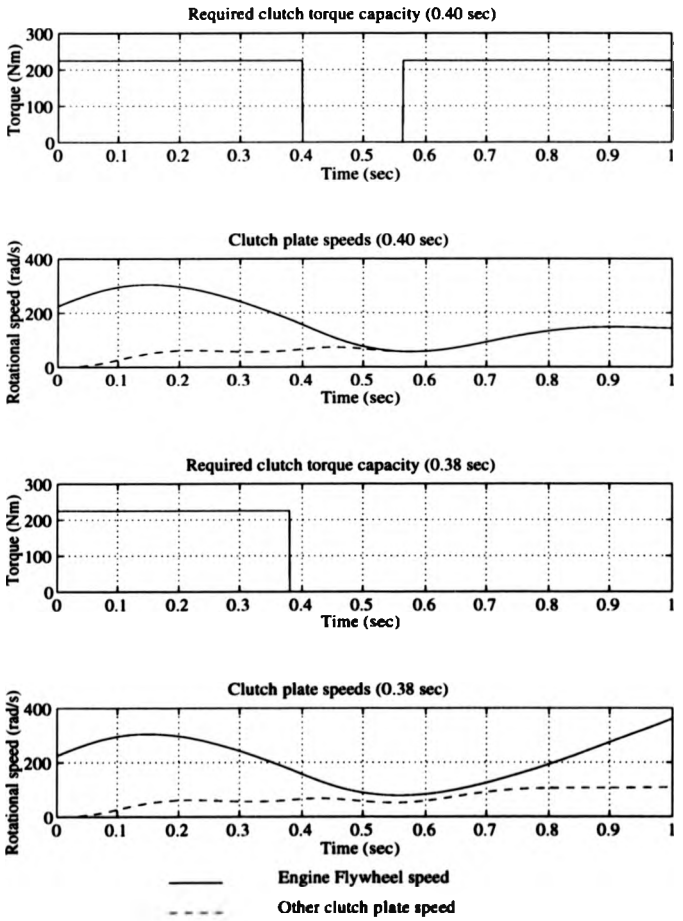


Figure 3.2: Bang-bang control simulations with banging times of 0.40 and 0.38 seconds

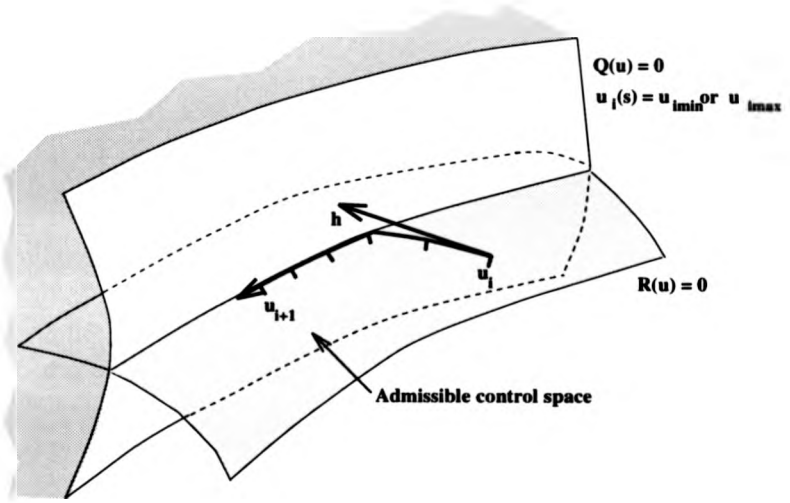


Figure 3.3: Pictorial representation of algorithm iteration

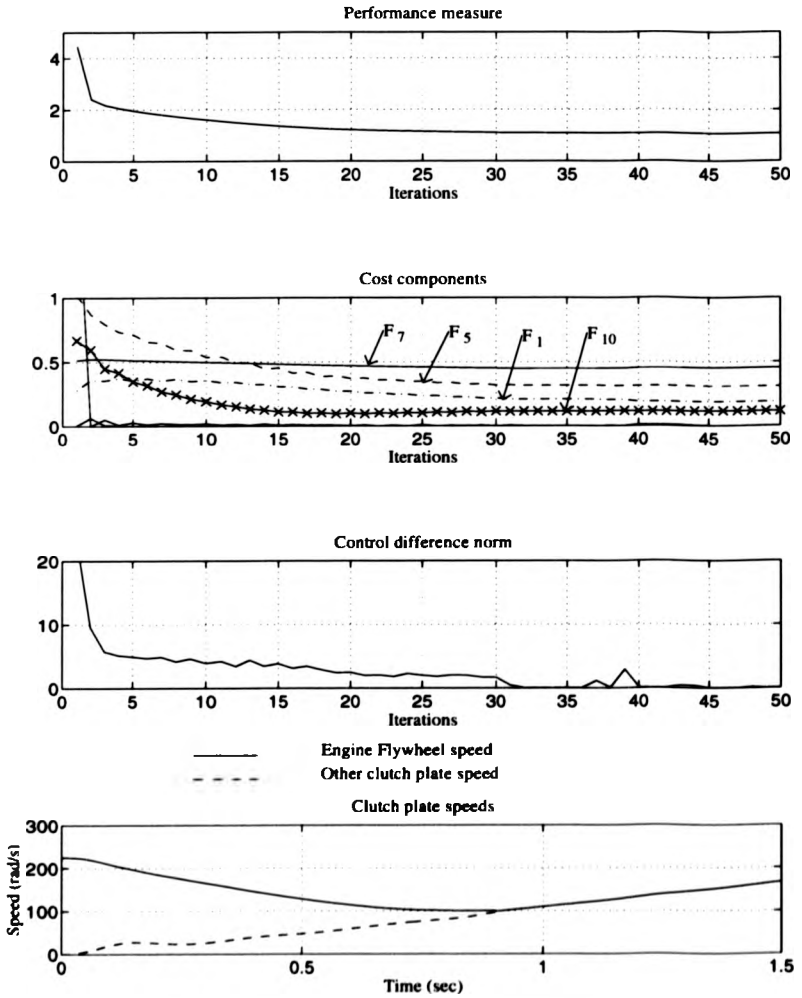


Figure 3.4: Optimal control problem solution: weighted sum; vehicle A; steepest descent method with initial control defined in equation (3.78)

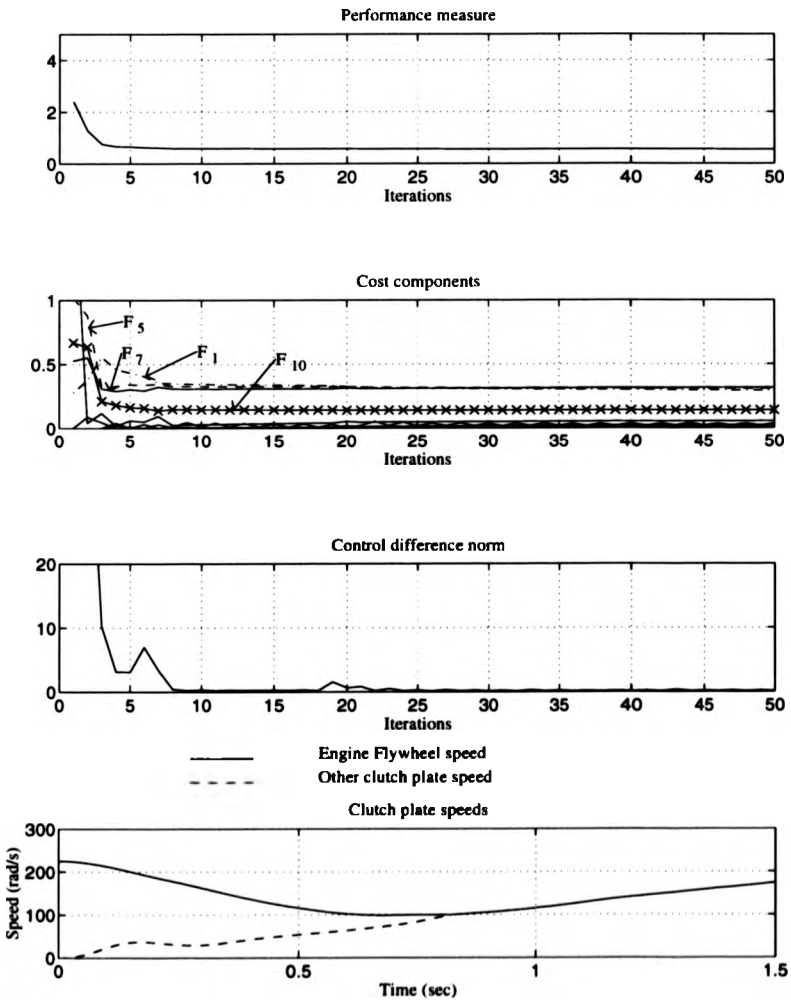


Figure 3.5: Optimal control problem solution: quadratic goal attainment; vehicle A; steepest descent method with initial control defined in equation (3.78)

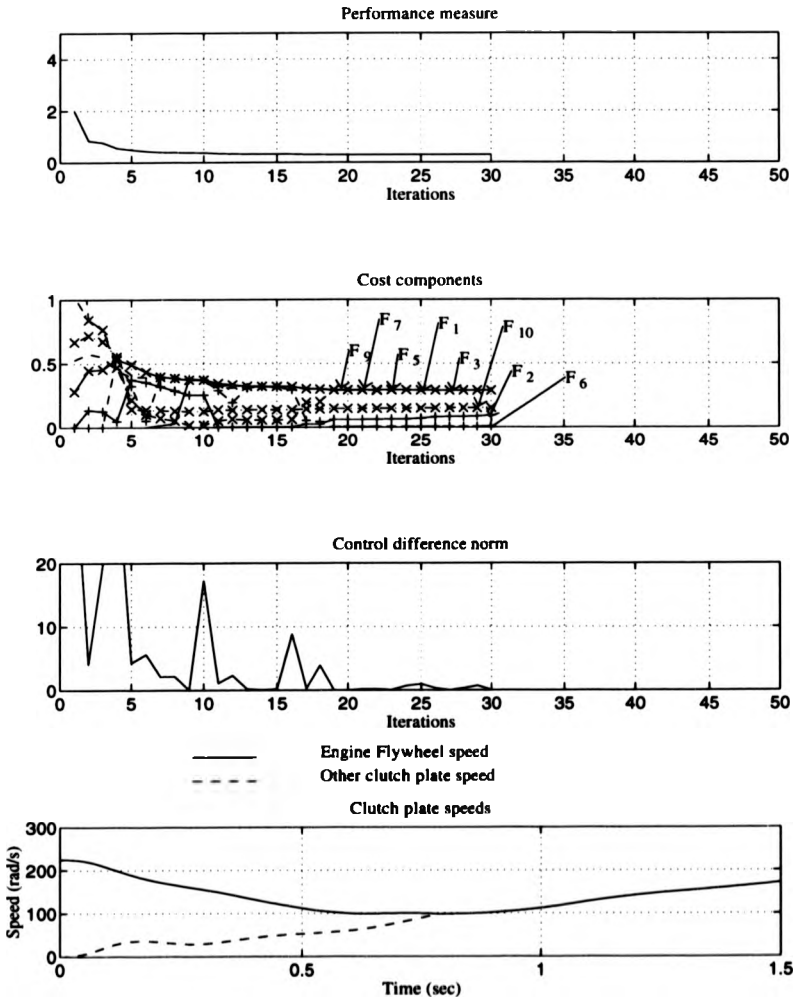


Figure 3.6: Optimal control problem solution: goal attainment; vehicle A; steepest descent method with initial control defined in equation (3.78)

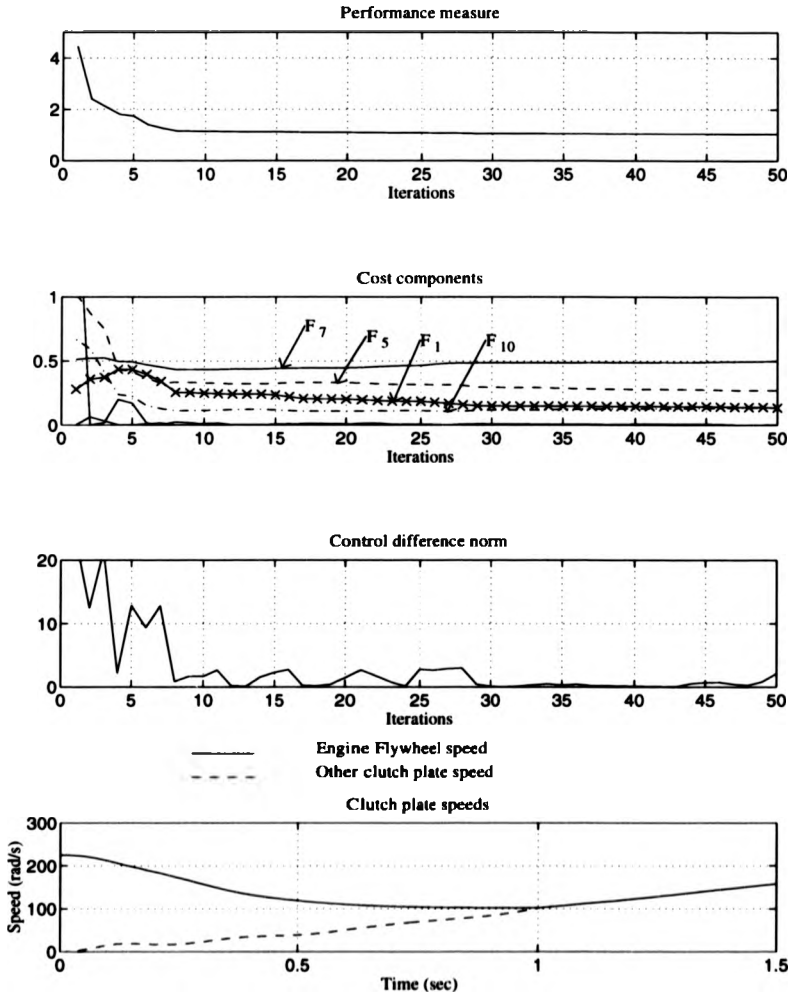


Figure 3.7: Optimal control problem solution: weighted sum; vehicle A; conjugate gradient method with initial control defined in equation (3.78)

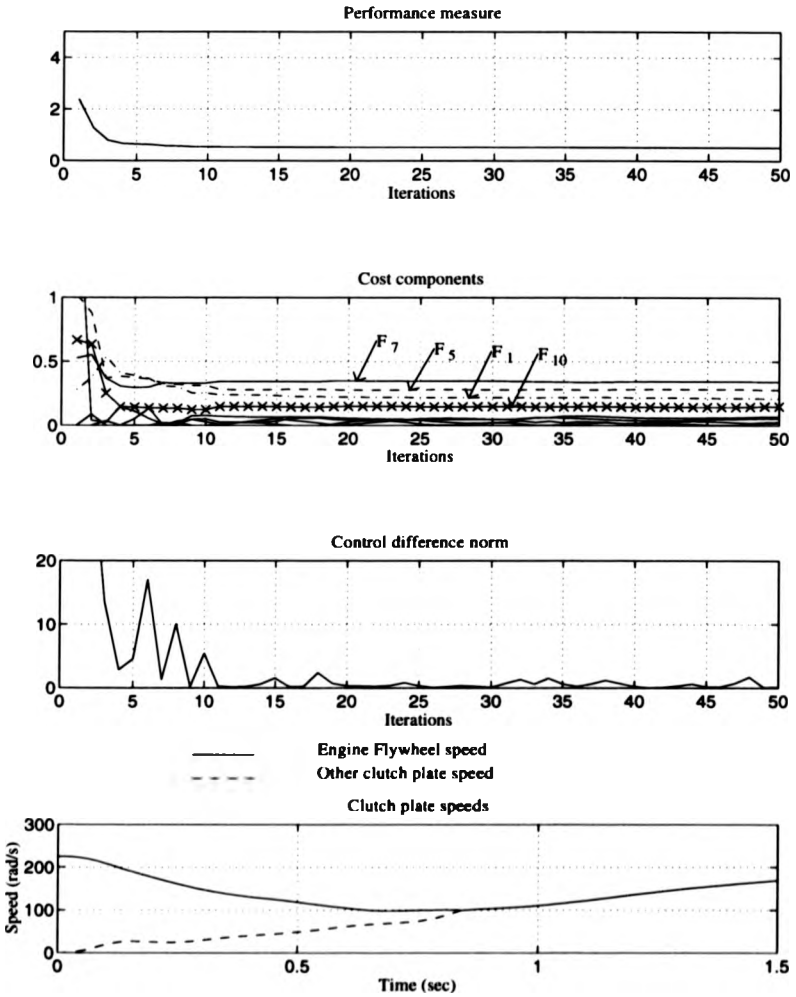


Figure 3.8: Optimal control problem solution: weighted sum; vehicle A; conjugate gradient method with initial control defined in equation (3.78)

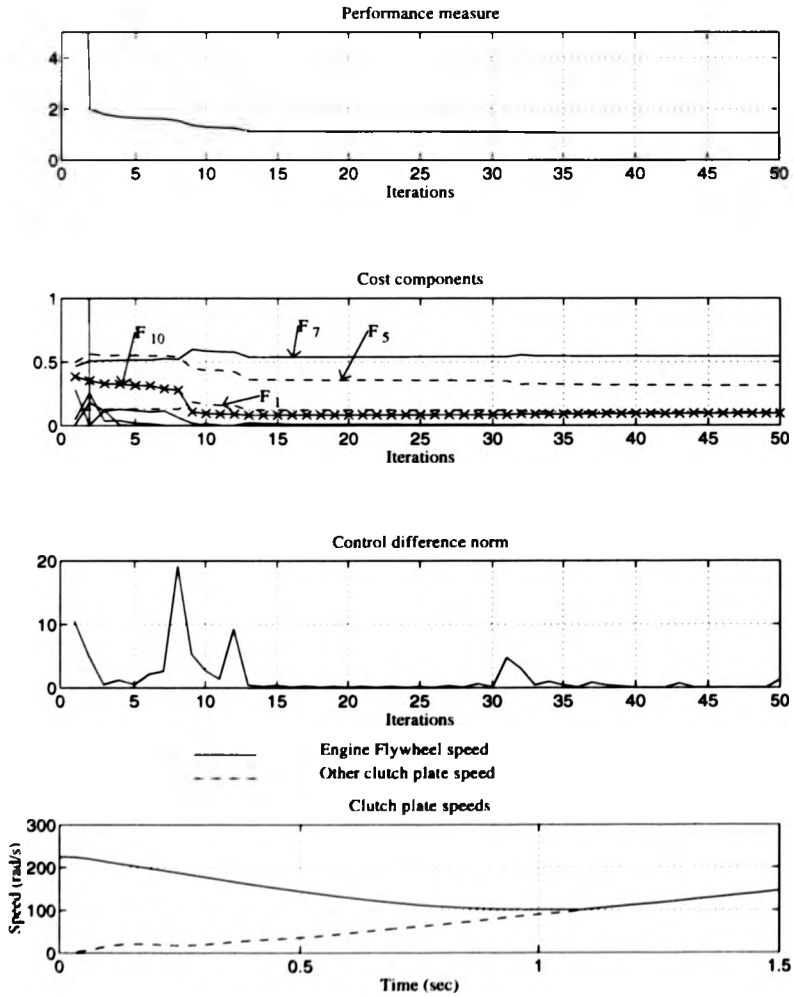


Figure 3.9: Optimal control problem solution: weighted sum; vehicle A; steepest descent method with different initial control defined in equation (3.79)

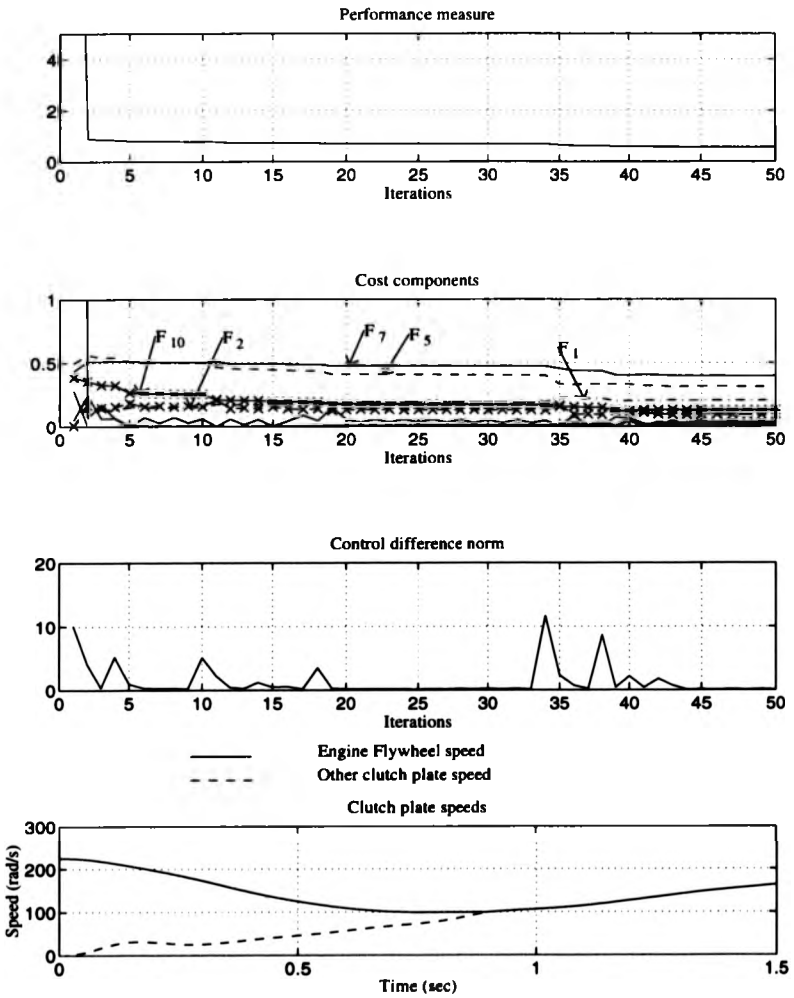


Figure 3.10: Optimal control problem solution: quadratic goal attainment; vehicle A; steepest descent method with different initial control defined in equation (3.79)

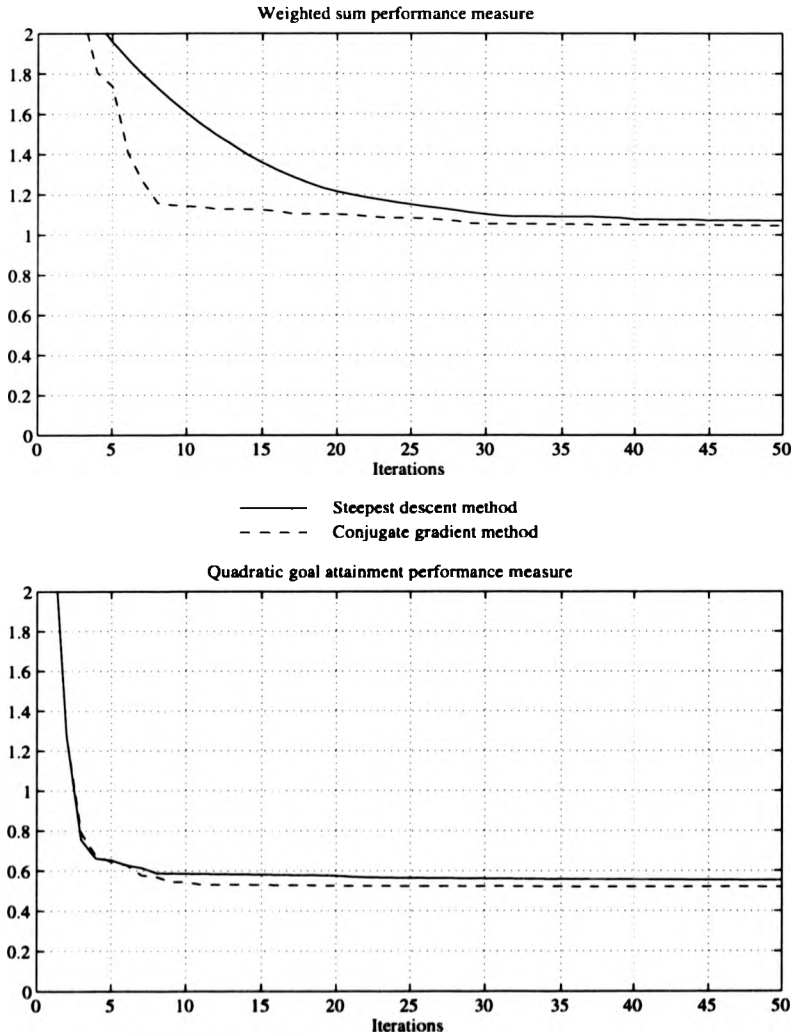


Figure 3.11: Comparison of performance measures for 1st and 2nd order gradient techniques

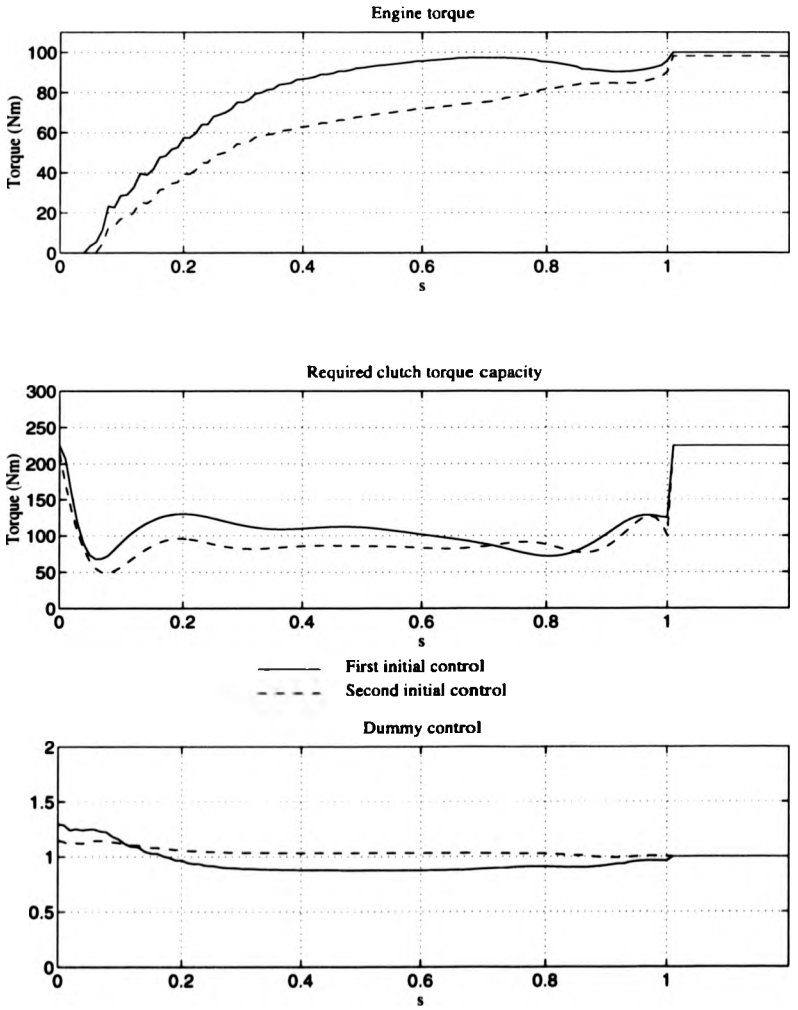


Figure 3.12: Different controls resulting from different initial controls: weighted sum method

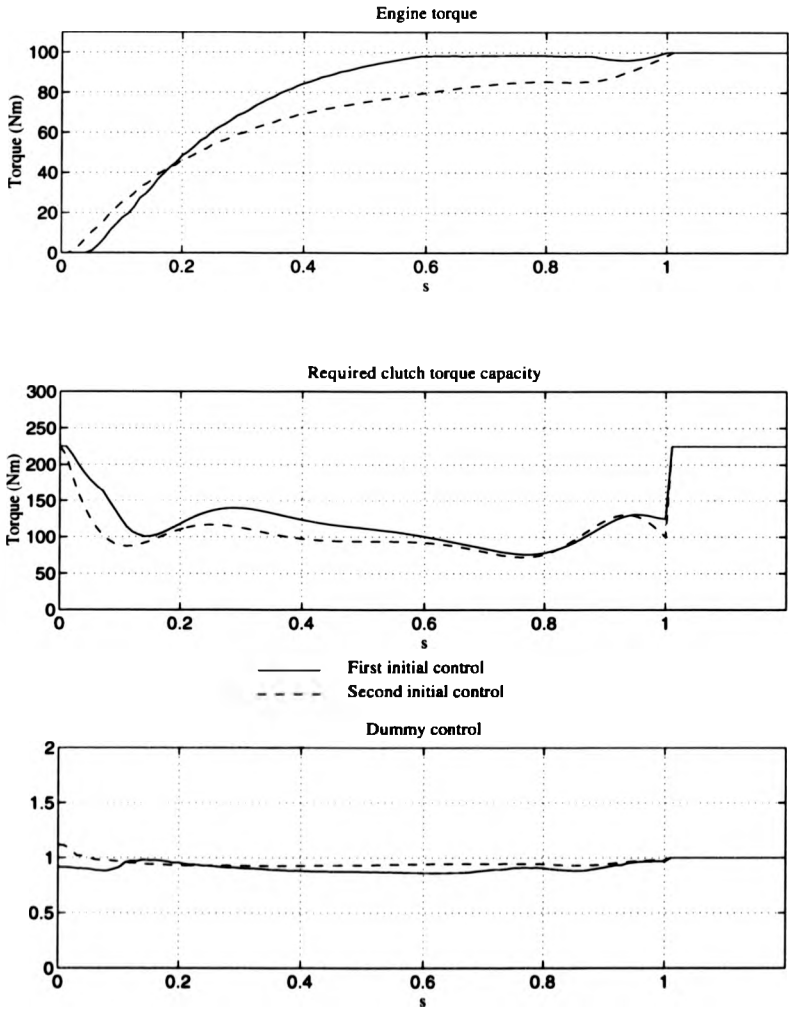


Figure 3.13: Different controls resulting from different initial controls: quadratic goal attainment method

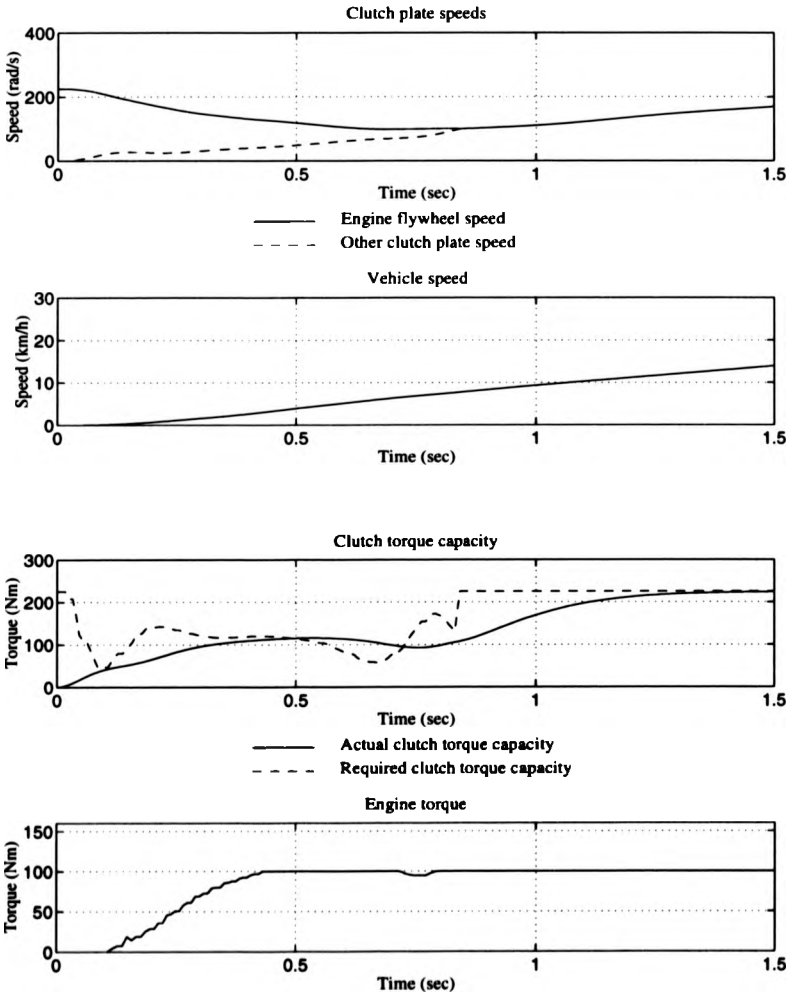


Figure 3.14: Powertrain performance of control solution: optimal control problem number 4 in table 3.5; vehicle A; two controls active.

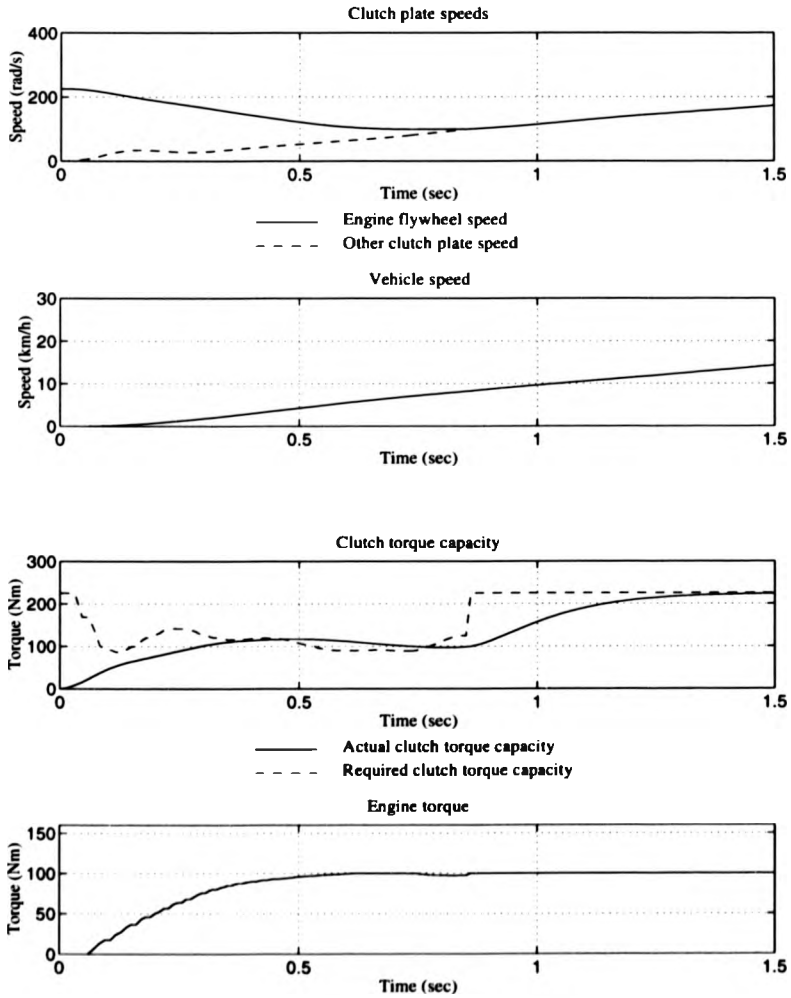


Figure 3.15: Powertrain performance of control solution: optimal control problem number 7 in table 3.5; vehicle A; two controls active.

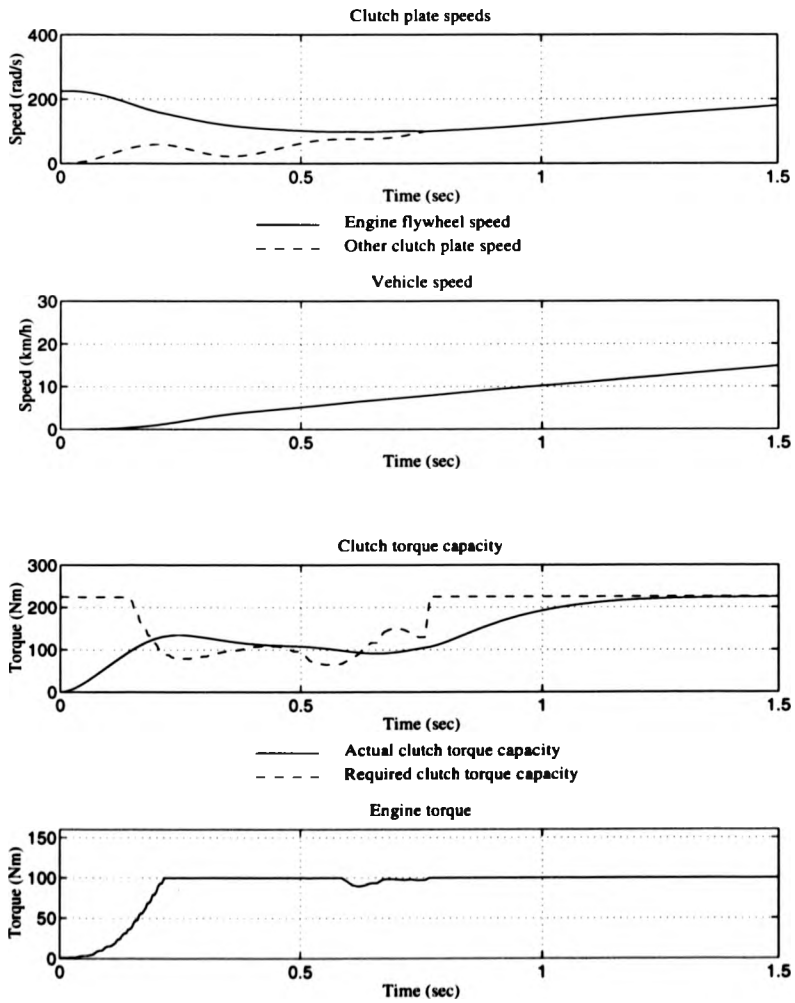


Figure 3.16: Powertrain performance of control solution: optimal control problem number 3 in table 3.5; vehicle A; two controls active.

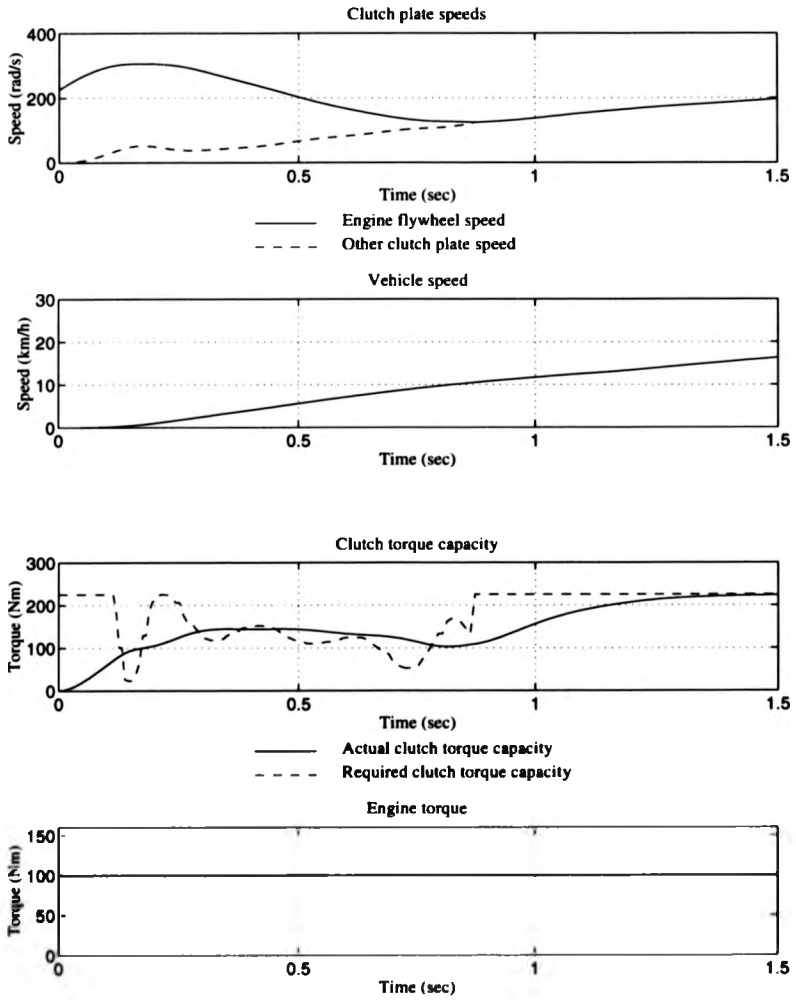


Figure 3.17: Powertrain performance of control solution: optimal control problem number 4 in table 3.5; vehicle A; one control active.

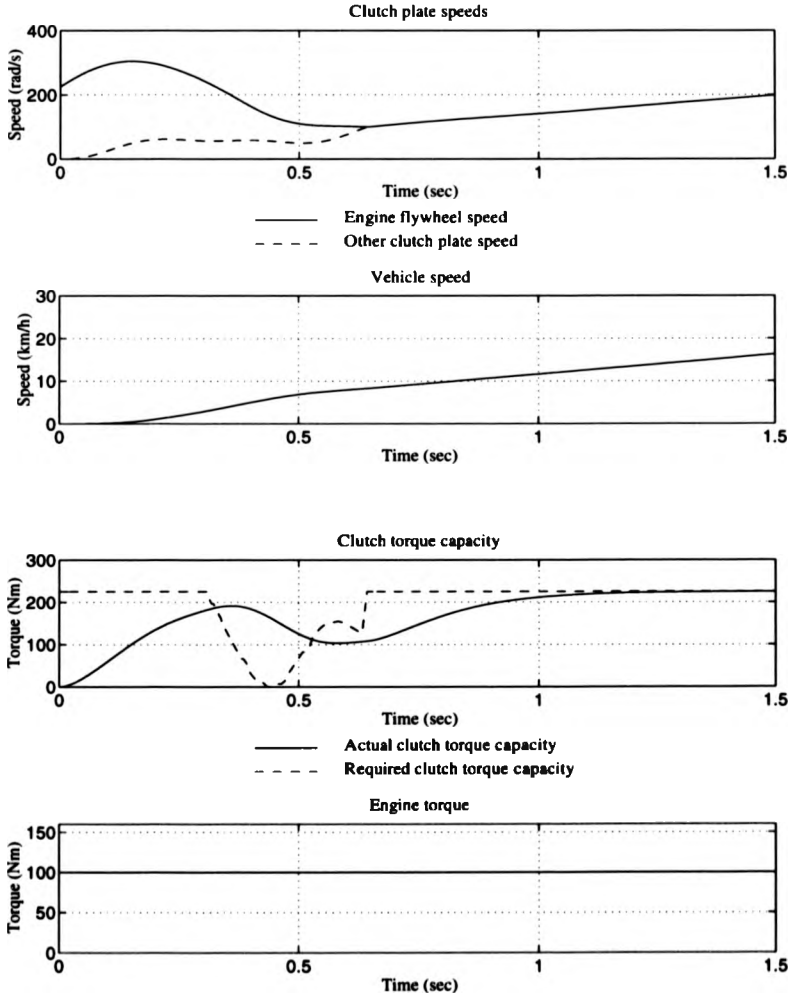


Figure 3.18: Powertrain performance of control solution: optimal control problem number 3 in table 3.5; vehicle A; one control active.

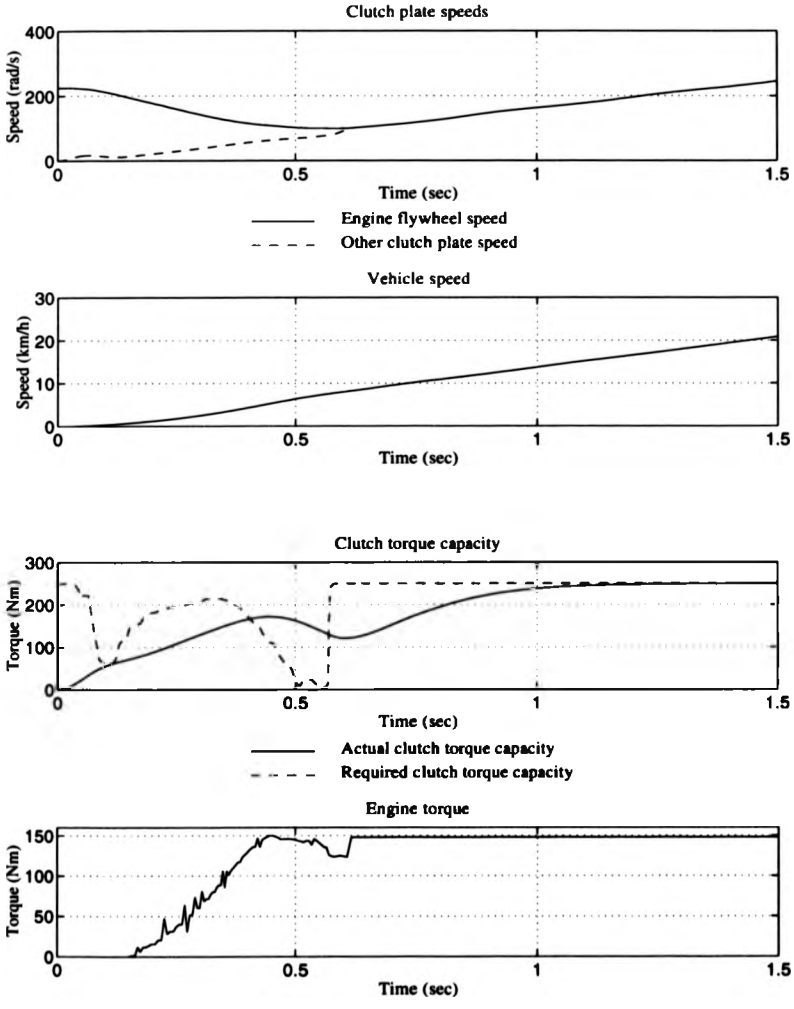


Figure 3.19: Powertrain performance of control solution: optimal control problem number 4 in table 3.5; vehicle B; two controls active.

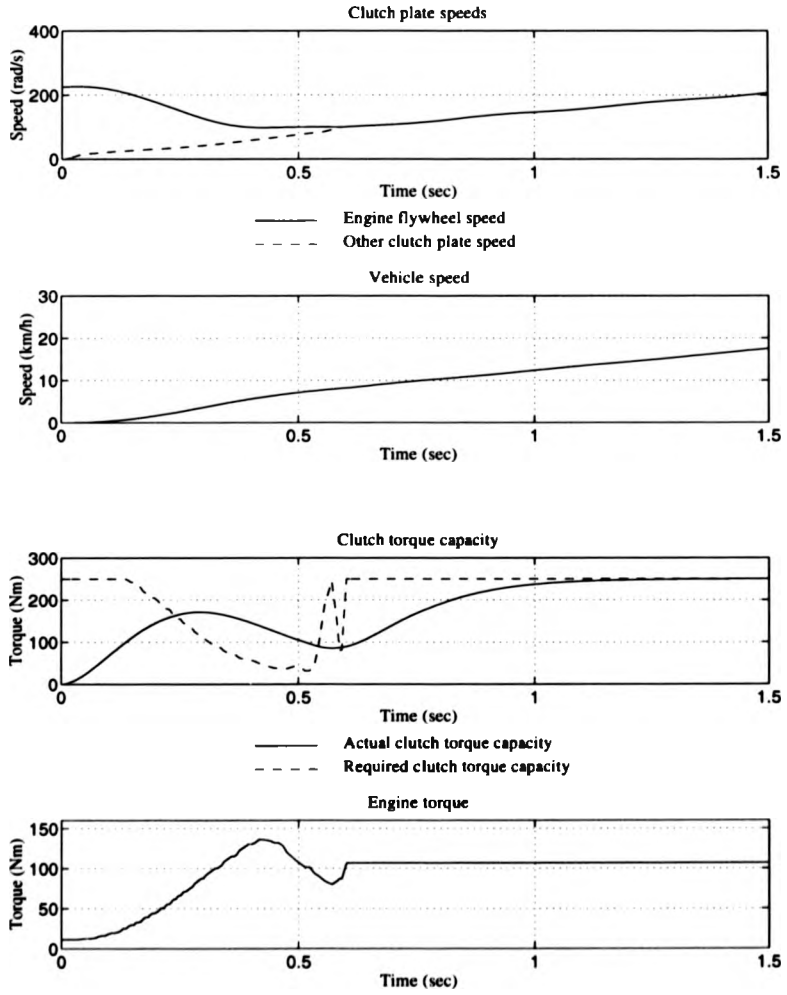


Figure 3.20: Powertrain performance of control solution: optimal control problem number 3 in table 3.5; vehicle B; two controls active.

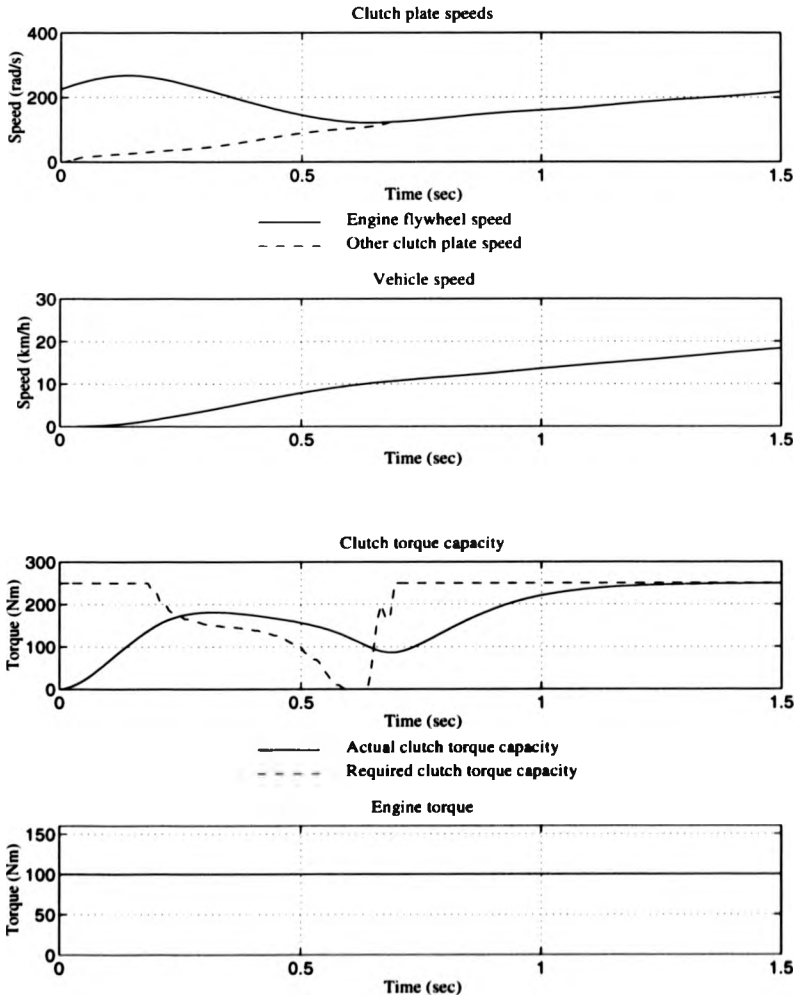


Figure 3.21: Powertrain performance of control solution: optimal control problem number 4 in table 3.5; vehicle B; one control active.

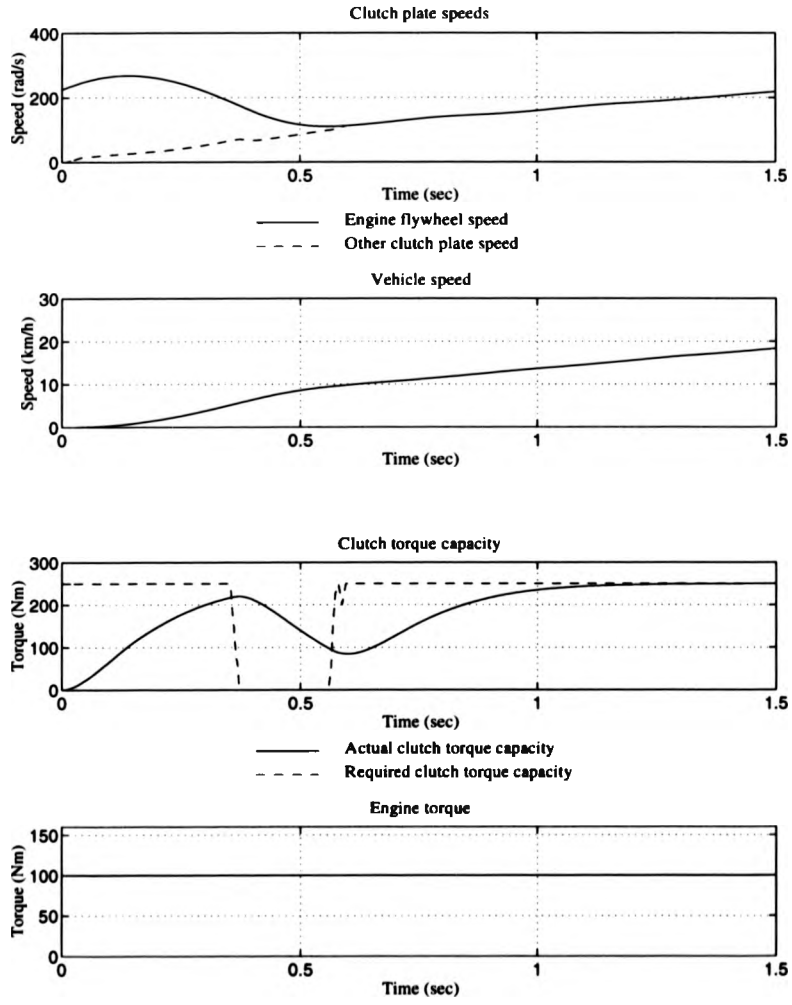


Figure 3.22: Powertrain performance of control solution: optimal control problem number 3 in table 3.5; vehicle B; one control active.

Costs	Weights
$(\dot{x}_2 - \dot{x}_1)\|_{t_f}$	0.03
$\int_0^{t_f} \mu x_5 \|x_1 - x_2\| dt$	0.00005
$\int_0^{t_f} 1 dt$	1.0

Table 3.1: Optimal control problem example: weights

Banging time	Performance measure	$0.03 \times (\dot{x}_2 - \dot{x}_1)\|_{t_f}$	$0.00005 \times \int_0^{t_f} \mu x_5 \|x_1 - x_2\| dt$	$1.0 \times \int_0^{t_f} 1 dt$
0.44	1.1629	0.1935	0.4958	0.4735
0.42	1.1047	0.1240	0.4971	0.4835
0.40	1.0774	0.0112	0.5035	0.5625
0.38	∞	?	?	∞

Table 3.2: Bang points example

Controls	Symbol	Normalisation values
Engine torque	u_1	100.0
Required clutch torque capacity	u_3	225.0
'Dummy' control	T	1.0

Table 3.3: Control normalisation values

Costs and constraints	Normalisation values
$F_1 = \int_0^{t_f} (x_3 - u_2 x_2)^2 dt$	40.0
$F_2 = \dot{x_4}^2(t_f)$	40.0
$F_3 = \ddot{x_4}^2(t_f)$	400.0
$F_5 = \int_0^{t_f} \mu u_3 \lvert x_1 - x_2 \rvert dt$	10000.0
$F_6 = \int_0^{t_f} r(t) dt$	100.0
$F_7 = \int_0^{t_f} 1 dt$	1.0
$F_9 = r(t_f)$	100.0
$F_{10} = \int_0^{t_f} \dot{x}_1^2 dt$	200000.0

$$\text{with } r(t) = \begin{cases} (x_1(t) - x_{1min})^2 & \text{if } x_1(t) < x_{1min} \\ 0 & \text{otherwise} \end{cases}$$

Table 3.4: Cost normalisation values

Performance measure	Multi-objective method	(weights, goals)							
		F_1	F_2	F_3	F_5	F_6	F_7	F_9	F_{10}
1	ρ_1	$(0,0)$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(0,0)$
2	ρ_1	$(1,0)$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(1,0)$
3	ρ_2	$(0,0)$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(1,\frac{1}{2})$	$(10,0)$	$(0,0)$
4	ρ_2	$(1,0)$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(1,\frac{1}{2})$	$(10,0)$	$(1,0)$
5	ρ_∞	$(0,0)$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(1,\frac{1}{2})$	$(10,0)$	$(0,0)$
6	ρ_∞	$(1,0)$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(1,\frac{1}{2})$	$(10,0)$	$(1,0)$
7	ρ_2	$(10,\frac{1}{4})$	$(1,0)$	$(1,0)$	$(\frac{1}{2},0)$	$(10,0)$	$(1,\frac{1}{2})$	$(10,0)$	$(10,\frac{1}{8})$

Table 3.5: Optimal control problem performance measures

Constraints	Normalisation values
$x_1 - x_2$	80.0
$(\dot{x}_2(t_f) - \dot{x}_1(t_f))\text{sign}(x_1(0) - x_2(0))$	600.0

Table 3.6: Constraint normalisation values

Tolerance description	Symbol	Value
Equality constraint relaxation parameter	ϵ_1	0.005
Inequality constraint relaxation parameter	δ_1	0.005
Maximum absolute error in control projection	d_{max1}	0.001
Maximum relative error in control projection	d_{max2}	0.001
Cost equality relaxation tolerance	ϵ_{cost}	0.01
Maximum error in search direction projection	d_{max3}	0.001
Rate of step increase	λ_{up}	1.25
Rate of step decrease	λ_{down}	5.0
Initial step size	ϵ_{step}	0.01
Maximum absolute error in interval halving procedure	ΔF	0.005
Maximum relative error in interval halving procedure	ρ	0.1

Table 3.7: Algorithm tolerances

Multiobjective method	Symbol	Approximate Matlab CPU time per iteration	Approximate real time per iteration
Weighted sum	ρ_1	20 sec	70 sec
Quadratic goal attainment	ρ_2	20 sec	70 sec
Goal attainment	ρ_∞	60 sec	120 sec

Table 3.8: Algorithm computational time

Chapter 4

Optimal feedback control of clutch engagement

In this chapter, drawbacks of the open loop optimal control problem for clutch engagement from rest are identified. Estimations of perturbations in the optimal control problem are then used to create a variational optimal control problem. The solution of this variational optimal control problem is used to modify the open loop solution, resulting in a feedback control strategy. Finally, the performance of this feedback strategy is analysed for a variety of different situations for which the open loop strategy is unsatisfactory. Some of this work at been presented at a recent international conference on control [67].

4.1 Open loop solution drawbacks

In the previous chapter, the solution of the optimal control problem results in a 'good' control strategy for a given problem. This control strategy is open loop, calculated off-line, being dependent on the initial states, the model parameters, the model equations as well as the performance measure. For this reason, if the initial conditions and the model do not adequately portray the clutch engagement problem and the powertrain dynamics, then the resulting clutch performance may deteriorate. In particular, if the initial flywheel speed is different, or if the car

is initially on a slight gradient, or even if the powertrain model is different, say because the vehicle is fully laden, or the clutch is worn, the calculated open loop control strategy may no longer maintain 'good' clutch engagement.

In order to assess the performance of clutch engagement under such perturbations an example is taken. This example is the case studied in the previous chapter, the case of engagement from rest, with the initial engine flywheel speed at 225 rad/s, both controls active and applied to the dataset representing vehicle A. The open loop solution to the optimal control problem, designed using the weighted sum method with all the costs included (problem 2 in table 3.5), results in good clutch performance as shown in figure 4.1. That is when simulated using the model equations used in its calculation. For this case, three model and problem perturbations are made, to the original problem, with simulations conducted to assess how the clutch performance is maintained. The three perturbations represent an engagement from rest with

1. the initial engine flywheel speed perturbed to only 175 rad/s
2. the car initially on a 1:10 gradient up hill
3. and the car fully laden (modelled by increasing inertia J_4 by 2 kgm^2) with a worn clutch (only 95% effective).

For these three perturbations, graphs of the clutch plate speeds resulting from simulations are detailed in figure 4.2, with the changes in the calculated costs detailed in table 4.1 along with the costs when no perturbations have been applied. From the figure, for the first perturbation, it can be seen that the clutch locks up very quickly, with only a small increase in the powertrain oscillations. However, in accomplishing this the engine flywheel speed, drops significantly below 100 rad/s, sufficiently low that the engine would probably stall. This stalling problem is also demonstrated by the cost functional values, with the costs measuring the

likelihood of stalling, costs F_6 and F_9, dramatically increasing when the initial flywheel speed perturbation is made. For the second perturbation, the engine flywheel speed remains above 100 rad/s, up until the time at which the clutch would lock up, if no perturbation had been made, at a time of approximately 1 second. At this time, as the speed of the clutch plate attached to the gearbox is lower than when the vehicle is on level ground, clutch engagement has not been completed. After this time, the clutch torque capacity is increased which would have ensured that the clutch remains locked up when no perturbation is made. This increase causes the clutch plates speeds to quickly converge, resulting in the engine flywheel speed dropping below 100 rad/s and the generation of oscillations when the clutch finally locks up a short time later. Again, the cost functional values indicate theses observations with increases in the costs measuring the likelihood of stalling and the costs measuring the oscillations after clutch lock up, costs F_2 and F_3. For the third perturbation, due to the increase in vehicle mass, the speed of the clutch plate attached to the gearbox is lower than expected, and due to the clutch being worn, the clutch plate speeds are further apart than expected. Hence, as before, clutch engagement is not completed at the time when clutch lock up is achieved under no perturbations. Again, due to increase in clutch torque capacity, the clutch plates speeds come together quickly, with large oscillations after engagement resulting. Once again, the individual costs demonstrate these findings, with increases in cost F_2 and F_3.

The above example demonstrates that the performance of the clutch engagement can deteriorate for realistic variations. Furthermore, it demonstrates that the performance measures adequately measure this deterioration. This result justifies the development of a refinement to the open loop control strategy which adjusts the controls in order to maintain the good clutch engagement performance by maintaining a small value of the performance measure.

4.2 Variational optimal control problem

As demonstrated, perturbations in the clutch engagement problem can cause the performance measure to change as the states and controls, on which the performance measure is dependent, change from their open loop calculated values. The obvious question is, can the changes in the performance measure be estimated? Fortunately such changes can be estimated using calculus of variations. This well established theory allows variations of functionals of the form

$$F(u) = \int_0^1 T^2 f_0(\underline{x},\underline{u})ds + f_t(\underline{x},\underline{u})|_{s=1} \tag{4.1}$$

$$\text{subject to} \quad \frac{d\underline{x}}{ds} = T^2(s)f(\underline{x},\underline{u}),\ \underline{x}(0) = \underline{x}_0 \tag{4.2}$$

where f, f_0 and f_t are all continuously differentiable, to be calculated. The functional is of the fixed terminal time form as variations of the higher order fixed terminal time problem will be used to construct a variational optimal control problem, although theory exists allowing variations of a free terminal time form to be used. The use of this higher order problem form is for similar reasons to the justifications for its development in chapter 3. The first and second variations of such functionals with respect to variations of the states and controls at a given control u^* are

$$\delta F_* = p_0^T \delta\underline{x}_0 + \begin{pmatrix} f_{tx} & f_{tu} \end{pmatrix}_* \begin{pmatrix} \delta\underline{x}_f \\ \Delta\underline{u}_f \end{pmatrix} + \int_0^1 H_u \Delta\underline{u}\, ds \tag{4.3}$$

$$\begin{aligned} \delta^2 F_* = p_0^T \delta^2\underline{x}_0 + \begin{pmatrix} \delta\underline{x}_f^T & \Delta\underline{u}_f^T \end{pmatrix} \begin{pmatrix} f_{txx} & f_{txu} \\ f_{tux} & f_{tuu} \end{pmatrix}_* \begin{pmatrix} \delta\underline{x}_f \\ \Delta\underline{u}_f \end{pmatrix} \\ + \int_0^1 \begin{pmatrix} \delta\underline{x}^T & \Delta\underline{u}^T \end{pmatrix} \begin{pmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{pmatrix}_* \begin{pmatrix} \delta\underline{x} \\ \Delta\underline{u} \end{pmatrix} ds \end{aligned} \tag{4.4}$$

$$\text{subject to } \dot{\delta\underline{x}} = f_x \delta\underline{x} + f_u \Delta\underline{u},\ \delta\underline{x}(0) = \delta\underline{x}_0. \tag{4.5}$$

In these expressions: $H = T^2(\underline{p}^T f + f_0)$ is the Hamiltonian; $\underline{p}$ are the costates evaluated as usual using $\dot{\underline{p}} = H_x^T$ and $p(1) = f_{tx}$; $\delta\underline{x}$ and $\delta^2\underline{x}$ are first and second order variations of the states; $\Delta\underline{u}$ represents the variations of the controls; subscript 0 denotes the evaluation at $s = 0$; subscript f denotes the evaluation at $s = 1$ and subscript $*$ denotes the evaluation of the Hamiltonian and functional derivatives at u^*, calculated by evaluating corresponding state x^* and costates p^*. Note that the constraint, equation (4.5), is just the linearisation of the state equations. The calculation of such variations is well described in a tutorial paper [68]. This allows the functional to be approximated by

$$\hat{F} = F_* + \delta F_* + \frac{1}{2}\delta^2 F_* \tag{4.6}$$

at u^* with respect to variations of the controls and initial states, neglecting all variations of third order or higher. This in turn allows approximations of the individual cost functionals and the individual constraint functionals to be calculated.

In order, to obtain a variational optimal control problem, variational approximations of the performance measure are required. In the case of optimal control problems designed using the weighted sum method, first and second order variations of $F = \rho_1(\underline{w}, \underline{F})$, as defined in equation (2.58), are just linear combinations of the first and second order variations of the individual cost functionals, being

$$\delta F = \rho_1(\underline{w}, \delta\underline{F}) \tag{4.7}$$

$$\delta^2 F = \rho_1(\underline{w}, \delta^2\underline{F}) \tag{4.8}$$

where $\delta\underline{F}$ and $\delta^2\underline{F}$ are vectors of the first and second order variations of the individual cost functionals. For optimal control problems designed using the quadratic goal attainment method, in the region where $F = \rho_2(\underline{w}, \underline{F} - \underline{c}) > 0$, as in equation (2.62), the first and second order variations are

$$\delta F = \frac{\sum_i w_i \delta F_i \max\{0, w_i(F_i - c_i)\}}{F} \tag{4.9}$$

$$\delta^2 F = \frac{\sum_i \delta^2 G_i - \delta F^2}{F} \tag{4.10}$$

$$\text{where} \quad \delta^2 G_i = \begin{cases} w_i^2 \delta F_i^2 + w_i^2 \delta^2 F_i (F_i - c_i) & \text{if } w_i F_i > w_i c_i \\ 0 & \text{if } w_i F_i < w_i c_i \end{cases} \tag{4.11}$$

where δF_i and $\delta^2 F_i$ are variations of the individual cost functionals. When $F_i = c_i$ the second variation is dependent on the variation δF_i or $\delta^2 F_i$ if $\delta F_i = 0$, with $\delta^2 G_i$ obtaining the first value if $\delta F_i > 0$ and the second value if $\delta F_i < 0$. Apart from the discontinuity of the second variation, the form of the second variation is particularly messy, containing a product of two integrals in the term δF^2. Finally, for the optimal control problem designed using the goal attainment method with $F = \rho_\infty(\underline{w}, \underline{F} - \underline{c})$, equation (2.60), both the first and second order variations of performance measures are dependent on the variations δF_i and $\delta^2 F_i$ of the individual cost functionals, with the values of the variations being

$$\delta F = \max_i \{F_i + \delta F_i\} - F \tag{4.12}$$

$$\delta^2 F = 2(\max_i \{F_i + \delta F_i + \frac{1}{2}\delta^2 F_i\} - \max_i \{F_i + \delta F_i\}). \tag{4.13}$$

With these variations now calculated, a variational optimal control problem can be constructed by taking variational approximations of the performance measure and the constraints in the higher order optimal control problem. For the constraints, if it is assumed that the variations about the calculated values are small, then the first order variation will dominate, allowing the second order term to be neglected. For the performance value, the first variation, calculated at the control strategy resulting from the solution of the optimal control problem in the previous chapter, will be very small. This is due to the small gradients that result when a 'good' solution has been obtained. However for generality, this term is still accounted for allowing variations about non-optimal solutions to be taken, so long as the higher order variations are not significant. This results in an approximation

of the optimal control problem

$$\min_{u^*+\Delta u \in \mathcal{U}} \left\{ F_* + \delta F_* + \frac{1}{2}\delta^2 F_* \right\} \tag{4.14}$$

$$\text{subject to } \underline{R}_* + \delta \underline{R}_* = \underline{0} \tag{4.15}$$

$$\underline{Q}_* + \delta \underline{Q}_* \leq \underline{0}. \tag{4.16}$$

with u^*, with no underscore, representing a column vector of $\underline{u}$, T and $\underline{u}_f$, as in the previous chapter. In this approximation the values of F_*, $\underline{R}_*$ and $\underline{Q}_*$ are unaffected by the control variations. This allows the values of F_* and $\underline{R}_*$ to be immediately removed as they have no other affect on the solution, with the value of $\underline{Q}_*$, being eliminated by assuming that the variations δQ_{k_2} are small and equating $Q_{k_2*} + \delta Q_{k_2} \leq 0$ with $\delta Q_{k_2} \leq 0$ if $Q_{k_2*} + \delta_{k_2} \geq 0$ where δ_{k_2} is the tolerance used in the open loop solution algorithm to define the boundaries of the inequality constraints. Similarly, the controls u_* can be eliminated by equating the bounds $u_{i*} + \Delta u_i \in [a_i, b_i]$ with $\Delta u_i \geq 0$ if $u_{i*} - \epsilon_{icon} \leq a_i$ and $\Delta u_i \leq 0$ if $u_{i*} + \epsilon_{icon} \geq b_i$, where ϵ_{icon} is a parameter relaxing the control bounds. This results in a variational optimal control problem

$$\min_{\Delta u} \left\{ \delta F_* + \frac{1}{2}\delta^2 F_* \right\} \tag{4.17}$$

$$\text{subject to } \delta \underline{R}_* = \underline{0} \tag{4.18}$$

$$\delta Q_k \leq 0 \text{ if } Q_{k_2*} + \delta_{k_2} \geq 0 \tag{4.19}$$

$$\Delta u_i \geq 0 \text{ if } u_{i*} - \epsilon_{icon} \leq a_i \tag{4.20}$$

$$\Delta u_i \leq 0 \text{ if } u_{i*} + \epsilon_{icon} \geq b_i. \tag{4.21}$$

4.3 Variational optimal control problem solution

The variational optimal control problem in its general form is very difficult to solve. In particular, for the cases when the optimal control problem has been designed using either the goal attainment method or the quadratic goal attain-

ment method, no easy solution has been found. However, for problems where the weighted sum method has been used, the resulting variational optimal control problem is of a form that allows a solution to be obtained, with a little work. The use of such variational optimal control problem solutions for developing optimal feedback control is well known, as discused in [65], with details of such solutions for constrained optimal control problems discused in [69], along with an application to space shuttle guidance.

For variational optimal control problems resulting from an optimal control problem designed using the weighted sum method, substituting the variations in equations (4.3), (4.4), (4.7) and (4.8) into equations (4.17) - (4.21) yields a variational optimal control problem which is a Linear Quadratic Regulator (LQR) type problem. In this optimal control problem the variational performance measure is quadratic in terms of first order variations of states and controls subject to linearised state space equations and restrictions in the form of linear variational constraints and control variational bounds. In the substitution of equations (4.3) and (4.4), the first terms can be neglected as the initial states are taken as given, they are inherent in the problem and not adjustable. Lagrange multiplier variations can be introduced in order to manage the constraints in equations (4.18) and (4.19). The variational Lagrange multipliers add varying magnitudes of the variational constraint functionals to the performance measure, so that the solution of the new unbounded problem is equal to the solution of the original constrained optimal control problem. Furthermore, from Lagrange multiplier theory, in particular the Kuhn-Tucker theorem for inequality constraint Lagrange multipliers, each Lagrange multiplier must satisfy

$$\delta R_j \delta\lambda_j = 0 \tag{4.22}$$

$$\text{or } \delta Q_k \delta\lambda_k = 0 \text{ and } \delta\lambda_k \geq 0 \tag{4.23}$$

depending on whether the Lagrange multiplier equates to an equality or inequality

constraint, where $\delta\lambda_j$ and $\delta\lambda_k$ are the variational Lagrange multipliers. This just leaves the control bounds which prevent the variational optimal control problem from being a standard LQR problem. For the control bounds, referring back to the open loop solutions results, the calculated controls prior to engagement often fail to achieve their bounds, as in figure 4.1, furthermore, if required, a certain degree of control bound margin can be introduced by artificially tightening the control bounds in the open loop solution procedure, without significantly affecting the performance measure value of the control located. For this reason it is assumed that the controls fail to achieve their bounds prior to engagement. For the controls after engagement, only the engine torque affects the performance measure value, with this control often achieving it's maximum bound after engagement, as in figure 4.1. Fortunately due to the finite dimension of the constraint this bound can be easily managed by introducing the variational constraint that the variation of the engine torque must be negative, $\Delta u_1|_{s=1} \leq 0$, when the engine torque achieves its maximum bound (i.e. if $u_1 + \epsilon_{1+} \geq b_1$), which like the other inequality constraints can be managed with a variational Lagrange multiplier. Typically, ϵ_{1+} is taken to be one percent of the range of the engine torque.

The resulting problem is now a time variant LQR problem, with the terms in the LQR problem dependent on the independent variable s. If a column vector $\underline{L}$ of cost and constraint integrands with a corresponding column vector $\underline{\psi}$ of terminal terms is introduced this variational optimal control problem can be expressed by

$$\min\left[\int_0^1 \begin{pmatrix}\delta\underline{x}\\ \Delta\underline{u}\\ \Delta T\end{pmatrix}^T \left\{ P_a \begin{pmatrix}\delta\underline{x}\\ \Delta\underline{u}\\ \Delta T\end{pmatrix} + \begin{pmatrix} T^2\underline{L}_x^T\\ T^2\underline{L}_u^T\\ 2T\underline{L}^T\end{pmatrix}_{\bullet} \delta\underline{\beta}\right\} ds + \begin{pmatrix}\delta\underline{x}_f\\ \Delta\underline{u}_f\end{pmatrix}^T \left\{ P_b \begin{pmatrix}\delta\underline{x}_f\\ \Delta\underline{u}_f\end{pmatrix} + \begin{pmatrix}\underline{\psi}_x^T\\ \underline{\psi}_u^T\end{pmatrix}_{\bullet} \delta\underline{\beta}\right\}\right] \tag{4.24}$$

$$\text{subject to } \delta\dot{\underline{x}} = T^2 f_x \delta\underline{x} + T^2 f_u \Delta\underline{u} + 2Tf\Delta T, \delta\underline{x}(0) = \delta\underline{x}_0 \tag{4.25}$$

$$\text{where } \delta\underline{\beta} = W\delta\underline{\lambda} + \underline{w} \tag{4.26}$$

$$H = \underline{w}^T \underline{L} + \underline{p}^T f \tag{4.27}$$

$$P_a = \frac{1}{2}\begin{pmatrix} T^2 H_{xx} & T^2 H_{xu} & 2T H_x^T \\ T^2 H_{ux} & T^2 H_{uu} & 2T H_u^T \\ 2T H_x & 2T H_u & 2H \end{pmatrix}, \tag{4.28}$$

$$P_b = \frac{1}{2}\begin{pmatrix} f_{txx} & f_{txu} \\ f_{tux} & f_{tuu} \end{pmatrix}. \tag{4.29}$$

In this variational optimal control problem the controls notation refers back to the notation where $\underline{u}$ refers to a vector of the controls in the powertrain model, with δ referring to a first variation of a variable, subscript $*$ denoting the evaluation at u^* with the corresponding state x^* and costates p^* calculated as usual, $\underline{w}$ and W being the performance measure weights and constraint determination matrix relating the weighted sum performance measure and the constraints to the vectors $\underline{L}$ and $\underline{\psi}$, and $\delta\underline{\lambda}$ being a vector of the Lagrange multipliers. The question posed in this optimal control problem, is that given an initial state perturbation $\delta\underline{x}_0$, what is the best control variation that minimise the performance measure. For a solution to be obtained for such a problem, matrices P_a and P_b must be positive semi-definite with matrices $\begin{pmatrix} T^2 H_{uu} & 2T H_u^T \\ 2T H_u & 2H \end{pmatrix}$ and f_{tuu} positive definite. Unfortunately, due to the linearity of some of the controls in the optimal control problem, these last two matrices, often fail to be positive definite, with the first matrix often being close to the zero matrix. This problem is overcome by introducing a tracking condition. Physically, this says that the modifications of the open loop solution are more desirable if they are close to the open loop solution. This tracking condition is formulated by adding the term

$$\Delta\underline{u}_f^T V_t \Delta\underline{u}_f + \int_0^1 \left(\Delta\underline{u}\Delta T\right) V_0 \begin{pmatrix} \Delta\underline{u} \\ \Delta T \end{pmatrix} ds \tag{4.30}$$

to the variational performance measure, equation (4.24), with the normally diagonal positive definite matrices V_0 and V_t, define the meaning of closeness as well

as defining a compromise between tracking along the open loop controls and the minimisation of the optimal control problem performance variation. This ensures that a solution is theoretically solvable, as by including this tracking condition, matrices $\begin{pmatrix} T^2 H_{uu} & 2T H_u^T \\ 2T H_u & 2H \end{pmatrix} + V_0$ and $f_{tuu} + V_t$ can be made positive definite. In practice, only a very small tracking condition is required as the matrices are at least positive semi-definite, although larger tracking conditions aid the computation of the solution by reducing ill-conditioning problems. This tracking condition, ensuring that the control variations are never too large, also helps to ensure that the variations remain in a region where the variational approximations used in creating the variational optimal control problem are reasonably accurate, with the neglected higher order terms negligible.

The solution of this LQR problem is accomplished by assuming that the first order variations of the costates satisfy $\delta \underline{p} = K \delta \underline{x} + C \delta \underline{\beta}$, and by using the maximum principle reducing the problem to solving a time variant matrix Riccati equation, as described in [10]. This assumption can be validated with a trivial extension to the validation of a similar assumption made in [10]. For the LQR problem detailed in equations (4.24) - (4.29) with equation (4.30) added to the performance measure, this results in the time variant matrix differential equations

$$\begin{bmatrix} \dot{K} \dot{C} \end{bmatrix} = \begin{bmatrix} -K I_n \end{bmatrix} \begin{bmatrix} M_d - M_c M_a^{-1} M_b \end{bmatrix} \begin{bmatrix} I_n & 0 \\ K & C \\ 0 & I_l \end{bmatrix} \tag{4.31}$$

where

$$\begin{bmatrix} K C \end{bmatrix}_{s=1} = \begin{bmatrix} f_{tx} \underline{\psi}_x^T \end{bmatrix} - f_{tu} (f_{tuu} + V_t)^{-1} \begin{bmatrix} f_{tux} \underline{\psi}_u^T \end{bmatrix} \tag{4.32}$$

$$M_a(s) = \begin{bmatrix} T^2 H_{uu} & 2T H_u^T \\ 2T H_u & 2H \end{bmatrix} + V_0 \tag{4.33}$$

$$M_b(s) = \begin{bmatrix} T^2 H_{ux} & T^2 f_u^T & T^2 \underline{L}_u^T \\ 2T H_x & 2T f^T & 2T \underline{L}^T \end{bmatrix} \tag{4.34}$$

$$M_c(s) = \begin{bmatrix} T^2 f_u & 2Tf \\ -T^2 H_{xu} & -2T H_x^T \end{bmatrix} \tag{4.35}$$

$$M_d(s) = \begin{bmatrix} T^2 f_x & 0 & 0 \\ -T^2 H_{xx} & -T^2 f_x^T & -T^2 \underline{L}_x^T \end{bmatrix} . \tag{4.36}$$

where n is the number of states in the powertrain model and l is the number of elements in the column vectors $\underline{L}$ and $\underline{\psi}$. These equations can be solved, by simulating them backwards in time in order to calculate the matrices, $K(s)$ and $C(s)$. In practice, this is accomplished by using ACSL, with the fourth order Runge-Kutta integration procedure and an integration step size of 0.01 in s. Again, in order to prevent ill-conditioning problems when inverting the positive definite matrices, and in order to reduce computational errors, the controls, states, costs and constraints are normalised. The control and state normalisation values used are just the root mean squared value over the interval [0,1] in s of the control and state values resulting from the optimal control problem solution in the previous chapter. The cost and constraint normalisation values are as in table 3.4. For the additional variational constraint bounding the value of the engine torque if it achieves its maximum bound, the normalisation value used is the engine torque maximum value. With the above matrices now calculated, the solution of the variational optimal control problem can be formed, with the optimal control variations determined by

$$\begin{pmatrix} \Delta \underline{u}(s) \\ \Delta T(s) \end{pmatrix} = -M_a^{-1} M_b \begin{bmatrix} I_n & 0 \\ KC & \\ 0 & I_l \end{bmatrix} \begin{pmatrix} \delta x(s) \\ \delta \beta \end{pmatrix} \tag{4.37}$$

$$\Delta u|_f = -(f_{tuu} + V_t)^{-1} \left[f_{tux} \underline{\psi}_u^T \right] \begin{pmatrix} \delta \underline{x}_f \\ \delta \underline{\beta} \end{pmatrix} . \tag{4.38}$$

These last formulae provide a solution to the variational optimal control problem for initial state perturbations if the Lagrange multipliers are known, deter-

mining $\delta\underline{\beta}$. For initial state perturbations, the Lagrange multiplier values are constant, but if they are allowed to vary then they can be used to cope with other more complex perturbations such as parameter variations, which can be interpreted as continual state perturbations. This arises from the observation that the variational optimal control problem solution at any given point s, equations (4.37) and (4.38), is independent from the variation values of the states and controls prior to that point. This allows the state perturbations at any given point in s to be treated as initial state perturbations. At such a point, the calculation of these Lagrange multipliers is achieved by estimating the variations of the constraints. Given a value of the current independent variable value, $\hat{s}$, the variations in the constraints up until that point can be calculated directly from

$$\begin{pmatrix} \Delta\underline{R}_0(\hat{s}) \\ \Delta\underline{Q}_0(\hat{s}) \end{pmatrix} = W^T \int_0^{\hat{s}} T^2(\underline{L}(\underline{x},\underline{u}) - \underline{L}(\underline{x}^*,\underline{u}^*))ds' \tag{4.39}$$

where W is as in equation (4.26) and superscript $*$ indicates the open loop solution states and controls, assuming that all the states and controls are known up to that point. The variations of the constraints after such a point in s can be approximated using the first variations of the constraints and eliminating the controls by substituting in equations (4.37) and (4.38), with the values of $\delta\underline{\beta}$ being subsequently eliminated by substituting in equation (4.26). Similarly for the linearised state equations, the controls can be eliminated by the substitution of these equations into equation (4.25). This results in the variational approximations of the constraints and state equations

$$\begin{pmatrix} \widetilde{\delta\underline{R}} \\ \widetilde{\delta\underline{Q}} \end{pmatrix} = \int_{\hat{s}}^{1} N_1(s)\delta\underline{x}(s) + N_2(s)\delta\underline{\lambda}(\hat{s})ds + N_3\delta\underline{x}_f + N_4\delta\underline{\lambda}(\hat{s}) + \begin{pmatrix} \Delta\underline{R}_0(\hat{s}) \\ \Delta\underline{Q}_0(\hat{s}) \end{pmatrix} \tag{4.40}$$

$$\delta\dot{\underline{x}} = N_5(s)\delta\underline{x}(s) + N_6(s)\delta\underline{\lambda}(\hat{s}). \tag{4.41}$$

In the above equation, the constant terms resulting from an unconverged solution are neglected, as for now they can be represented by an additional Lagrange

multiplier, which will later be set to unity. The state equation approximation, can be used to eliminate $\delta\underline{x}(s)$ from the variational constraint approximations by simulating equation (4.41) backwards in s over a basis of the states in order to evaluate

$$\delta\underline{x}(s) = \Phi(s)\delta x_f + \Psi(s)\delta\underline{\lambda}(\hat{s}) \tag{4.42}$$

which is substituted into equation (4.40). The variations in the integrand in this equation are now no longer dependent on s allowing the integrals to be evaluated, resulting in the constraint variations being dependent on the Lagrange multipliers and the terminal state variations. Again using equation (4.42) the terminal state variations can be related back to the states variations at $\hat{s}$, as Φ by definition is invertible. This results in the variational approximations for the constraints of

$$\begin{pmatrix} \widetilde{\delta R} \\ \widetilde{\delta Q} \end{pmatrix} = M_1(\hat{s})\delta\underline{x}(\hat{s}) + M_2(\hat{s})\delta\underline{\lambda}(\hat{s}) + \begin{pmatrix} \Delta\underline{R}_0(\hat{s}) \\ \Delta\underline{Q}_0(\hat{s}) \end{pmatrix} \tag{4.43}$$

$$\text{where } M_1(\hat{s}) = \int_{\hat{s}}^{1} N_1(s)\Phi(s)\Phi^{-1}(\hat{s})ds + N_3 \tag{4.44}$$

$$M_2(\hat{s}) = \int_{\hat{s}}^{1} N_2(s) + N_1(s)\left[\Psi(s) - \Phi(s)\Phi^{-1}(\hat{s})\Psi(\hat{s})\right] ds + N_4 \tag{4.45}$$

In practice the evaluation of these approximations are performed using Matlab and ACSL. The computer package ACSL, is used to evaluate Φ and Ψ in equation (4.42) and the evaluation of the integrals in equations (4.44) and (4.45). Some difficulties in the evaluation of Φ and Ψ were encountered due the difference in order of some of the eigenvalues, with some values diverging off to infinity. This problem was solved by partitioning the interval $[0, 1]$ in s into segments $[s_i, s_{i+1}]$ along which equation (4.41) is solved backwards in time. At the boundaries, the initial conditions are reset, calculating a series of Φ_i's and Ψ_i's such that for $s' \in [s_i, s_{i+1}]$

$$\delta\underline{x}(s') = \Phi_i(s')\delta x(s_{i+1}) + \Psi_i(s')\delta\underline{\lambda}(\hat{s}). \tag{4.46}$$

From these values, $\Phi(s)\Phi^{-1}(\hat{s})$ and $\Psi(s)-\Phi(s)\Phi^{-1}(\hat{s})\Psi(\hat{s})$ are directly calculated, using the iterative formula coded in Matlab

1. Initially let $A = \Phi_j^+(\hat{s})$ and $B = -\Phi_j^+(\hat{s})\Psi_j(\hat{s})$ where $\hat{s}$ lies on the jth segment of the partition.

2. For i from $j+1$ to k, with s lying in the kth interval calculate

$$A = \Phi_i^+(s_i)A \tag{4.47}$$

$$B = B - \Phi_i^+(s_i)\Psi_i(s_i). \tag{4.48}$$

3. Finally set

$$\Phi(s)\Phi^{-1}(\hat{s}) = \Phi_k(s)A \tag{4.49}$$

$$\Psi(s) - \Phi(s)\Phi^{-1}(\hat{s})\Psi(\hat{s}) = \Phi_k(s)B + \Psi_k(s). \tag{4.50}$$

where Φ_i^+ is the Moore-Penrose pseudo inverse of Φ_i, allowing the evaluation of equations (4.44) and (4.45) using ACSL. The use of the pseudo inverse prevents ill-conditioning problems, but results in some loss of information in the calculation of the pseudo inverse, as explained in [70]. For the problem in question, this loss of information relates to neglecting terms which have very little effect on the evaluation of equations (4.44) and (4.45).

Now that approximations of the constraints have been calculated, the Lagrange multipliers can be evaluated. For the equality constraints, this is done by taking $\widetilde{\delta R} = 0$, whilst for the inequality constraints from condition (4.23), either the Lagrange multiplier $\delta\lambda_k = 0$ or for the corresponding inequality constraint $\widetilde{\delta Q_k} = 0$. This last condition results in 2^q cases being considered, where q is the number of inequality constraints in the variational optimal control problem, the permutations of each inequality constraint or Lagrange multiplier being zero. For each case,

equation (4.43), can be used to calculate 2^q Lagrange multipliers sets, using

$$\delta\underline{\lambda}_i(\hat{s}) = -\Lambda_i \left[\Lambda_i^T M_2(\hat{s})\Lambda_i\right]^+ \Lambda_i^T \left\{ \begin{pmatrix} \Delta\underline{R}_0(\hat{s}) \\ \Delta\underline{Q}_0(\hat{s}) \end{pmatrix} + M_1(\hat{s})\delta\underline{x}(\hat{s}) + \text{constant} \right\}. \quad (4.51)$$

where Λ_i is a $q+p$ by $\tilde{q}+p$ matrix, with p the number of equality constraints, q the number of inequality constraints and $\tilde{q}$ number of active inequality constraints (i.e number of none zero inequality Lagrange multipliers). This matrix selects the case and consists of column vectors with zero entries apart from one entry set to unity which identifies which constraints are active. The constant term, results from the additional Lagrange multiplier representing the neglected constant terms in equations (4.40) and (4.41), which is now set to unity. Again, for safety the Moore-Penrose pseudo inverse is used, although for variational constraints sufficiently linearly independent this should not be needed. For each case if the Lagrange multipliers are substituted back into equations (4.37) and (4.38), the variational optimal control solution

$$\begin{pmatrix} \Delta u(\hat{s}) \\ \Delta T(\hat{s}) \end{pmatrix} = A_i(\hat{s})\delta x(\hat{s}) + B_i(\hat{s}) \begin{pmatrix} \Delta\underline{R}_0(\hat{s}) \\ \Delta\underline{Q}_0(\hat{s}) \end{pmatrix} + c_i(\hat{s}) \quad (4.52)$$

$$\Delta u_f = A_{fi}\delta x_f + B_{fi} \begin{pmatrix} \Delta\underline{R}_0 \\ \Delta\underline{Q}_0 \end{pmatrix}_{s=1} + c_{fi} \quad (4.53)$$

results, where $\hat{s}$ denotes the current value of the independent variable and subscript i denotes the particular case resulting from a particular set of Lagrange multipliers. These final calculations are performed using Matlab. Note that the constant terms $c_i(\hat{s})$ and c_{fi} will disappear for a converged solution, or at least be negligible for controls at which the gradient of the performance measure is small. This last situation is often the case, for a control strategy calculated using the algorithm described in section 3.

4.4 Optimal feedback control

In section 4.2, the performance variation is approximated in terms of variations of the states and controls, with a solution in terms of control perturbations which minimise this variational performance measure being obtained in the previous section. Hence, if the controls are perturbed by the calculated amount, then the resulting performance measure should remain small, as small as is possible. In order to apply this solution to the clutch engagement problem, the terms in equations (4.52) and (4.53) must be calculated. If it is assumed that at any time after starting to engage the clutch, the current and past states and controls are known, along with the elapsed time and the open loop states and controls, then all the required terms can be calculated. In particular, the current independent variable $\hat{s}$, the current state perturbation, the current constraint perturbations $\begin{pmatrix}\Delta \underline{R}_0(\hat{s}) \\ \Delta \underline{Q}_0(\hat{s})\end{pmatrix}$ and the Lagrange multiplier set resulting from the 2^q cases considered, can all be determined.

The current time and the current independent variable are related by the 'dummy' control T, allowing the current independent variable to be calculated using the relationship

$$\hat{s} = \int_0^t \frac{1}{T^2} dt' = \int_0^t \frac{1}{(T^* + \Delta T)^2} dt'. \tag{4.54}$$

The calculation of the current independent variable then allows the current state perturbations to be calculated, by subtracting the current states from the open loop solution states. Similarly the current constraints variations can be calculated using equation (4.39), or far more simply by just integrating the difference in the constraint integrands up until the current time. Hence, in order to calculate the optimal variation, all that is required is to choose the Lagrange multiplier set. This last calculation is slightly more complex, but by substituting equation (4.51) for a particular case into equation (4.43), an estimate for the constraint variations

in terms of the current state perturbations and the current constraint perturbations, can be calculated. For each case, this allows the assessment of whether a particular Lagrange multiplier set results in a solution which satisfies the inequality constraints, with the cases that fail to satisfy the variational approximations of the constraints being rejected. For the remaining sets, estimates of the performance measure variations are obtained. This is done in an identical way to which equation (4.43) is evaluated, which when (4.51) is substituted in again determines an estimate for the performance measure which is dependent on the current state and constraint variations. This allows a unique Lagrange multiplier set to be chosen, the set which yields the smallest performance measure estimate and which satisfies the estimated constraints. Equation (4.51) could have been used to further refine the sets after rejecting the sets which fail the constraints, by rejecting the sets that fail to satisfy condition (4.23), but as little computational advantage results and as the use of a performance measure estimate has greater physical significance, this was not exploited.

The resulting optimal feedback control architecture is detailed in figure 4.3. In this feedback control strategy, given the Lagrange multiplier sets to be used, the real time evaluation requirements are just the on going evaluations of the integral terms in equations (4.39) and (4.54) and the evaluation of equations (4.52) and (4.53). These last evaluations just require the addition and multiplication of matrices dependent on the current independent variable $\hat{s}$, matrices which can be acquired from look-up tables. The real time requirements of the Lagrange multiplier set choice is, at most, the evaluation of 2^q expressions of the form

$$\begin{pmatrix} \delta\tilde{\underline{R}} \\ \delta\tilde{\underline{Q}} \\ \delta\tilde{F} \end{pmatrix} = M(\hat{s}) \begin{pmatrix} \delta\underline{x}(\hat{s}) \\ \Delta\underline{R}_0(\hat{s}) \\ \Delta\underline{Q}_0(\hat{s}) \\ 1 \end{pmatrix}, \tag{4.55}$$

the expressions that provide the cost and constraint variational estimates, where q is the number of inequality constraints. Again these evaluations only require the addition and multiplication of matrices dependent on the independent variable $\hat{s}$. These real time requirements are felt to be attainable for the clutch engagement problem.

4.4.1 Feedback clutch control analysis

In order to assess the performance of this feedback control strategy, simulations are carried out for the perturbations discused in section 4.1. Variations of the second optimal control problem performance measure in table 3.5 is used to construct the variational performance measure. This variational optimal control problem has two inequality constraints as well as the mandatory equality constraint, equation (2.53). These inequality constraints arise from the inequality constraint, equation (3.77), in the original problem, and the engine torque upper bound after clutch lock up. For the open loop solution to the optimal control problem, the control strategy detailed in figure 4.1, both of the inequality constraint bounds are reached by the calculated control strategy. Hence, neither constraint can be neglected resulting in four Lagrange multiplier sets. In the calculation of the required feedback matrices, the positive definite matrices, V_0 and V_t, are taken to be diagonal with the diagonal entries set to 10^4. The simulations of the feedback control strategies is performed using the computer package ACSL, which reads the calculated feedback matrices from memory and performs the required real time calculations as well as simulating the perturbed model equations. For each of the perturbations described in section 4.1, simulations of the feedback strategy are detailed in figures 4.4 - 4.6. In each figure four graphs are detailed, the first graph of the clutch plates speeds when the open loop control strategy is used to control the engagement, the second graph of the clutch plate speeds when the

feedback strategy is used to control the engagement, and the third and fourth graphs detailing the open loop and feedback controls. In these last two graphs the solid lines indicate the open loop controls with the dashed lines indicating the feedback or closed loop controls. In the third graph, both the required clutch torque capacity and the actual clutch torque capacity are detailed with the latter identified by its smooth transients.

For the first perturbation, clutch engagement with the initial engine flywheel speed lowered to 175 rad/s, when the feedback strategy is used, the engine flywheel speed still drops below its unperturbed level but not quite as far as when the open loop solution is used, a possible slight improvement. In doing this the clutch takes longer to engage with minor increases in the engine torque and a slight reduction in the clutch torque capacity prior to engagement. However, as the engine flywheel speed still drops below 100 rad/s when feedback is implemented, the feedback control strategy is not satisfactory. For the second perturbation, the hill start, when the feedback strategy is used, the clutch plates engage more smoothly with little oscillations after engagement. This improvement is achieved by small reductions in the engine torque and small increases in the clutch torque capacity, causing the clutch plate speeds to be nearly equal when the required clutch torque capacity is stepped up at the point when $s = 1$, preventing the occurrence of the oscillations after engagement. Note that the point when this step occurs has been moved to a slightly later time with modifications of the 'dummy' control. However, again the feedback clutch plate response is far from satisfactory with the engine flywheel speed dropping significantly below 100 rad/s, a deterioration on the open loop response. For the third perturbation, simulated by perturbing the model parameters in order to represent a fully laden vehicle with a worn clutch, the feedback response is improved with a reduction in the size of the oscillations after clutch lock up. Again this is accomplished by small

reductions in the engine torque and small increases in the clutch torque capacity. However, just prior to and after clutch lock up large control variations occur, in particular large variations of the 'dummy' control, causing a slight kink in the clutch plate speed at a time of about 1.2 seconds.

These three previous feedback simulations assume that both controls are active. The feedback procedure developed is just as applicable when only the clutch torque capacity is active as a control. To demonstrate the ability of the feedback procedure to control the three perturbation examples but with only variations in the clutch torque capacity, a further three simulations are detailed in figures 4.7 - 4.9. As in the open loop study, the engine torque is set to 100 Nm, both before and after clutch lock up. These one control simulations again show how, in general, the feedback control cause improvements in clutch control, reducing oscillations after engagement in all three simulations. However, in accomplishing this large control variations result, in particular, large control variations are prominent for the flywheel speed perturbation and the model parameter perturbation simulations, figures 4.7 and 4.9, with the control reaching its bounds. The greater size of the control variations with only one control active, indicates the greater difficulty in controlling engagement. Again these excessive control variations can cause undesirable dynamics, as in figure 4.7, with the clutch plates initially locking up, then briefly returning to a slipping state, before finally locking up, completing engagement. Note that due to the large value of engine torque, stalling problems do not occur in these simulations. However, for other cases with smaller engine torque values stalling can again be a problem with the feedback control again struggling to prevent stalling.

Simulations have also been conducted using the dataset set representing vehicle B. As already noted, clutch engagement is a little easier for this dataset as oscillations prior to clutch lock up are not significant, the reason for concentrating

on the dataset representing vehicle A. The only slight difficulty is the increased singularity of the matrices used in the calculation of the feedback matrices, with care having to be taken when inverting matrices by making greater use of the Moores-Penrose pseudo inverse. The results obtained from these simulations of vehicle B are very similar to the simulations presented, with no additional points to the points already noted being acquired.

To conclude, the simulations indicate that improvements in the clutch plates speeds do result from the use of the feedback control strategy, but that the cost measuring the likelihood of stalling is not always maintained at a small value, and that around the point of clutch lock up, in particular for model parameter perturbations and when only one control is active, large variations can result which can cause minor undesirable effects. These results have also been obtained for the simulation of other model and state perturbations.

4.4.2 Feedback control modifications

The failure of the feedback strategy to maintain the costs measuring the likelihood of stalling at small values can be easily explained. This failure is due to the variations of the costs measuring stalling, F_6 and F_9, being zero for control strategies which maintain the engine flywheel speed above its minimum permissible value. Hence for such control strategies, such as the open loop solution detailed in figure 4.1, the variational optimal control problem will not include any terms equating variations of the controls to terms measuring the likelihood of stalling. No ideal solution to this difficulty has been found, however, if the variational cost $w_x \delta x_1^2$ is added to the variational optimal control problem performance measure, then for the resulting closed loop solution, the flywheel speed should remain close to the open loop value, which if remaining above x_{1min} ensures that the closed loop flywheel speed will remain above x_{1min}. In other words, for cases where the fly-

wheel speed is maintained above x_{1min} in the open loop solution, the closed loop flywheel speed can be maintained above x_{1min} by introducing a flywheel speed tracking condition.

To demonstrate the success of this feedback control modification, a simulation of the first perturbation example, clutch engagement with a reduced engine flywheel speed, is conducted with the weighting parameter w_x set to 10^5 and both controls active. This simulation, detailed in figure 4.10, demonstrates that this modification, at least for this particular example, eliminates the problem of stalling, with the flywheel speed being maintained above 100 rad/s. This is accomplished, almost entirely, by large variations of the engine torque in the first 0.2 seconds, which cause the flywheel speed to acquire its open loop optimal value in this period. The marked behaviour of the engine torque unilaterally controlling the engine flywheel speed, results from the dynamics of the flywheel being decoupled from the dynamics of the rest of the powertrain when the clutch is slipping, as discussed in chapter 2.1. Furthermore, the engine torque variations only effect the flywheel dynamics, allowing the engine torque to be used to control the flywheel speed, preventing stalling, without interfering with the control aims for the remainder of the powertrain. The engine torque is still required to control the point of lock up, stepping up and down as required, with the dynamics no longer decoupled after this time. This minor difficulty can be overcome by ensuring that the majority of the engine torque variations, controlling the flywheel speed, occur well before the point of lock up, as in figure 4.10, in which only small variations occur after a time 0.2 seconds. This is possible for the example in figure 4.10, as no other perturbations are present, indicating the effectiveness of this modification for coping with flywheel speed perturbations, a very likely occurrence in clutch control. This modification, also has limited success in preventing stalling for other perturbations, as found by conducting other feedback simulations, but

unfortunately, it was found that a compromise between preventing stalling and achieving the other control aims sometimes results. For problems in which the prevention of stalling is extremely difficult, this feedback approach still struggles to prevent stalling. As the prevention of stalling is imperative, the suggestion for over coming this difficulty is to use a feedback control strategy designed specifically for preventing stalling, such as the centrifugal clutch control method discussed in the introduction, in these cases, switching back to the optimal feedback control strategy when the problem of stalling has diminished. In essence, this is what the weighting parameter w_x does, determining how much of the control strategy should be concerned with preventing stalling by tracking the open loop value, and how much concerned with minimising the variational performance measure.

The other problem identified with the feedback control strategy simulations, the large variations around the point of clutch lock up, is attributed to the requirement of the variational constraints being satisfied. These large variations are unsatisfactory as they can cause the control bounds to be reached, can cause the neglected higher order variations to become significant and together with fast transient fluctuations can cause undesirable responses as previously noted. The large control variations are required to satisfy the constraints when the independent variable s is approaching unity, if small but significant values of the constraint variations still persist. It can be immediately appreciated that for parameter perturbations, as the control strategies are determined using the unperturbed state equations, that such problems are more severe. The solution to this problem is a little more complex than the previous problem, but by relaxing the variational constraints it can be achieved. Equations (4.52) and (4.53) relate the optimal control variations to the state variations and the constraint variations, at any given independent variable value s. If instead of calculating the equations by setting the variational constraint approximations to zero, they are set to an error value $\underline{\epsilon}$

then by careful choice of this error term large control variations can be prevented whilst still achieving the constraints to a reasonable degree of satisfaction. This compromise between small constraint errors and control actuation is accomplished by minimising a weighted sum of the control perturbations and the error value, in other words by minimising

$$\frac{1}{2}w_{\Delta u}^2\left(\Delta\underline{u}^T\Delta T\right)\begin{pmatrix}\Delta\underline{u}\\ \Delta T\end{pmatrix}+\frac{1}{2}\underline{\epsilon}^T\underline{\epsilon}. \tag{4.56}$$

If it is noted that the constraint errors can be achieved by applying the error vector to the measured value, $\begin{pmatrix}\Delta\underline{R}_0(\hat{s})\\ \Delta\underline{Q}_0(\hat{s})\end{pmatrix}$, in equations (4.52) and (4.53), then it can be shown that the minimal of equation (4.56) is satisfied by the control variations

$$\begin{aligned}\begin{pmatrix}\Delta\underline{u}\\ \Delta T\end{pmatrix}=&\left[I_{m+1}-w_{\Delta u}^2B_i(\hat{s})\Lambda_i\left(w_{\Delta u}^2\Lambda_i^TB_i^T(\hat{s})B_i(\hat{s})\Lambda_i+I_{p+\bar{q}}\right)^{-1}\Lambda_i^TB_i^T(\hat{s})\right]\\ &\times\left(A_i(\hat{s})\delta x(\hat{s})+B_i(\hat{s})\begin{pmatrix}\Delta\underline{R}_0(\hat{s})\\ \Delta\underline{Q}_0(\hat{s})\end{pmatrix}+c_i(\hat{s})\right)\end{aligned} \tag{4.57}$$

$$\begin{aligned}\Delta\underline{u}_f=&\left[I_m-w_{\Delta u}^2B_{fi}\Lambda_i\left(w_{\Delta u}^2\Lambda_i^TB_{fi}^TB_{fi}\Lambda_i+I_{p+\bar{q}}\right)^{-1}\Lambda_i^TB_{fi}^T\right]\\ &\times\left(A_{fi}\delta x_f+B_{fi}\begin{pmatrix}\Delta\underline{R}_0\\ \Delta\underline{Q}_0\end{pmatrix}_{s=1}+c_{fi}\right)\end{aligned} \tag{4.58}$$

where m is the number of controls, p is the number of inequality constraints, $\bar{q}$ is the number of active inequality constraints, Λ_i is as in equation (4.51), $w_{\Delta u}$ is the weighting factor in the previous equation and A_i, B_i, A_{fi}, B_{fi} are as in equations (4.52) and (4.53). As with the final calculation of equations (4.52) and (4.53), these calculations are performed using Matlab. Note that these new equations are equivalent to equations (4.52) and (4.53) when the weighting factor $w_{\Delta u}$ is set to zero. Also note, that this modification requires no additional real time computations, as the calculations result in the control variations being dependent on matrix equations of the same form as equations (4.52) and (4.53).

To demonstrate the success of this modification, this new feedback strategy is applied to the three cases previously presented, for which large control variations result with the unmodified feedback strategy. That is, the parameter perturbation example, with both one and two controls active, and the initial flywheel speed perturbation with only the clutch torque capacity active as a control. In calculating the modified feedback matrices, the weighting parameter, $w_{\Delta u}$ is set to 0.01, with the new simulations detailed in figures 4.11 - 4.13. For the parameter perturbation example with two controls active, as previously noted in figure 4.6, a kink in caused by large variations just after the point of clutch lock up. For the same simulation, but using the modified feedback control strategy, this kink has disappeared, with the simulation detailed in figure 4.11. This, as expected, is as a result of smaller control variations. It should be noted that in figure 4.11, the feedback controls are shifted in time by variations of the 'dummy' control, making the control variations appear larger than they really are, with the feedback controls shadowing the open loop controls. For the parameter perturbation example with only one control active, again large control variations result, although no kink occurs. Once again, as shown in figure 4.12, the modification reduces the size of the variations, removing the sharp dip that occurs just before 1.2 seconds after starting engagement. It should be noted that this dip, can not be blamed for the oscillations after engagement, as oscillations still persist on its removal. However, the feedback control does reduce the size of the oscillations, an improvement on the open loop controls. For the final modified feedback simulation, the perturbation example with the reduces engine flywheel speed, quite large control variations still persist. However, the size of the variations has decreased significantly from when the modification is not implemented, and the problem of the clutch plates locking up twice before completing engagement has been removed. The proposed reason for large variations persisting, is that significant variations below the open

loop clutch torque capacity is required to prevent the clutch engaging too quickly, resulting in large oscillations. The size of the variations can still be reduced by distributing the variations over the whole interval in s. One way of achieving this, is by increasing the weight $w_{\Delta u}$, in practice, for this case a value of 0.1 was found to be sufficient to prevent excessive variations. To conclude, this last modification is extremely effective at reducing large control variations at or just prior to clutch lock up, removing undesirable dynamics that these large variations create.

4.4.3 Generalisation of clutch feedback control

The feedback control engagement strategy developed, arises from taking variations of the open loop optimal control problem. Unfortunately, the approach presented is only applicable to the case when the performance measure is constructed using the weighted sum multi-objective technique. For this reason, some effort has been applied to the problem of generalising the feedback control derived from a weighted sum performance measure to feedback control strategies relating to goal attainment performance measures. In other words, trying to extend the approach to a feedback technique which maintains variations of performance measures designed using one of the goal attainment methods. As an example, one such approach that was proposed, is to represent the goal attainment performance measure form by

$$F = H_i \tag{4.59}$$

$$\text{subject to } H_j - Hi < 0 \text{ for all } j \neq i \tag{4.60}$$

for some i, where H_i are the individual cost functionals. In other words, relating this back to the geometric representation of the goal attainment method described in section 2.4, representing the goal attainment value by the facets of the corner of points in the cost space of equal value. This allows the variations of the goal attainment performance measure to be represented by one of the n variations,

constructed by taking variations of the above performance measure form, where n is the number of cost functionals included in the multi-objective compromise. This is possible, as on each facet the performance measure is of the same form as when the weighted sum method has been employed. This fact also allows controls variations to be obtained for each facet, using exactly the same algebraic manipulation as previously presented. This just leaves the choice of which facet to use to generate the control variations. The proposed approach is to use estimates of the performance measure variation, as with the choice of Lagrange multiplier set, choosing the facet which results in the smallest estimated performance measure value. The problem with this approach is that for each facet, an extra $n - 1$ inequality constraints have been introduced. As the computational requirements increase exponentially with the number of constraints, this dramatically increases the real time computational requirements, one of the reasons for not continuing with this approach.

The main reason for not exploring the generalisations of feedback clutch engagement fully, is that as small values of performance measures equate to good clutch performance, whatever multi-objective technique is employed, a control strategy producing small values of a performance measure designed using a goal attainment method, will also produce small values for a performance measure designed using the weighted sum method. It is felt that the choice of multi-objective method is far more important in locating a 'good' control strategy than for estimating the performance of clutch engagement around such a control strategy. Indeed, if the feedback control strategy is applied to a perturbation example, where the open loop solution is obtained using one of the goal attainment methods, then the performance of the clutch is still maintained, maintained just as with a weighted sum performance measure. This validates the approach of using the described feedback approach whatever the multi-objective method used in

creating the performance measure.

To conclude, the feedback clutch control method described in this chapter, enables 'good' clutch control to be maintained for the majority of realistic state and parameter perturbations. This method can be applied to any 'good' solution obtained using the method described in the previous chapter, and indeed, due to the first variations remaining in the variational optimal control problem, can be applied to other non optimal solutions, so long as the neglected higher order variations are not significant. Finally, the real time computational requirements are felt to be achievable in implementing this feedback strategy.

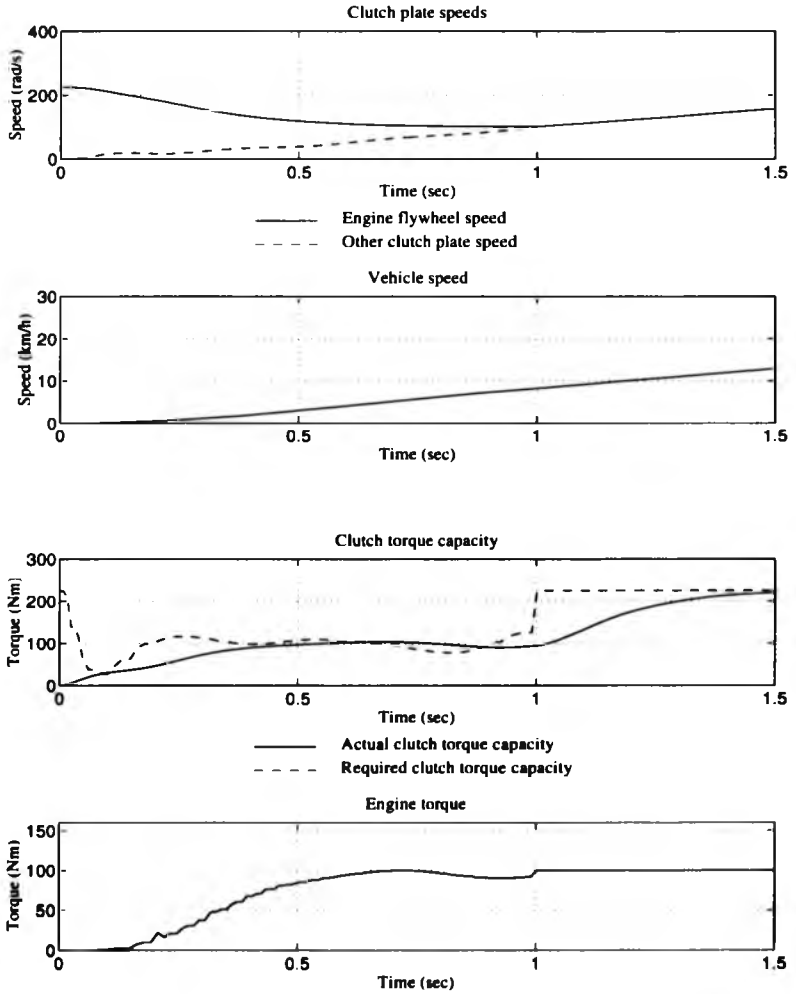


Figure 4.1: Clutch performance of open loop control strategy

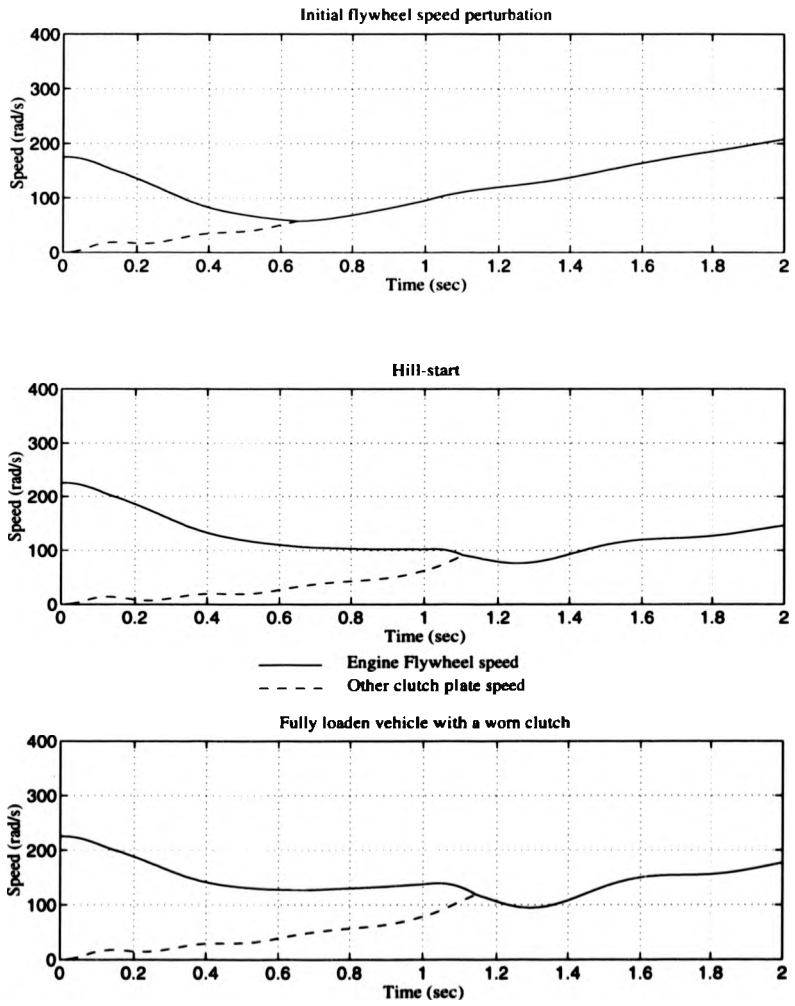


Figure 4.2: Clutch plate speed simulations for perturbed clutch engagement

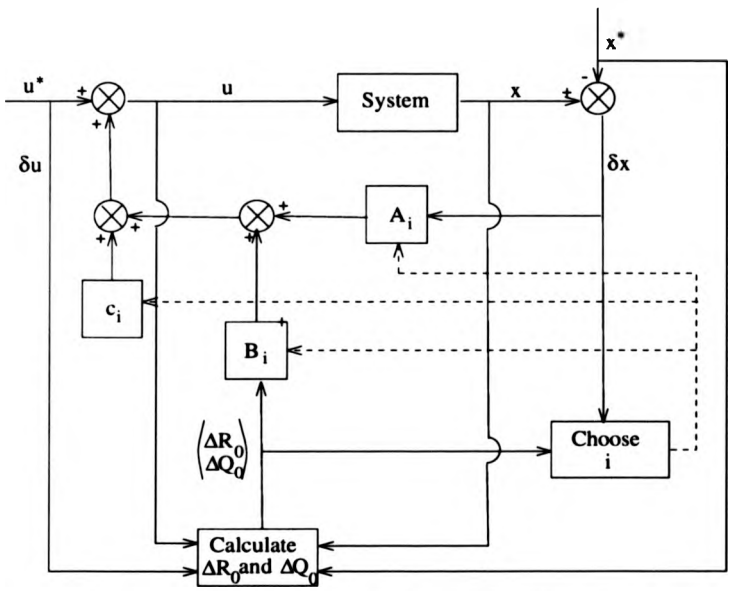


Figure 4.3: Optimal feedback control architecture

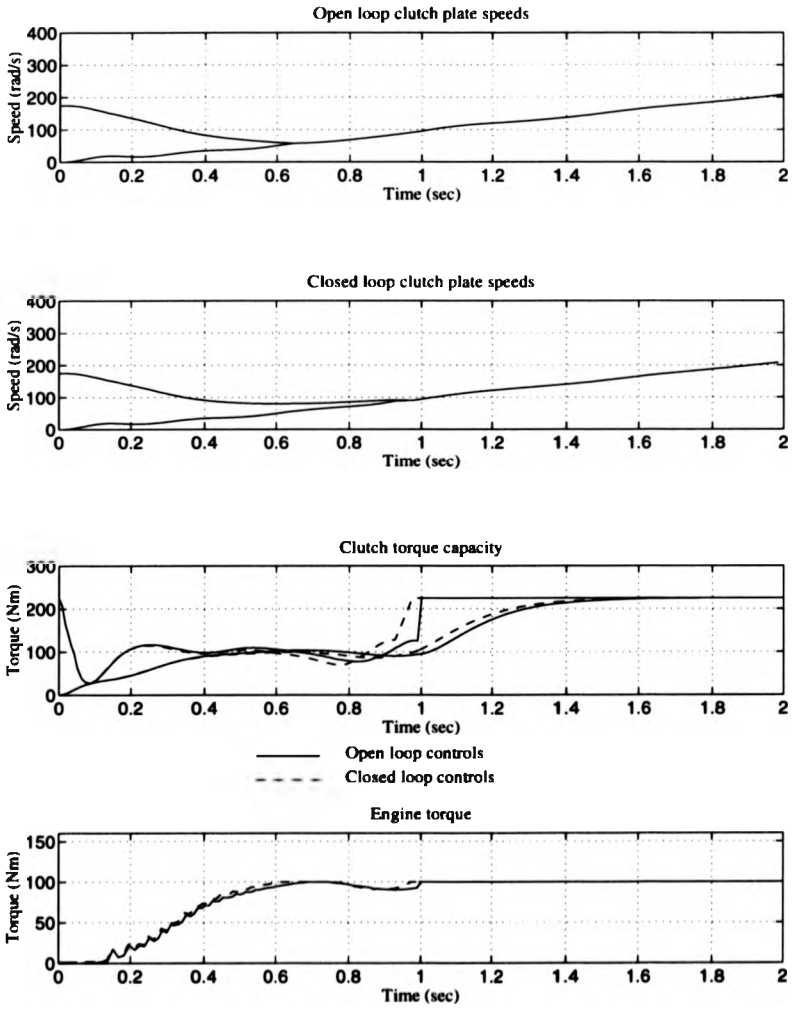


Figure 4.4: Feedback control with a reduced initial engine flywheel speed

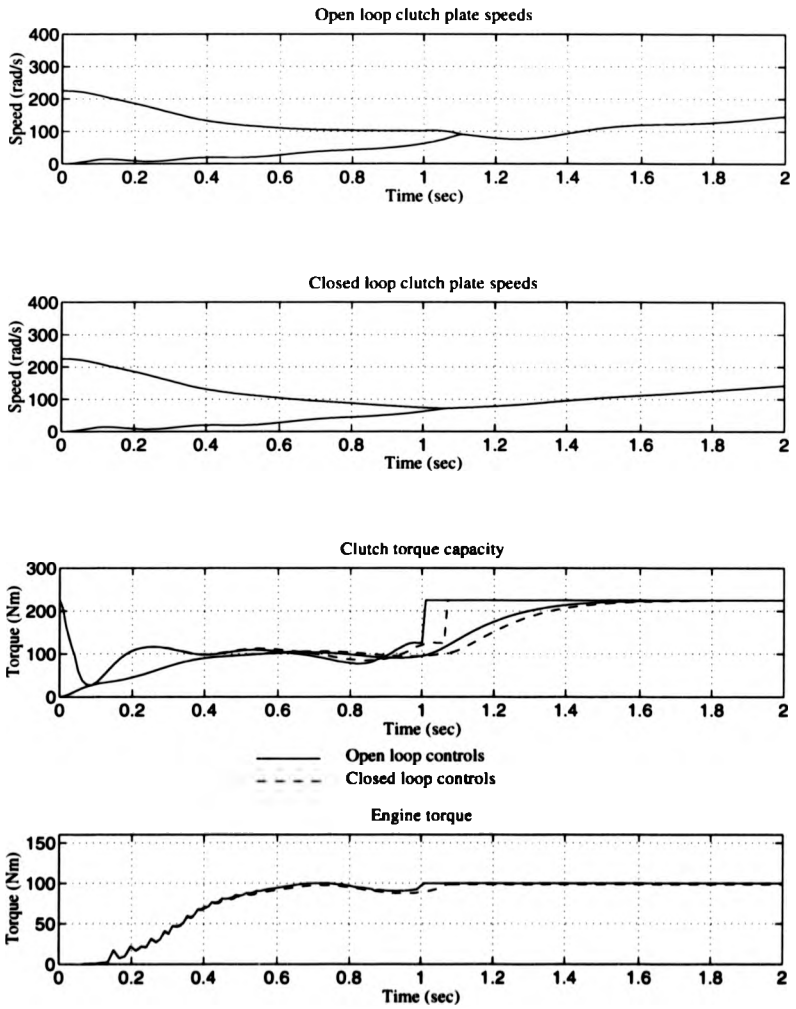


Figure 4.5: Feedback control of a hill start

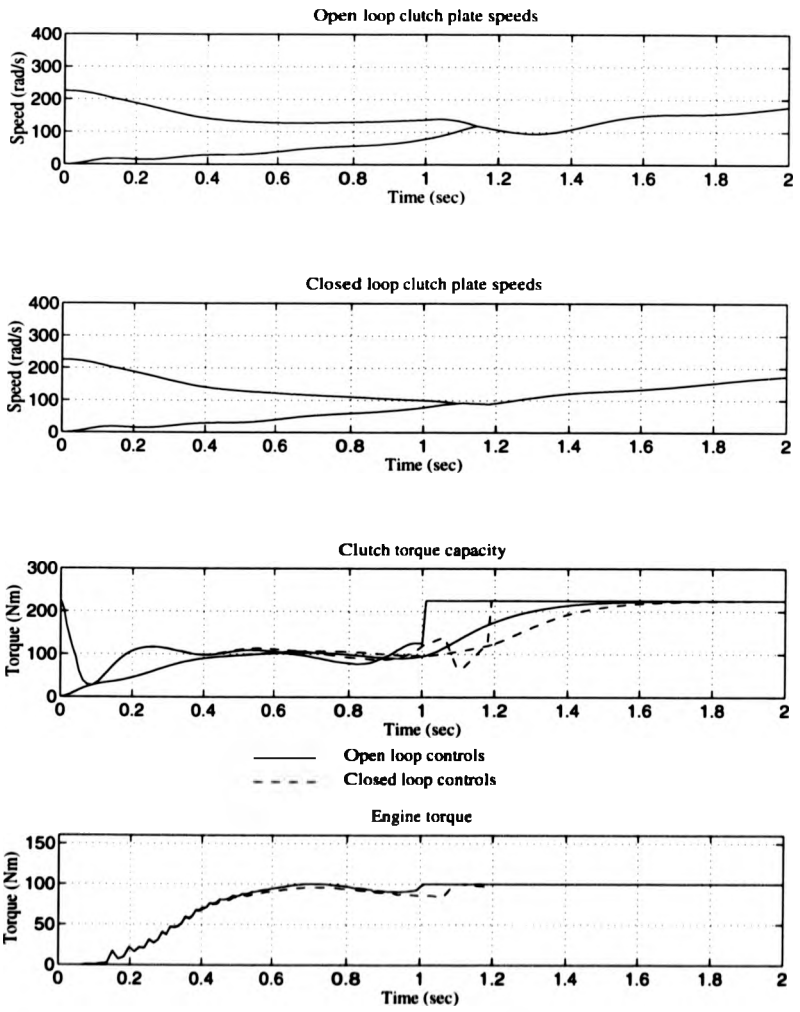


Figure 4.6: Feedback control of a fully laden vehicle with a worn clutch

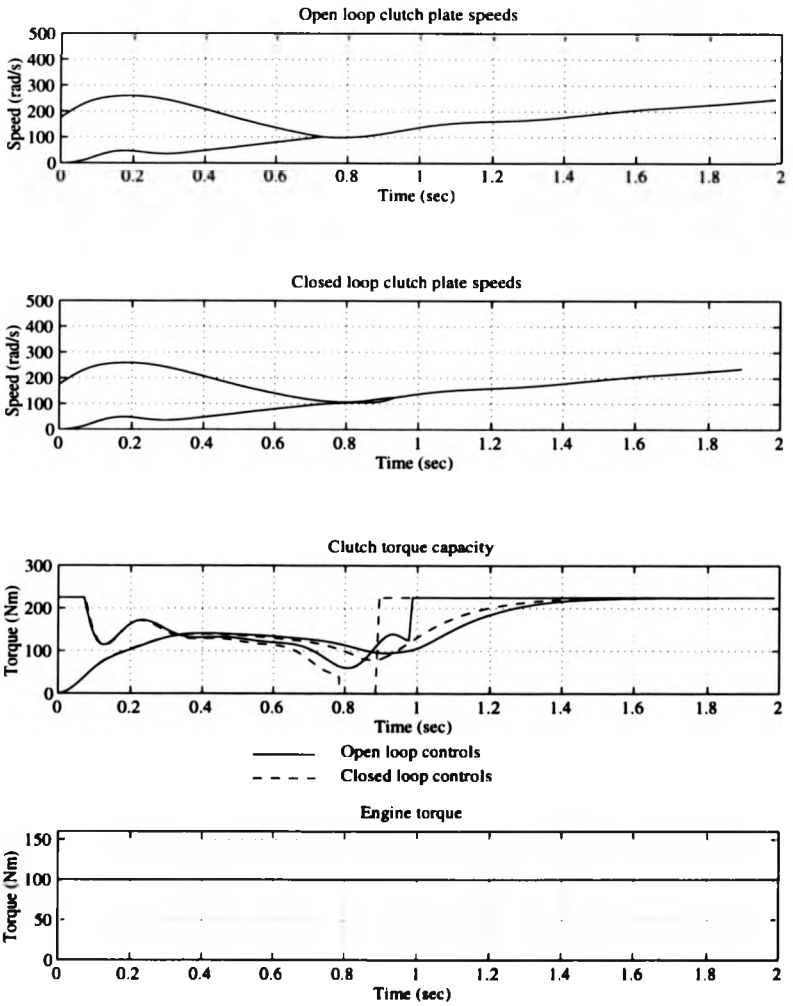


Figure 4.7: Feedback control with a reduced initial engine flywheel speed

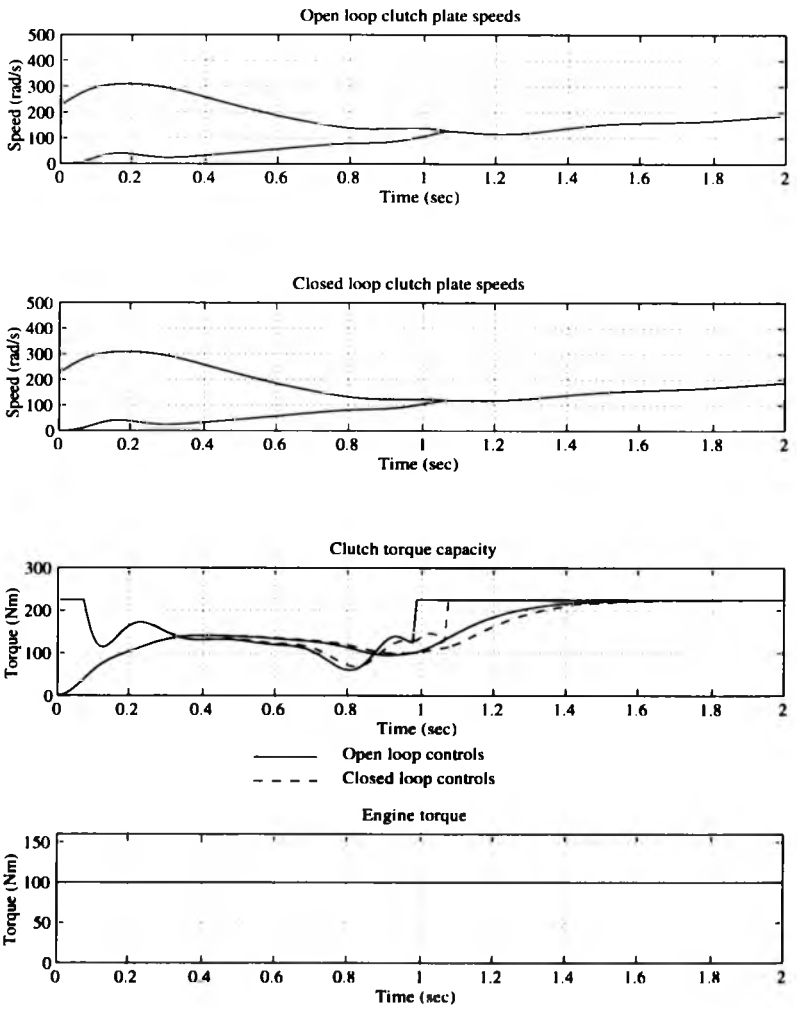


Figure 4.8: Feedback control of a hill start: only one control active

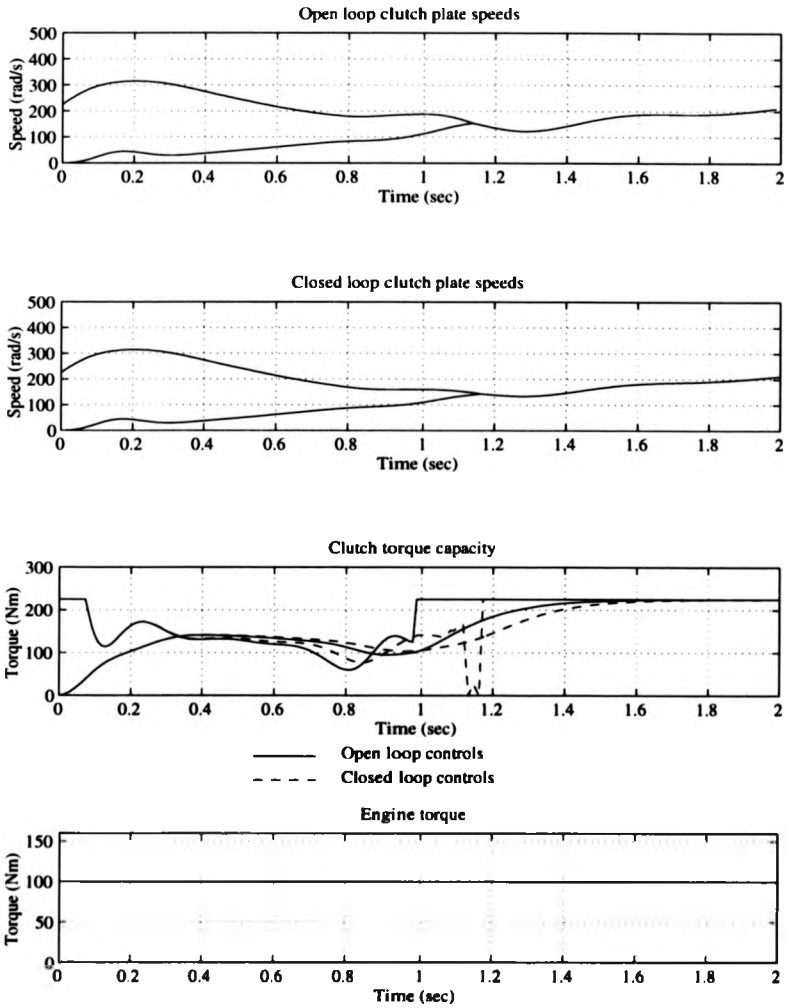


Figure 4.9: Feedback control of a fully laden vehicle with a worn clutch: only one control active

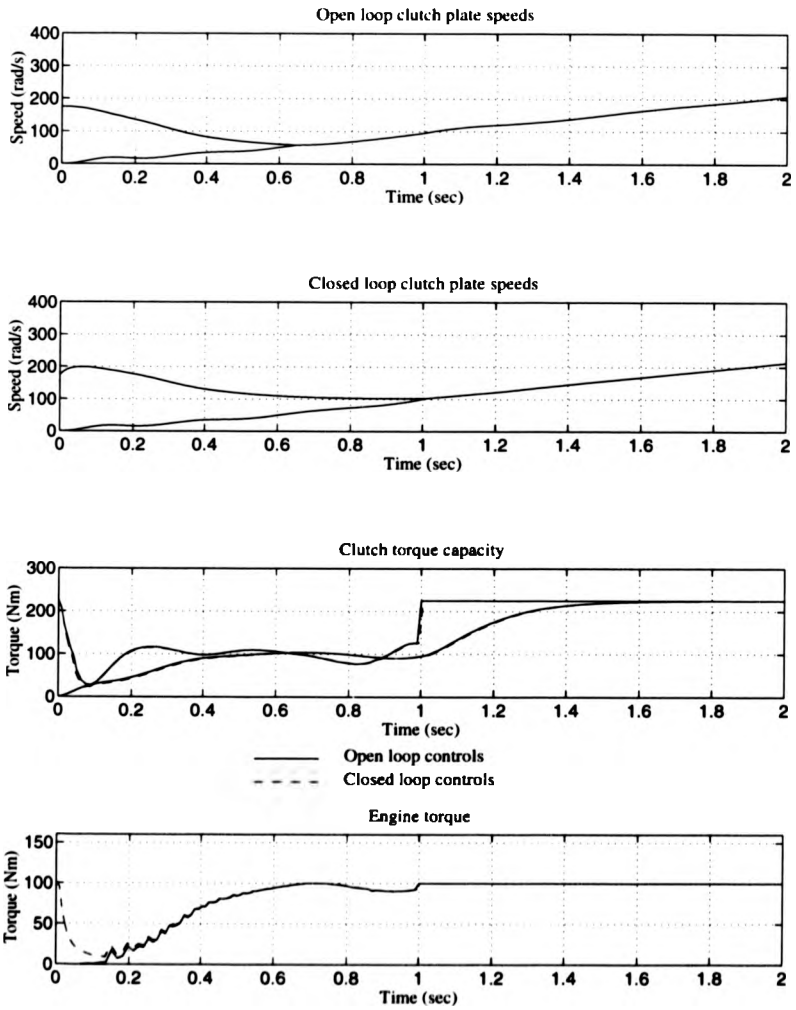


Figure 4.10: Feedback control with a reduced initial engine flywheel speed, with the flywheel speed tracking its open loop value

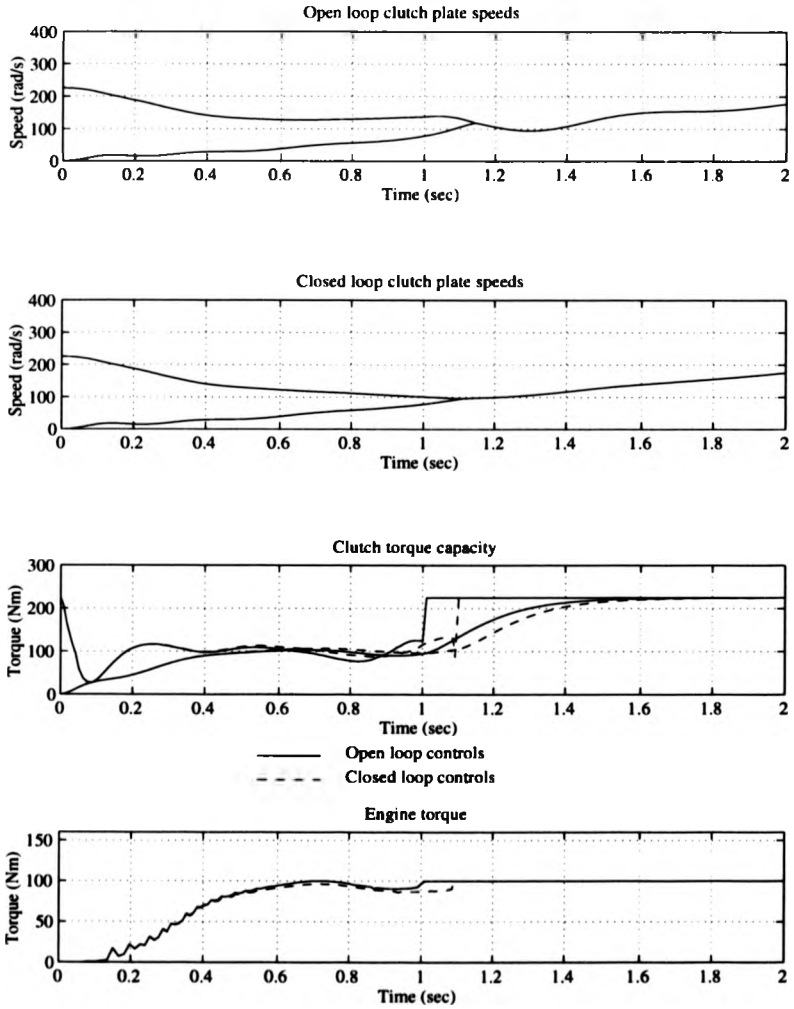


Figure 4.11: Modified feedback control: parameter perturbation example with both control active

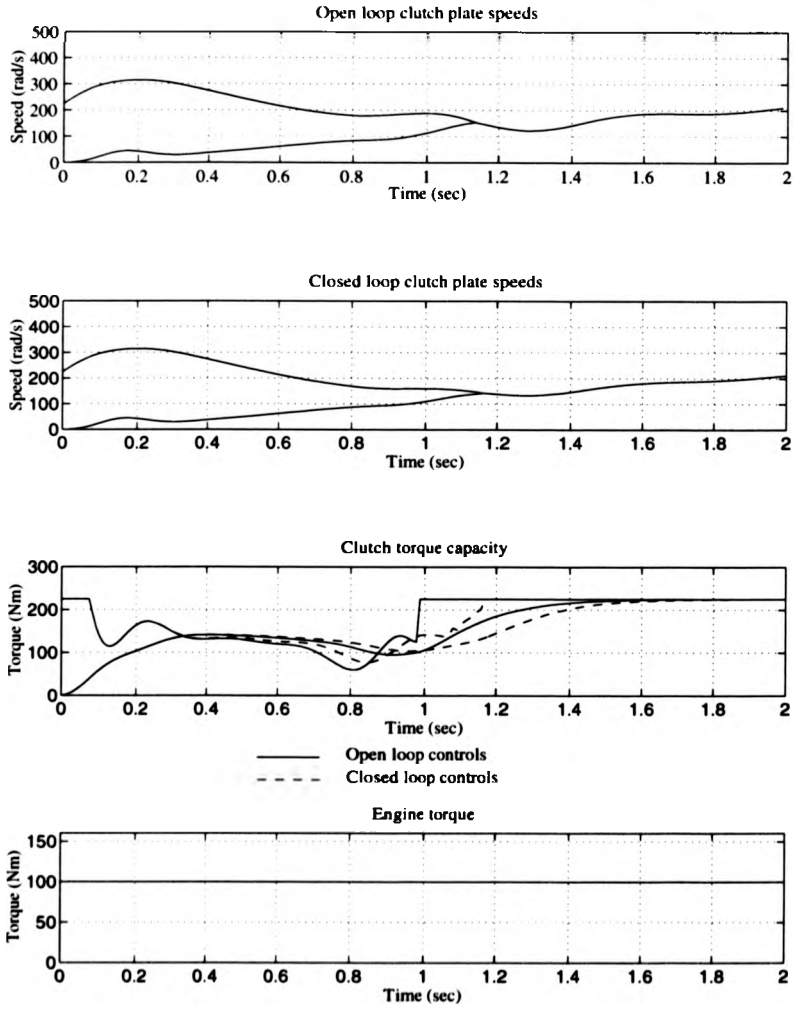


Figure 4.12: Modified feedback control: parameter perturbation example with one control active

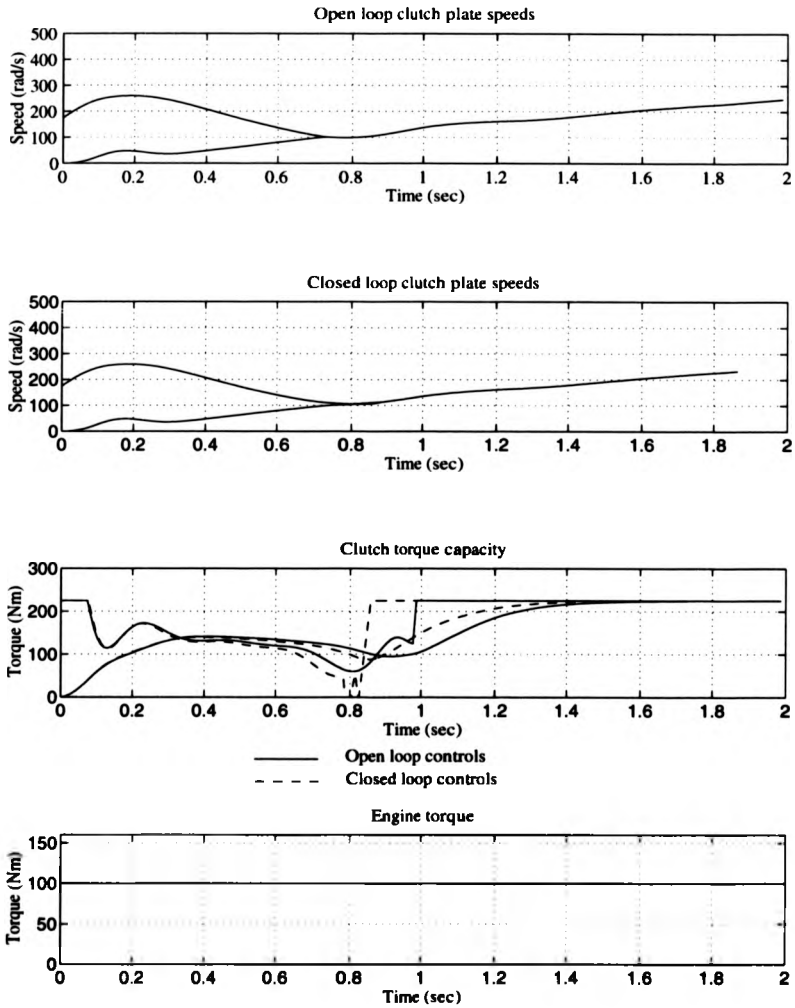


Figure 4.13: Modified feedback control: reduced initial flywheel speed with one control active

Perturbation	Performance measure	Individual cost functionals							
		F_1	F_2	F_3	F_5	F_6	F_7	F_9	F_{10}
None	1.059	0.128	0.004	0.019	0.273	0.000	0.501	0.000	0.134
Flywheel speed	209.868	0.121	0.006	0.312	0.120	27.805	0.323	181.048	0.133
Hill-start	11.637	0.162	0.916	1.098	0.377	0.091	0.552	8.296	0.146
Fully laden/ worn clutch	5.340	0.192	1.463	2.585	0.396	0.000	0.571	0.000	0.135

Table 4.1: Cost functionals for perturbed clutch engagement

Chapter 5

Optimal clutch engagement with variable model parameters

In this chapter, powertrain model parameters that may vary are identified, grouping them into classes describing how fast these parameters vary in time. Simulations are then conducted, illustrating the significance of each parameter variation, as well as the ability of the feedback control to maintain 'good' clutch engagement under such variations. As a result of these simulations, a revised feedback control strategy is developed, which utilises known parameter variations. As parameter variations may not be known, a method is then presented which yields estimates of such variations, one estimate for each clutch engagement. Finally, a brief outline of how such information can be used to design self learning/adaptive control strategies is discused.

5.1 Clutch engagement with realistic parameter perturbations

In the previous chapter, it was demonstrated that the feedback control is able to cope with certain parameter perturbations. In this section, a far more comprehensive investigation of the effects of parameter perturbations on feedback clutch control is conducted.

The dynamics of perturbations in automotive powertrains varies significantly from one type of perturbation to another. For this reason, the types of parameter perturbations have been grouped into four classes, with examples of particular reference to the powertrain model presented in chapter 2 given:

STATIC Perturbations present between different vehicles, but which are constant for the lifetime of a given vehicle, such as perturbations of

- the flywheel inertia
- and the inertias and ratios in the gearbox.

SLOW Perturbations that occur over the lifetime of the powertrain or powertrain components, often due to wear, such as perturbation of

- the friction between the clutch plates
- the dynamics of the clutch actuator
- and the performance of the engine.

DISCRETE Perturbations that occur between each vehicle journey or during the first part of the journey, but which remain constant or very similar for the remainder of the journey, such as perturbations of

- the vehicle mass
- and the performance of the engine say through temperature change.

FAST Perturbations that may occur sufficiently rapidly that from one clutch engagement to another, such perturbations appear to be uncorrelated, such as

- the gradient of the terrain
- the road condition

- and the vehicle status, perturbations such as the speed of the vehicle and how sharply the vehicle is cornering at the time of clutch engagement.

For the examples of model perturbations in each group, simulations of the powertrain model are conducted for realistic perturbations, using both the open loop control strategy and the feedback control strategy to control clutch engagement. In these simulations, an engagement from rest case applied to the dataset representing vehicle A with only the clutch torque capacity active as a control, is taken, again as it is felt that this is the worst case. The other control, the engine torque is again set to a constant value of $100Nm$. The feedback control strategy used is as described in the previous chapter, with the modification preventing large control variations around the point of clutch lock up implemented by setting the weighting parameter $w_{\Delta u}$ to 0.01. The other modification, tracking the closed loop flywheel speed to its calculated open loop solution, is not implemented as the problem of stalling is not felt to be significant for this case. This is due to the high engine torque, which is set to its maximum value of $100Nm$. For each model parameter perturbation, a figure containing two graphs is detailed, both of the clutch plate speeds, one of the open loop control strategy used to control engagement and the other of the feedback control strategy used to control engagement.

For the STATIC perturbations, simulations of perturbations in the flywheel inertia, the inertia of the gearbox shaft connected to the clutch and perturbations in the gear ratio are conducted, which are detailed in figures 5.1 - 5.3. For the flywheel inertia, its value has been perturbed from $0.109kgm^2$ to $0.119kgm^2$. From figure 5.1, it can be seen that this perturbation, of the order of 10 percent, fails to produce any deterioration in clutch performance. Similarly, for a perturbation of the gearbox inertia from $0.05kgm^2$ to $0.055kgm^2$, shown in figure 5.2, little

if any deterioration occurs. For these two model parameters a perturbation of 10 percent is felt to be excessive, indicating their unimportance in the control of the clutch. The final STATIC perturbation simulated, is a perturbation of the first gear ratio from a value of 0.3 to 0.35. From the simulation, detailed in figure 5.3, it can be observed how the perturbation causes the open loop control performance to deteriorate, with the clutch plate speed failing to converge quickly enough and oscillations resulting after engagement. For this perturbation, the feedback strategy copes very well, bringing the clutch plates speeds together more quickly and preventing any significant oscillations after engagement. With again the perturbation felt to be large for a change in first gear ratio, this suggests that in this instance the feedback strategy is adequate.

For the SLOW perturbations, simulations of a worn clutch, a clutch with a slower actuator and of a powertrain with an engine producing less torque than expected, are conducted, with the simulations detailed in figures 5.4 - 5.6. The first simulation of a worn clutch, is simulated as before by multiplying the clutch torque capacity by a given value μ, chosen to be 0.9, a reduction of 10 percent in the friction between the clutch plates. From figure 5.4, the marked deterioration in the clutch performance with the open loop controls can be observed, with the clutch plates failing to converge quickly enough resulting in very large oscillations when clutch lock up is finally achieved. The feedback control strategy again does quite well, eliminating most of the oscillations after engagement. However, some small oscillations after engagement still remain, a deterioration of the calculated optimal open loop clutch performance. The simulation of the slower clutch actuator, figure 5.5, is conducted by reducing the natural frequency of the critically damped actuator from $10 rads^{-1}$ to $8 rads^{-1}$. Again a deterioration in the clutch performance results with the open loop controls, which the feedback control corrects, in this case with only very small oscillations after engagement. The final

simulation of a powertrain with an engine producing less torque than required, is modelled by multiplying the engine torque control by a value of 0.9 to represent a 10 percent reduction. Referring back to the engine models discused in chapter 2, this is to say that for a given flywheel speed, the throttle angle that should produce the required torque, calculated from the modelled engine map, in practice produces a torque 10 percent lower than expected. In other words, the actual engine map is 10 percent lower than the modelled map over the entire range of throttle angles and flywheel speeds. From figure 5.6, this perturbation causes the flywheel speed to drop faster than expected with the open loop controls, causing premature clutch engagement, resulting in oscillations after engagement. The feedback strategy tries to correct this by increasing the engine torque and decreasing the clutch torque capacity, resulting in a slowing down of the convergence of the clutch plates, most noticeable between 0.6 and 0.8 seconds. Unfortunately, the feedback control over corrects the problem, with the clutch plate failing to engage at a time of 0.8 seconds, engaging later after 1.0 second has elapsed. Between these two times, the clutch plates speeds move further apart again before coming together, with large oscillations after engagement resulting. For this perturbation, the feedback control performance is worse than the open loop control performance.

For the DISCRETE perturbations, the simulation of clutch engagement with a fully laden vehicle is conducted, with the simulation of engine variations omitted as it has already been covered. As before, a fully laden vehicle is represented by increasing the inertia representing the vehicle mass from $7.945 kgm^2$ to $9.945 kgm^2$. From figure 5.7, it can be seen that such variations cause some deterioration in clutch performance with the open loop controls, which the feedback strategy corrects satisfactorily, with the elimination of significant oscillations after engagement. Again, as there is a clear limit to the magnitude of this perturbation, it

can be stated that the feedback control strategy is satisfactory for perturbations in vehicle mass.

Finally for the FAST variations, three simulations are conducted, clutch engagement up and down 1 : 10 gradients, and of clutch engagement with reduced damping in the compliance, the dynamics between the gearbox output shaft and the vehicle mass. The first simulation detailed in figure 5.8, of engagement up a 1 : 10 gradient with the open loop control strategy employed, demonstrates how the failure of the clutch plate attached to the gearbox to reach the other flywheel speed, cause clutch lock up to be delayed and oscillations after engagement . For the simulation in figure 5.9, of engagement down a 1 : 10 gradient, a similar deterioration in clutch performance occurs with the open loop controls. This time, the speed of the clutch plate attached to the gearbox increases too quickly, causing oscillations when clutch lock up occurs prematurely. For both of these simulations, the feedback controls improve clutch engagement, resulting in satisfactory clutch performance. The final simulation of FAST parameter perturbations, in figure 5.10, of reduced damping in the powertrain compliance, shows how the reduced damping cause larger oscillations before clutch lock up. This reduced damping is modelled by reducing the model damping coefficient from $7.3Nmsrad^{-1}$ to $3.3Nmsrad^{-1}$. This last perturbation, represents changes in the road condition and/or vehicle status. With reference to a typical tyre modelling, linearised about a given operating condition, it can be observed that the damping in the tyres vary with vehicle speed, road condition and factors such as how tightly the vehicle is cornering at the time of clutch engagement. Despite the larger oscillations, the general dynamics of the clutch plates remain unaltered with engagement being completed in a similar time to the calculated optimal value, and little in any oscillations generated after engagement. The feedback controls produce nearly identical dynamics, failing to improve on the open loop

controls. The reason for the lack of improvement is attributed to the difficulty in preventing oscillations just after starting to engage the clutch. With the original damping coefficient these oscillations are noticeable but not significant, but with the reduced damping they become more prevalent. It is felt that the open loop solution does as well as is possible for the engagement problem, with the oscillations only able to be reduced by significantly increasing the time taken to complete engagement. This apparent independence of the controls on the damping coefficient, helps justify the approximation used in the modelling of damping in the compliance, previously modelled by a linear damper.

To conclude, the feedback strategy maintains 'good' clutch performance under most realistic model parameter perturbations. However, 'poor' clutch performance still results from the use of the feedback control strategy under some model perturbations. In particular, for engine perturbations and to a certain extent for changes in the friction between the clutch plates. This motivates the development of a revised feedback control strategy which maintains clutch performance in these extreme cases.

5.2 Clutch engagement with known parameter perturbations

In the previous section, it was noted that the performance of clutch engagement can deteriorate with the feedback clutch control strategy. It may be the case that such model parameter perturbations are known. In the case of STATIC perturbations, perturbations might be known from measurements at manufaction, or for other perturbations measured directly via a sensor, or even estimated from measurements of existing sensors. For example, sensors exist that are capable of measuring the vehicle mass and the gradient or tilt of the vehicle. Such perturbation information might be useful in modifying the feedback control strategy to

cope with parameter perturbations, motivating the development of such a control strategy.

The feedback control strategy arises by constructing a variational optimal control problem, equations (4.24) - (4.29). In particular, by taking variations of the states and controls. If in addition to taking variations of the states and controls, variations are also taken of the model parameters then a variational optimal control problem of the form

$$\min\left[\int_0^1 \begin{pmatrix}\delta\underline{x}\\ \Delta\underline{u}\\ \Delta T\\ \Delta\underline{z}\end{pmatrix}^T \left\{ P_a \begin{pmatrix}\delta\underline{x}\\ \Delta\underline{u}\\ \Delta T\\ \Delta\underline{z}\end{pmatrix} + \begin{pmatrix}T^2\underline{L}_x^T\\ T^2\underline{L}_u^T\\ 2T\underline{L}^T\\ T^2\underline{L}_z^T\end{pmatrix}_* \delta\underline{\beta} \right\} ds + \begin{pmatrix}\delta\underline{x}_f\\ \Delta\underline{u}_f\\ \Delta\underline{z}\end{pmatrix}^T \left\{ P_b \begin{pmatrix}\delta\underline{x}_f\\ \Delta\underline{u}_f\\ \Delta\underline{z}\end{pmatrix} + \begin{pmatrix}\underline{\psi}_x^T\\ \underline{\psi}_u^T\\ \underline{\psi}_z^T\end{pmatrix}_* \delta\underline{\beta} \right\}\right] \tag{5.1}$$

$$\text{subject to } \delta\dot{\underline{x}} = T^2 f_x \delta\underline{x} + T^2 f_u \Delta\underline{u} + 2Tf\delta T + T^2 f_z \Delta\underline{z}, \delta\underline{x}(0) = \delta\underline{x}_0 \tag{5.2}$$

$$\text{where } \delta\underline{\beta} = W\delta\underline{\lambda} + \underline{w} \tag{5.3}$$

$$H = \underline{w}^T\underline{L} + \underline{p}^T f \tag{5.4}$$

$$P_a = \frac{1}{2}\begin{pmatrix} T^2H_{xx} & T^2H_{xu} & 2TH_x^T & T^2H_{xz}\\ T^2H_{ux} & T^2H_{uu} & 2TH_u^T & T^2H_{uz}\\ 2TH_x & 2TH_u & 2H & 2TH_z\\ T^2H_{zx} & T^2H_{zu} & 2TH_z^T & T^2H_{zz}\end{pmatrix}_* \tag{5.5}$$

$$P_b = \frac{1}{2}\begin{pmatrix} f_{txx} & f_{txu} & f_{txz}\\ f_{tux} & f_{tuu} & f_{tuz}\\ f_{tzx} & f_{tzu} & f_{tzz}\end{pmatrix}_* \tag{5.6}$$

results. In this new LQR type problem, as in the previous problem: $\underline{u}$ refers to a vector of the controls in the powertrain model; subscript $*$ denotes the evaluation at u^* with the corresponding states x^* and costates p^* calculated as usual; $\underline{w}$ and W are the performance measure weights and constraint determination matrix

relating the weighted sum performance measure and the constraints to the vectors $\underline{L}$ and $\underline{\psi}$; and $\delta\underline{\lambda}$ is a vector of the Lagrange multipliers variations, with the new vector $\Delta\underline{z}$ denoting a vector of known model parameter variations. Again, this variational optimal control problem can be made solvable by the addition of the term in equation (4.30). The solution of this variational optimal control problem can be obtained by solving a time variant matrix Riccati equation, this time obtained by making the assumption that the optimal variational costates satisfy $\delta\underline{p} = K\delta\underline{x} + C\delta\underline{\beta} + D\Delta\underline{z}$, which can be justified with a minor modification to the argument given in [10]. This assumption, results in the time variant matrix Riccati equation

$$\left[\dot{K}\,\dot{C}\,\dot{D}\right] = \left[-K\,I_n\right]\left[M_d - M_c M_a^{-1} M_b\right]\begin{bmatrix} I_n & 0 & 0 \\ K & C & D \\ 0 & I_l & 0 \\ 0 & 0 & I_p \end{bmatrix} \tag{5.7}$$

where

$$\left[K\,C\,D\right]_{s=1} = \left[f_{txx}\underline{\psi}_x^T f_{txz}\right] - f_{txu}(f_{tuu} + V_t)^{-1}\left[f_{tux}\underline{\psi}_u^T f_{tuz}\right] \tag{5.8}$$

$$M_a(s) = \begin{bmatrix} T^2 H_{uu} & 2T H_u^T \\ 2T H_u & 2H \end{bmatrix} + V_0 \tag{5.9}$$

$$M_b(s) = \begin{bmatrix} T^2 H_{ux} & T^2 f_u^T & T^2 \underline{L}_u^T & T^2 H_{uz} \\ 2T H_x & 2T f^T & 2T \underline{L}^T & 2T H_z \end{bmatrix} \tag{5.10}$$

$$M_c(s) = \begin{bmatrix} T^2 f_u & 2T f \\ -T^2 H_{xu} & -2T H_x^T \end{bmatrix} \tag{5.11}$$

$$M_d(s) = \begin{bmatrix} T^2 f_x & 0 & 0 & T^2 f_z \\ -T^2 H_{xx} & -T^2 f_x^T & -T^2 \underline{L}_x^T & -T^2 H_{xz} \end{bmatrix}. \tag{5.12}$$

where, n is the number of states in the powertrain model, l is the number of elements in the column vectors $\underline{L}$ and $\underline{\psi}$, and p is the number of model parameters

for which variations of the optimal control problem are taken. These differential equations are of the same form as for the feedback control strategy, which when solved backwards in time evaluates the matrices $K(s)$, $C(s)$ and $D(s)$. These matrices then allow the optimal control variations to be calculated in terms of the current state variations, the Lagrange multiplier variations and the parameter variations with

$$\begin{pmatrix} \Delta \underline{u}(s) \\ \delta T(s) \end{pmatrix} = -M_a^{-1} M_b \begin{bmatrix} I_n & 0 & 0 \\ K & C & D \\ 0 & I_l & 0 \\ 0 & 0 & I_p \end{bmatrix} \begin{pmatrix} \delta x(s) \\ \delta \underline{\beta} \\ \Delta \underline{z} \end{pmatrix} \tag{5.13}$$

$$\Delta u|_f = -(f_{tuu} + V_t)^{-1} \left[f_{tux} \underline{\psi}_u^T f_{tuz} \right] \begin{pmatrix} \delta \underline{x}_f \\ \delta \underline{\beta} \\ \Delta \underline{z} \end{pmatrix} . \tag{5.14}$$

As before, all that is required is to evaluate the Lagrange multiplier variations, as it is assumed that variations of the states and parameters are known, and that the current value of the independent variable can be evaluated using (4.54). This evaluation of the Lagrange multiplier variations is accomplished is a similar way to the method employed with the feedback strategy, allowing the value to change with time. As before, this evaluation is based on estimating the constraint variations in terms of known variations and variations of the Lagrange multipliers. This time additional variations of the constraints and state equations are taken with respect to the parameters. Furthermore, equations (5.13) and (5.14) contain variations of the parameters which when substituted into variations of the constraints and state equations adds the terms $\int_i^1 N_7(s)\Delta \underline{z} ds + N_8 \Delta \underline{z}$ and $N_9(s)\Delta \underline{z}$ to equations (4.40) and (4.41) respectively. Again, the constant terms are represented by an additional Lagrange multiplier, which will later be set to unity. The state variations at any s can be evaluated by simulating over a suitable basis, as

with the evaluation of equation (4.42), but this time with the additional dependence on the parameter variations. This new expression can be used to evaluate estimates of the constraints in terms of the current state perturbations, the current Lagrange multipliers, the parameter variations and the known variations of constraint integrands up until the current time, an expression of the form

$$\begin{pmatrix} \widetilde{\delta \underline{R}} \\ \widetilde{\delta \underline{Q}} \end{pmatrix} = M_1(\hat{s})\delta \underline{x}(\hat{s}) + M_2(\hat{s})\delta \underline{\lambda}(\hat{s}) + M_3(\hat{s})\Delta \underline{z} + \begin{pmatrix} \Delta \underline{R}_0(\hat{s}) \\ \Delta \underline{Q}_0(\hat{s}) \end{pmatrix}. \tag{5.15}$$

In this expression, the matrices $M_i(\hat{s})$ are evaluated using a similar procedure to the procedure described in the previous chapter. As before, 2^q cases are taken to evaluate 2^q Lagrange multiplier sets calculated using the slightly modified equation

$$\delta \underline{\lambda}_i(\hat{s}) = -\Lambda_i \left[\Lambda_i^T M_2(\hat{s}) \Lambda_i\right]^+ \Lambda_i^T \left\{ \begin{pmatrix} \Delta \underline{R}_0(\hat{s}) \\ \Delta \underline{Q}_0(\hat{s}) \end{pmatrix} + M_1(\hat{s})\delta \underline{x}(\hat{s}) + M_3(\hat{s})\Delta \underline{z} + \text{const} \right\} \tag{5.16}$$

which when substituted into equations (5.13) and (5.14) yields the result

$$\begin{pmatrix} \Delta u(\hat{s}) \\ \Delta T(\hat{s}) \end{pmatrix} = A_i(\hat{s})\delta x(\hat{s}) + B_i(\hat{s}) \begin{pmatrix} \Delta \underline{R}_0(\hat{s}) \\ \Delta \underline{Q}_0(\hat{s}) \end{pmatrix} + c_i(\hat{s}) + D_i(\hat{s})\Delta \underline{z} \tag{5.17}$$

$$\Delta u_f = A_{fi}\delta x_f + B_{fi} \begin{pmatrix} \Delta \underline{R}_0 \\ \Delta \underline{Q}_0 \end{pmatrix}_{s=1} + c_{fi} + D_{fi}\Delta \underline{z}. \tag{5.18}$$

In this final expression, $\hat{s}$ is the current independent variable and i denotes the Lagrange multiplier set, with all the other terms as previously defined. Again the choice of Lagrange multiplier set is performed by substituting the Lagrange multipliers into equation (5.15), and a similar equation estimating the performance measure variation, with the set chosen to be the set minimising the performance measure variation which also satisfy the estimated constraints.

As this new control strategy is just an extension to the feedback approach, both of the modifications discused in section 4.4.2 can still be included. In particular, the reduction of large control variations just prior to clutch lock up can

be performed by pre-multiplying the control variations calculated using equations (5.17) and (5.18) by the same terms that the control variations in equations (4.52) and (4.53) have been multiplied by in equations (4.57) and (4.58).

In practice the evaluation of the matrices in equations (5.17) and (5.18) is achieved using exactly the same programs as in the solution of the feedback problem, but with the required modifications made to the coding. In carrying out this computation, the model parameters are normalised by the non-perturbed parameter values. The resulting control architect is identical to the architecture show in figure 4.3, but with an additional known input, the parameter variations. This new control architecture is shown in figure 5.11. Note that, setting $\Delta \underline{z}$ to zero in this new control strategy, is equivalent to using the original feedback control strategy.

In order to demonstrate the ability of this new control strategy to cope with parameter perturbations, simulations of this new control strategy are performed. For these simulations, the matrices required in the evaluation of the control variations and the choice of Lagrange multiplier set are calculated with the entries of the diagonal matrices V_0 and V_t set to 10^4. The modification preventing large control variations just prior to clutch lock up is also implemented by setting $w_{\Delta u}$ to 0.01. All the other parameters, such as normalisation constants are as used in the previous chapters. Again, as with the feedback simulations, the computer package ACSL is used to carry out these simulations, simulating the perturbed model equations, as well as calculating the control variations. As before, this last calculation only requires the on going evaluation of variations of the constraint integrals prior to the current time, and the multiplication of matrices. These matrices are the matrices calculated prior to the simulation and are thus read from disk.

The two parameter perturbations presented are the parameter perturbations

studied in section 5.1, which are less than satisfactory with the feedback control strategy. In particular, perturbations of the engine map and to a lesser extent perturbations of the clutch plate friction produce unsatisfactory clutch engagement. In both cases, graphs of the clutch plate speeds, using the open loop control, using the feedback strategy and using the new feedback strategy utilising the known parameter perturbations, are shown in figures 5.12 and 5.13. From figure 5.12, the improvement in clutch performance for the perturbed engine map is dramatic, with much smoother engagement resulting with the new approach than with either the open loop controls or the feedback strategy. In fact, the engagement is seemingly as good as the open loop optimal solution. Figure 5.13, also demonstrates an improvement when the parameter variations are used to improve the feedback control strategy. In this case, controlling clutch engagement with a worn clutch, the oscillations present after engagement with this new approach are noticeably less than with either the open loop control or the feedback control strategy. This improvement is not as dramatic as with the engine perturbation, but it is still felt to be worthwhile. Other small differences between feedback control, with and without making use of the parameter variations have also been identified. In particular, making use of the parameter variations tends to reduce the size of the control variations, especially when the modification, preventing large control variations just prior to clutch lock up is not employed. In other words, it seems that the use of the parameter variations, makes the feedback control problem easier.

5.3 Estimation of parameter perturbations

In the previous section, it was demonstrated that knowledge of parameter perturbations, can be used to aid the control of clutch engagement. Unfortunately, such parameter perturbations might not be known, or might not be measurable, with no sensor available to measure a particular parameter. Even if such sensors

do exist, the implementation of the sensors may be expensive, it might be preferable to estimate the parameter variations from existing measurements. Existing theory does exist, which describe how this estimation might be conducted. For instance, extended Kalman filtering might be applicable, where the states and parameters are simultaneously estimated, performed by treating the parameters as additional states. This approach is well described in [71]. However, as an alternative approach, a different technique is presented, a technique which concentrates on estimating parameter variations which describe the perturbations of the model dynamics. As will be discused, these parameter variations might not be the real variations, but for the purpose of controlling clutch engagement this is unimportant.

5.3.1 Parameter estimation during clutch engagement

The aim of this new approach is to estimate the state variations, in terms of known variables and variations of the model parameters. This relationship can then be used to choose the parameter variations which predict the actual state variations in some optimal way. Such an approach is applicable over the entire operating range of the powertrain. However, as previously noted additional dynamics are present when the clutch plates are slipping, which allow parameters to be estimated that could not be estimated by examining the powertrain dynamics when the clutch is locked up. In particular, no estimation of clutch wear could be achieved. For this reason it is proposed than the estimation of the state dynamics is performed on an interval in time, when the clutch plates are slipping. In particular, during clutch engagement, with the engagement being controlled using the feedback control strategy.

As before, first variations of the state equations can be taken with respect to

variations of the states, controls and model parameters by

$$\delta\dot{\underline{x}} = f_x\delta\underline{x} + f_u\delta\underline{u} + f_z\delta\underline{z}. \tag{5.19}$$

As before, $\delta\underline{x}$ and $\delta\underline{u}$ are first order variations of the states and controls, with $\delta\underline{z}$ the first order variation of the actual parameter variations, which is not the same as $\Delta\underline{z}$, the known parameter variations used in the previous section. For generality, the control variations are calculated using an initial estimate of the parameter variations, $\Delta\underline{z}_0$. This allows $\delta\underline{u}$ to be eliminated in equation (5.19), by substituting in equation (5.13). This in turn allows the estimate of the state equations

$$\delta\underline{x}(s) = \Phi(s)\delta\underline{x}_0 + \Psi(s)\begin{pmatrix}\delta\underline{\lambda}_0\\ \Delta\underline{z}_0\\ 1\end{pmatrix} + \Omega(s)\delta\underline{z}. \tag{5.20}$$

to be obtained, acquired by simulating the variations of the state equation over a suitable basis, calculated for all s in the interval used for the parameter estimation. In this last equation, $\delta\underline{x}_0$ and $\delta\underline{\lambda}_0$, are the state and Lagrange multiplier variations at the start of the estimation interval. The problem left is to calculate the Lagrange multiplier variations, enabling the state variations to be estimated in terms of variations of model parameters, and known terms. The difficulty in evaluating the Lagrange multiplier variations is that the feedback control procedure allows them alter in time, both in value and choice of Lagrange multiplier set. This Lagrange multiplier alteration, allows the feedback control strategy to adapt to unmodelled dynamics and external influences. However, if the perturbations in the model parameters adequately portray these effects, then the Lagrange multiplier set and value should remain constant or at least only vary slightly, when the true parameter perturbation is used to calculate the control variations. This of course will not be the case when an incorrect parameter estimation is used, as with poor initial estimations. However, it is still felt that with the assump-

tion that the Lagrange multipliers remain constant, equation (5.20) is still a good approximation of the state variations. The difficulty with the variability of the Lagrange multiplier variations, can be helped by making the assumption that the choice of Lagrange multiplier set remains constant. This not only removes the computationally expensive choice of the Lagrange multiplier set but also helps simplify the algebra. This simplification is not unjustified, as from simulations of the feedback strategy it can be observed that for most engagement problems the bounds of all of the inequality constraints are reached over the entire interval $[0,1]$ in s. Even when this is not the case, the restriction that the inequality constraints achieve their bounds, results in almost identical control variations. Hence, for this parameter estimation, the Lagrange multiplier set corresponding to non-zero Lagrange multipliers for all constraints is chosen. This is equivalent to setting the inequality constraints whose bounds are achieved by the open loop solution to equality constraints in the following feedback calculations.

With this last assumption, the required estimation of the state variations can be found by substituting equation (5.16), with $\hat{s}$ corresponding to the start of the estimation interval, into equation (5.20), resulting in the relationship

$$\delta \underline{x}(s) = \Phi'(s)\delta \underline{x}_0 + \Psi'(s) \begin{pmatrix} \delta \underline{R}_0 \\ \delta \underline{Q}_0 \\ \Delta \underline{z}_0 \\ 1 \end{pmatrix} + \Omega(s)\delta \underline{z}. \tag{5.21}$$

In this equation, the terms $\delta \underline{R}_0$ and $\delta \underline{Q}_0$ refer to the integral of the constraint integrands up until the start of the estimation interval, and are assumed to be known. Now that the state variations are estimated, parameter variations are chosen so that the estimated variations remain as close to the actual state variations, as close as is possible. Closeness can be measured using the standard norm

of continuous functions on a fixed interval, measured by

$$\int_{t_1}^{t_2} ||\delta\underline{x}(t) - \underline{x}(t) + \underline{x}^*(t)||^2 \, dt = \int_{s_1}^{s_2} T^2 \, ||\delta\underline{x}(s) - \underline{x}(s) + \underline{x}^*(s)||^2 \, ds. \qquad (5.22)$$

The two forms, question how the interval on which the estimation is conducted is defined. As the calculations are performed with s as the independent variable, s_1 and s_2 are used for this definition. This also enables an interval to be chosen for which the clutch plates are slipping, by choosing $[s_1, s_2] \subset [0, 1]$. Substituting equation (5.21) into equation (5.22), differentiating with respect to $\delta\underline{z}$, and setting this derivative to zero, yields the equality condition

$$M\delta\underline{z} + N \begin{pmatrix} \delta\underline{x}_0 \\ \delta R_0 \\ \delta Q_0 \\ \Delta_{\underline{z}_0} \\ 1 \end{pmatrix} = \int_{s_1}^{s_2} T^2 \Omega^T (\underline{x} - \underline{x}^*) ds \qquad (5.23)$$

$$\text{where } M = \int_{s_1}^{s_2} T^2 \Omega^T \Omega ds \qquad (5.24)$$

$$N = \int_{s_1}^{s_2} T^2 \Omega^T \left[\Phi' \Psi' \right] ds \qquad (5.25)$$

that must be satisfied for $\delta\underline{z}$ to be the parameter variations that minimise the distance of the estimated variations from the actual state variations. The difficulty in using this equality condition to calculate an estimation of the parameter variations is that the matrix M might be singular. Fortunately this problem can be overcome using the Moores-Penrose pseudo inverse, resulting in

$$\delta\underline{z} = M^+ \left\{ \int_{s_1}^{s_2} T^2 \Omega^T (\underline{x} - \underline{x}^*) ds - N \begin{pmatrix} \delta\underline{x}_0 \\ \delta R_0 \\ \delta Q_0 \\ \Delta_{\underline{z}_0} \\ 1 \end{pmatrix} \right\} \qquad (5.26)$$

The use of the Moores-Penrose pseudo-inverse, means that the parameter estimations calculated might not be the real parameter perturbations when M is singular. However, in this case, the effect of the calculated perturbation and the real perturbation on the states are indistinguishable, and there use in correcting

the control variations is just a valid. For the case when M is non-singular, then the estimation should be an explicit estimation of the parameter variations.

It is questionable whether differences in all of the states is necessarily the best measure to use to achieve the parameter estimates. For instance, differences between both clutch plate speeds only without including the differences between the other states, or differences between real system measurements, might be preferable. For this reason, the approach is generalised slightly, by examining the differences between the estimated and actual variations of

$$\underline{y} = g(\underline{x}), \tag{5.27}$$

where g is a differentiable function. With this generalisation, estimates of first variation of $\underline{y}$ can be calculated, by pre-multiplying estimates of the state variations in equation (5.21) by g_x, the derivative of g with respect to the states. An expression describing the closeness of the estimated $\underline{y}$, to the actual $\underline{y}$, can be obtained in a similar way, which when differentiated with respect to the parameter variations, allows the new estimate

$$\delta \underline{z} = M^{+} \left\{ \int_{s_1}^{s_2} T^2 \Omega^T g_x^T (\underline{y} - g(\underline{x}^*)) ds - N \begin{pmatrix} \delta \underline{x}_0 \\ \delta R_0 \\ \delta Q_0 \\ \Delta_{\underline{z}_0} \\ 1 \end{pmatrix} \right\} \tag{5.28}$$

$$\text{where } M = \int_{s_1}^{s_2} T^2 \Omega^T g_x^T g_x \Omega ds \tag{5.29}$$

$$N = \int_{s_1}^{s_2} T^2 \Omega^T g_x^T g_{\underline{x}} \left[\Phi' \Psi' \right] ds \tag{5.30}$$

of the parameter variations to be obtained, where the actual value, $\underline{y}$, is assumed to be known. Note that, in both parameter estimations, the terms M and N can be calculated off-line, only being dependent on the predefined interval $[s_1, s_2]$, with the integral in equations (5.26) and (5.28), being calculated on-line.

As a summary of the approach, a geometric interpretation of this estimation procedure is now given. This is done by representing the states, controls and

model parameters as three vertices, describing co-ordinates in a 3-D space. In order to maintain this representation as simple as is possible, the constraints and the initial parameter estimates used in the calculation of the feedback controls, are ignored. With this representation, the calculation of the parameter estimates precedes as follows, with a diagram of the representation shown in figure 5.14.

In this 3-D space, the state equations are represent by a surface or manifold. The calculation of the open loop optimal solution locates a point $(\underline{u}^*, \underline{x}^*, \underline{z}^*)$ on this manifold, the point which minimises the performance measure, with $\underline{z}^*$ the model parameters used in this calculation. The feedback control strategy, describes a linear relationship between variations in the states and variations in the controls, and is thus represented by a hyperplane, passing through the open loop optimal solution. The intersection, of the manifold describing the state equations and the hyperplane describing the feedback relationship, describes the set of all possible combinations of states, controls and model parameters, which can occur when the feedback control strategy is implemented. In reality, this intersection is a manifold of the same dimensions as the number of model parameters allowed to vary. Furthermore, as the feedback strategy is independent of the model parameter perturbations, any point in this manifold can be uniquely described by the model parameters. Hence, if the real model parameters are $\underline{z}'$, then a unique point $(\underline{u}', \underline{x}', \underline{z}')$, on this manifold, is determined. The parameter estimation procedure, approximates this manifold of all possible points, by its tangent space at the open loop optimal point $(\underline{u}^*, \underline{x}^*, \underline{z}^*)$. The new point $(\underline{u}^{est}, \underline{x}^{est}, \underline{z}^{est})$ on this tangent space can then be located, the point closest to $(\underline{u}', \underline{x}', \underline{z}')$, in some pre-defined way. The parameter estimate can now be read off as $\underline{z}^{est}$.

5.3.2 Simulations of model parameter estimation

In order to demonstrate the effectiveness of this estimation procedure, simulations of clutch engagement are conducted, with parameter perturbations applied and with the parameter estimation calculations performed. In particular, the simulations of engagement with a worn engine, a worn clutch, a fully laden vehicle and of a hill-start, are conducted. These are four of the simulations used in section 5.1 to motivate the parameter adaptation of the feedback strategy, and are repeated, this time with the integral in equation (5.26) calculated. In order to evaluate this integral, and the matrices M and N, the interval $[s_1, s_2]$ must be chosen. For these simulations, the interval $[0, 0.9]$, was chosen. The last part of the interval for which the clutch is slipping is not used, as in this region, larger control variations are more likely and the assumption that the Lagrange multiplier remain constant is probably incorrect. Indeed, is has been found that using this region, can result is poor parameter estimates. The feedback control used, is as described in section 5.2, with initial parameter estimates being used to adapt the controls. The restriction that the inequality constraints remain on their bounds is also applied. In all other aspects the simulations are exactly as in section 5.1. For each simulation, the parameter estimation procedure is acquired using all of the states, as in equation (5.26).

For each parameter perturbation, three simulations are conducted. The first of feedback control, with the controls adjusted using no initial parameter variation, the second of feedback control using the correct parameter variation, and the final simulation of feedback control using the parameters estimated in the first simulation. Graphs of the clutch plate speeds acquired from these simulations along with graphs of the clutch plate speeds when the open loop controls are used, are detailed in figures 5.15 and 5.16, for the two perturbations for which unadapted feedback control is unsatisfactory. For all three simulations, and for

all four perturbations, estimates of the parameter variation are obtained. In this estimation it is assumed that the parameter which has been altered is the only parameter that can vary. In other words, the perturbed parameter, is the only parameter estimated. These parameter estimates are detailed in table 5.1.

From the table, it can be seen that the estimation procedure produces good estimates for all of the parameter perturbations. In particular, the drop in engine torque is estimated within 9% of its true value, the clutch wear within 16% of its true value, the vehicle mass within with 16% of its true mass, and the gradient of the terrain within 3% of its true value. The discrepancies occur mainly as a result of the non-linearity of the states and model parameters in the state equations. Note that, the parameter expressing the gradient of the terrain is the only parameter with a linear relationship in the state equations. This is the probable reason for its accurate estimation. It can also be seen that the effect of using initial parameter estimates in calculating the controls, on the calculated estimates is negligible. Figures 5.15 and 5.16, show how these estimates can be used to correct any problems with the engagement dynamics, with the engagement using the estimated parameters as good as engagement using the actual parameter variations. In fact, there are no noticeable differences between these engagements for the two problematic perturbations.

Unfortunately, it is unlikely that just one parameter will vary. For this reason a further three simulations are conducted, of engagement with a worn engine, engagement with a worn clutch, and of engagement with a worn engine with a fully laden vehicle, this time with four parameters estimated. In particular, with the engine and clutch wear parameters estimated, along with the parameters measuring the vehicle mass and the gradient of the terrain. Again for each perturbation, three simulations are conducted, as with the single parameter estimation, with the estimated values for the perturbations detailed in tables 5.2 - 5.4. From these

tables, it can be seen that the estimation of the engine perturbation, is achieved, just as successfully as when one parameter is estimated. The estimations of the other parameters, not perturbed, also remain small with the their values less the 10% of typical variations of these parameters. The estimation of the clutch wear is not as quite as accurate, as is shown in table 5.3, getting within 20% of the true values and with larger estimates of the non-perturbed parameters being found. However, the estimates are reasonable. For the perturbation of engine torque and of vehicle mass, again reasonable but not good estimates are achieved, with the estimation of vehicle mass almost 40% out for the first estimation. For all three perturbations, engagement using the estimated parameters located with the first simulation, is just as good as with engagement using the actual parameter perturbations, as shown in figures 5.17 - 5.19.

As a final investigation, simulations of clutch engagement with a worn engine and with the initial engine flywheel speed lowered to $175 rad/s$, are conducted, in order to demonstrate parameter estimations under state perturbations. For this case, the parameter estimation is poorer, not in estimating the engine wear, but in estimating the other parameters, as shown in table 5.5. In particular, the vehicle mass is estimated 10% lighter than it is, with the gradient of the terrain estimated at a 1:40 up hill. The difficulty with this estimation, is that the effect of a heavier vehicle and of an up hill gradient on the state dynamics are very similar. This means that in essence the lighter vehicle estimate cancels out with the up hill gradient estimate in the adaptation of the feedback controls. Hence, the use of these estimates in controlling clutch engagement, is still valid, with just as good engagement resulting, as shown in figure 5.20.

Other simulations, using just the clutch plate speeds to estimate the parameters have also been conducted. Initial results suggested that there was little difference in the ability of the procedure to estimate the parameter variations,

with more work being required to firmly establish this hypotheses. Further work, identifying the minimal state information enabling successful parameter estimation is felt appropriate.

To conclude, the estimation procedure acquires good estimates in most parameter perturbations simulated. Furthermore, even when the estimates are slightly different from the real values, the use of them to modify the feedback controls is just as effective.

5.4 Adaptive/self-learning control

The previous section has shown that good estimates of the model parameters can be found. Using the technique described in the previous section, one estimate can be achieved for each clutch engagement. Furthermore, it has been shown that this parameter estimate can be used to improve clutch engagement. For the simulations, one estimation from a single clutch engagement was sufficient to correct the feedback strategy. In practice, a variety of variations are present, all varying at different rates, making the task more complex. For this reason, a brief outline of techniques for creating adaptive or self-learning control strategies from these estimations is discused.

As mentioned, in reality, variations of states, other parameter variations, and of different external influences occur from one clutch engagement to another. This will cause the estimations of any particular model parameter, to vary with such variations, even if its value is static. However, its true underlying value may also change in time. Hence, the problem is to extract this trend from any sequence of estimated model parameters, in order to obtain accurate estimates of the model parameters. These estimates can then be used to modify clutch engagement. One way of achieving this is to update a running estimate of a model parameter with the estimates acquired from each clutch engagement. Typically this can be done

using to formula

$$z_{new} = (1 - \lambda)z_{old} + \lambda z_{est} \tag{5.31}$$

where $\lambda \in [0, 1]$ is a relaxation parameter and z_{est} is a parameter estimation acquired from a single clutch engagement. In essence, this filters out, high frequency variations, higher than would be expected for the parameter in question. In section 5.1, the model parameters where grouped into classes describing how fast they can vary. This classing can now be used to design values of the relaxation parameters for each model parameter being estimated. For instance, for:

- STATIC model parameters, a small relaxation parameter could be chosen, just small enough to filter out all dynamics present in the powertrain, but large enough that parameter learning is not excessively long,
- SLOW model parameters, a small relaxation parameter could again be chosen, possible slightly larger, but probably very similar in magnitude to the value for STATIC model parameters,
- DISCRETE model parameters, a larger value needs to be chosen during the first part a journey, with the value decreasing for lengthy journeys.

Note that in practise, that the relaxation parameter is not only dependent on the speed of the dynamics but on the intervals between clutch engagements. Also note, that the first order discrete adaptation, equation (5.31), can be generalised to higher order equations, using discrete filter theory, to design a low pass filter.

There is still a difficulty with FAST model parameters. Recall from section 5.1, that these model parameters vary sufficiently fast, that from one clutch engagement to another, their value appears to be uncorrelated. Even if the relaxation parameter is set to 1, equivalent to correcting engagement with the previous estimate, the estimate used may be significantly different from its true value. A possible solution to this is to carry out more than one estimation for ever clutch

engagement. The theory used in calculating these estimates can use any interval during clutch engagement. This allows the value used in correcting clutch engagement to be updated during the engagement, by partitioning the interval $[0, 1]$ in s into several intervals on which parameters estimation is performed. Initial simulations, suggest that this approach might work, although more work is required the establish this approach.

To conclude, it is felt that the parameter adaptation and estimation techniques presented in this chapter, can be used to design adaptive or self-learning clutch engagement controllers.

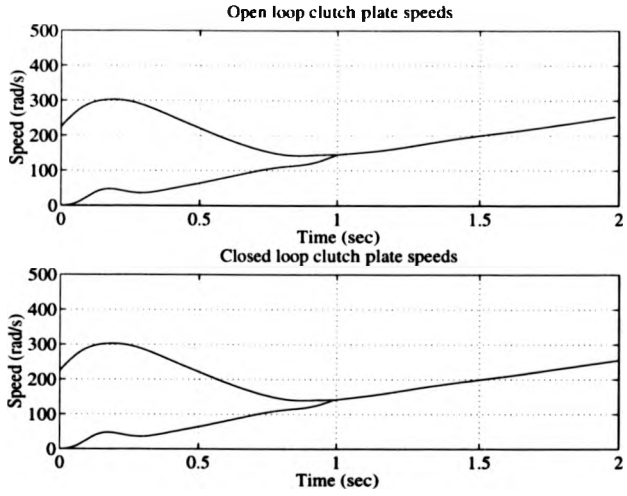


Figure 5.1: Clutch performance of feedback control wwith a heavier flywheel inertia

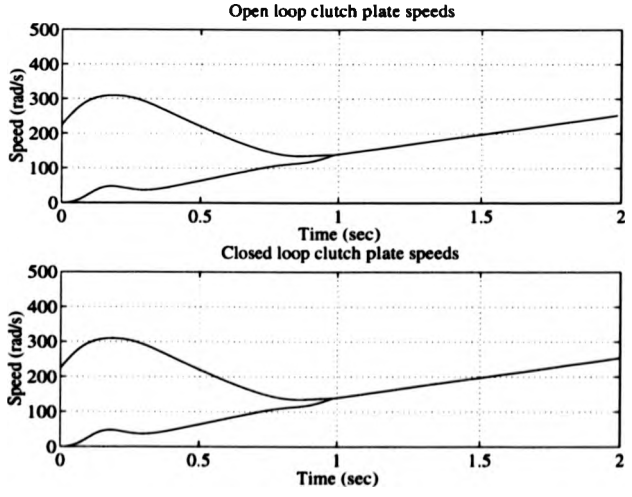


Figure 5.2: Clutch performance of feedback control with a heavier gearbox input shaft

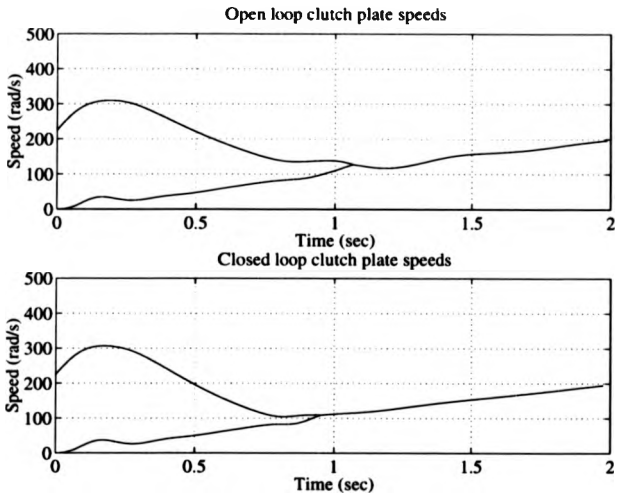


Figure 5.3: Clutch performance of feedback control with a larger gearbox 1st gear ratio

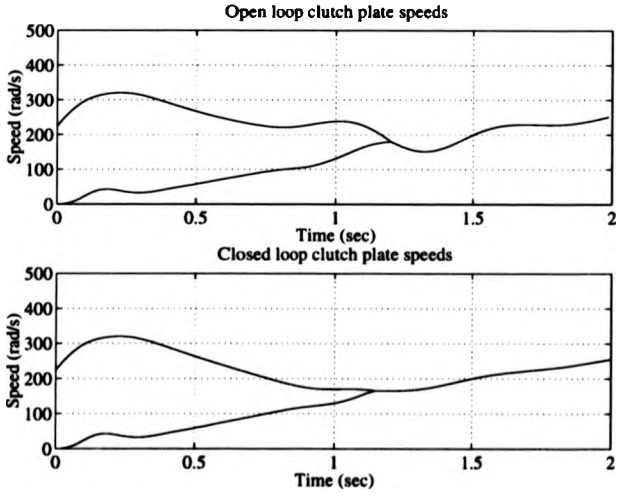


Figure 5.4: Clutch performance of feedback control with reduced clutch plate friction

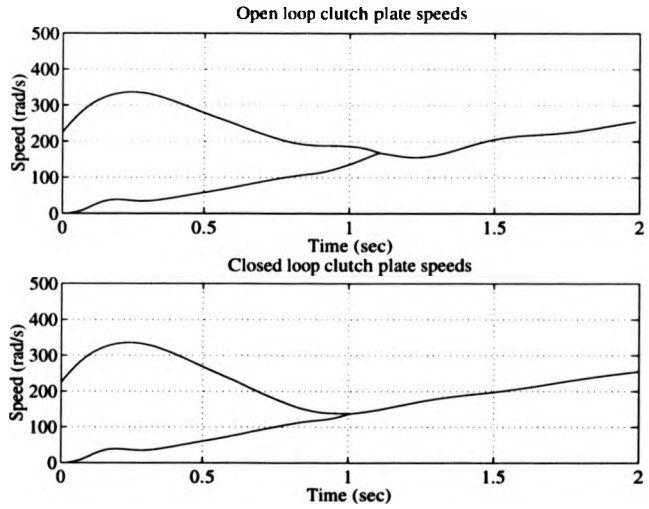


Figure 5.5: Clutch performance of feedback control with a slower actuator

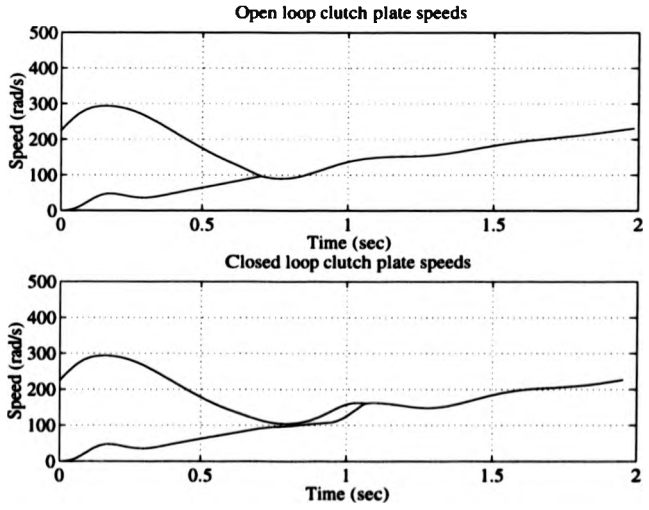


Figure 5.6: Clutch performance of feedback control with an engine producing less torque than expected

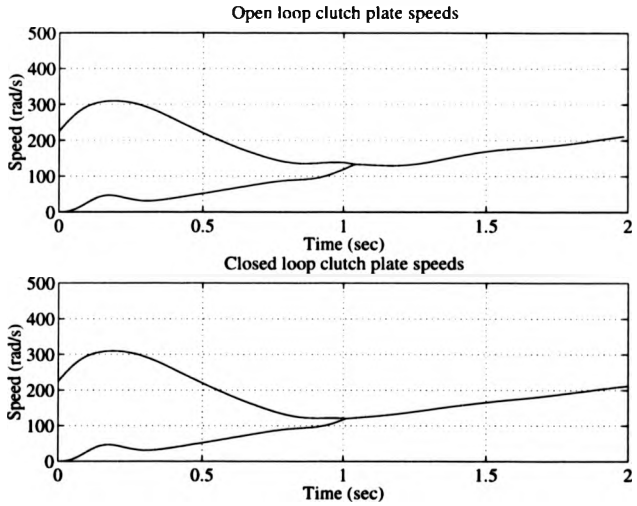


Figure 5.7: Clutch performance of feedback control for a fully laden vehicle

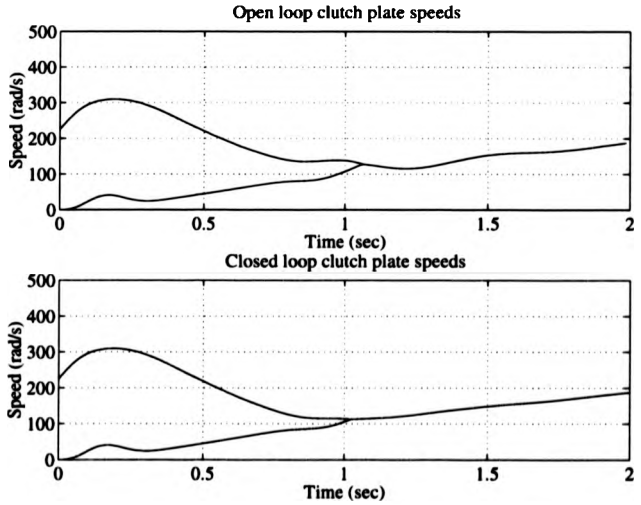


Figure 5.8: Clutch performance of feedback control up a 1 : 10 gradient

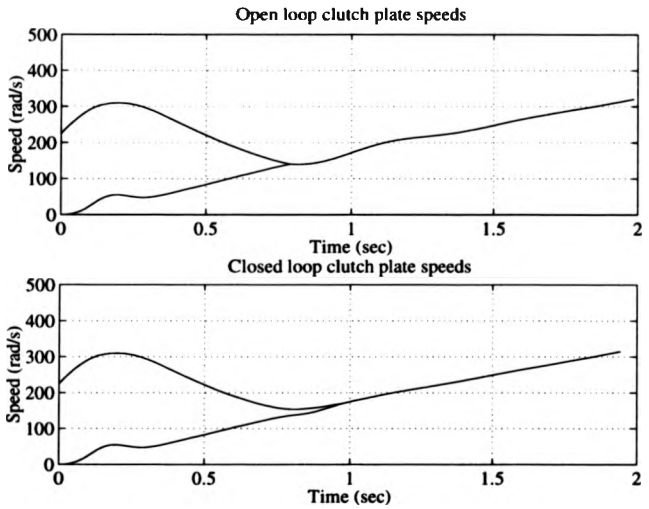


Figure 5.9: Clutch performance of feedback control down a 1 : 10 gradient

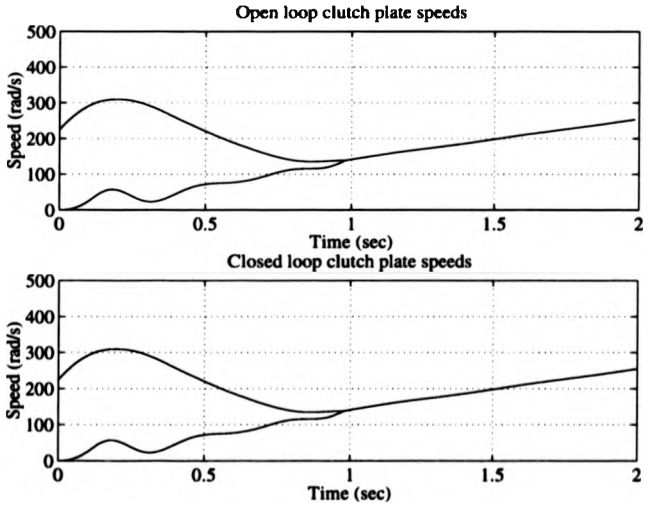


Figure 5.10: Clutch performance of feedback control with less damping

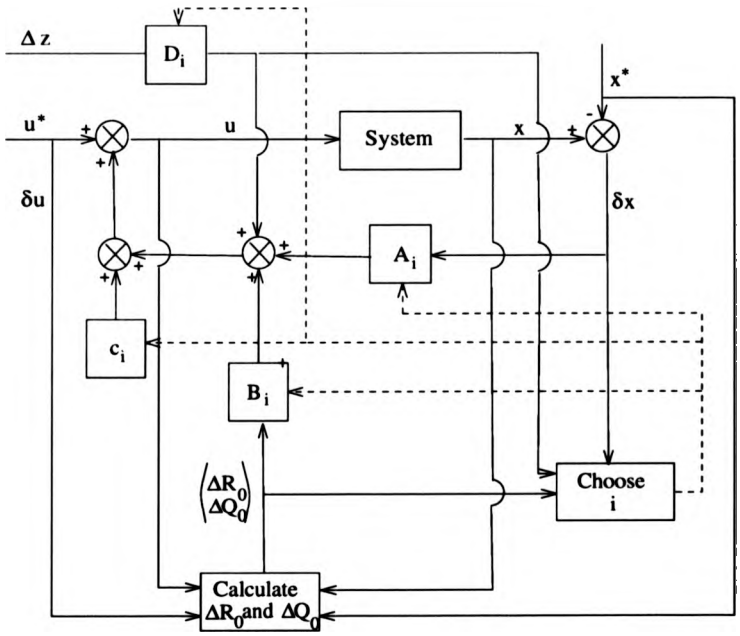


Figure 5.11: Optimal feedback control architecture with known parameter perturbations

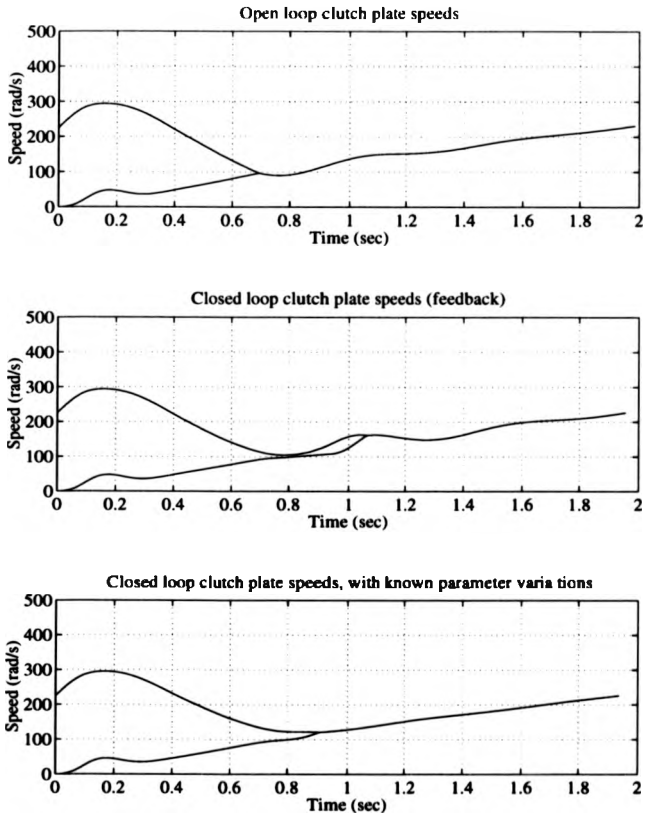


Figure 5.12: Clutch performance with known parameter variations of an engine producing less torque than expected

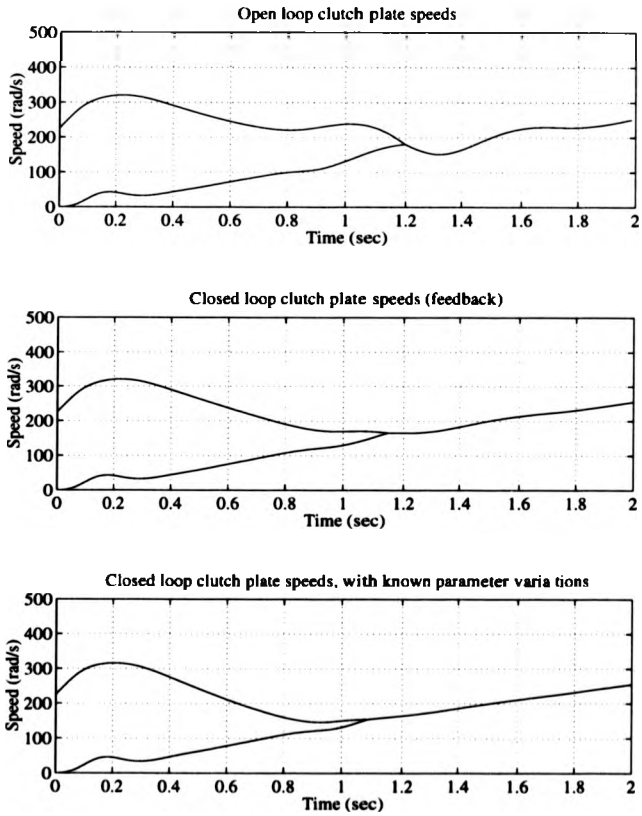


Figure 5.13: Clutch performance with known parameter variations of engagement with reduced clutch plate friction

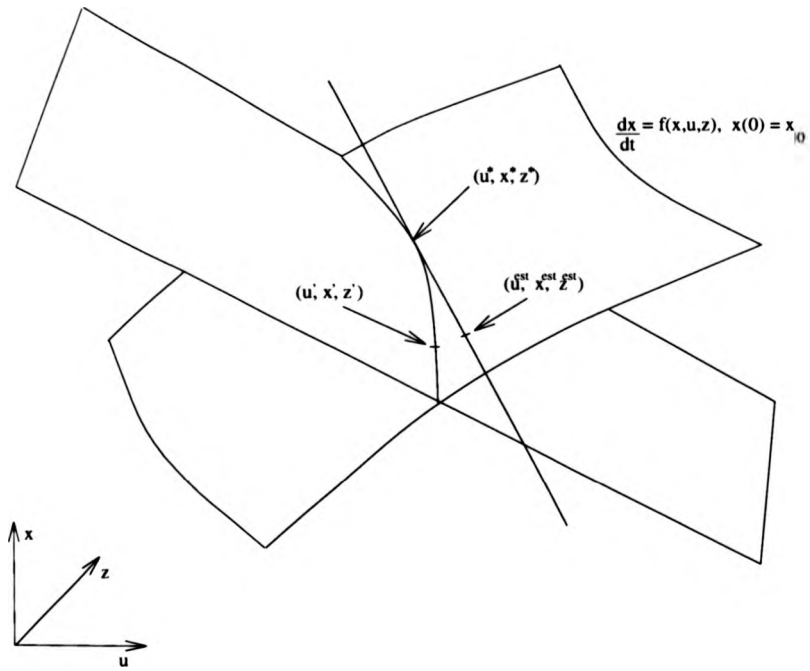


Figure 5.14: Geometric interpretation of model parameter estimation

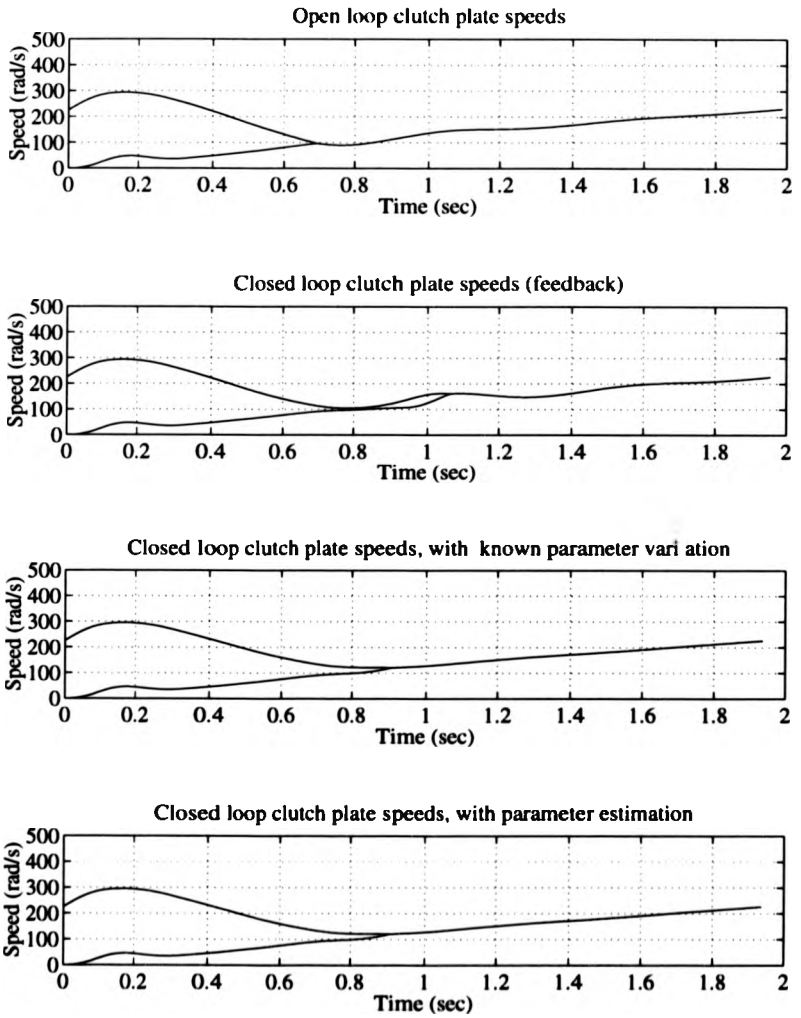


Figure 5.15: Feedback control of an engine producing less torque than expected, using estimated parameter variations

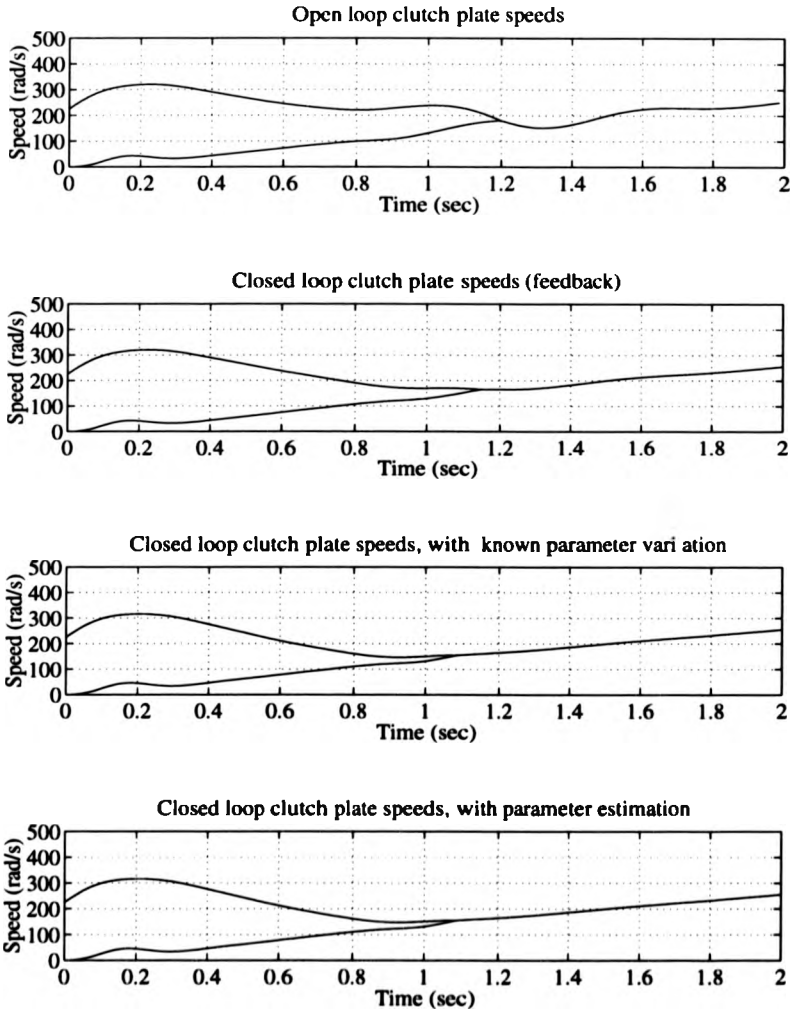


Figure 5.16: Feedback control with a worn clutch, using estimated parameter variations

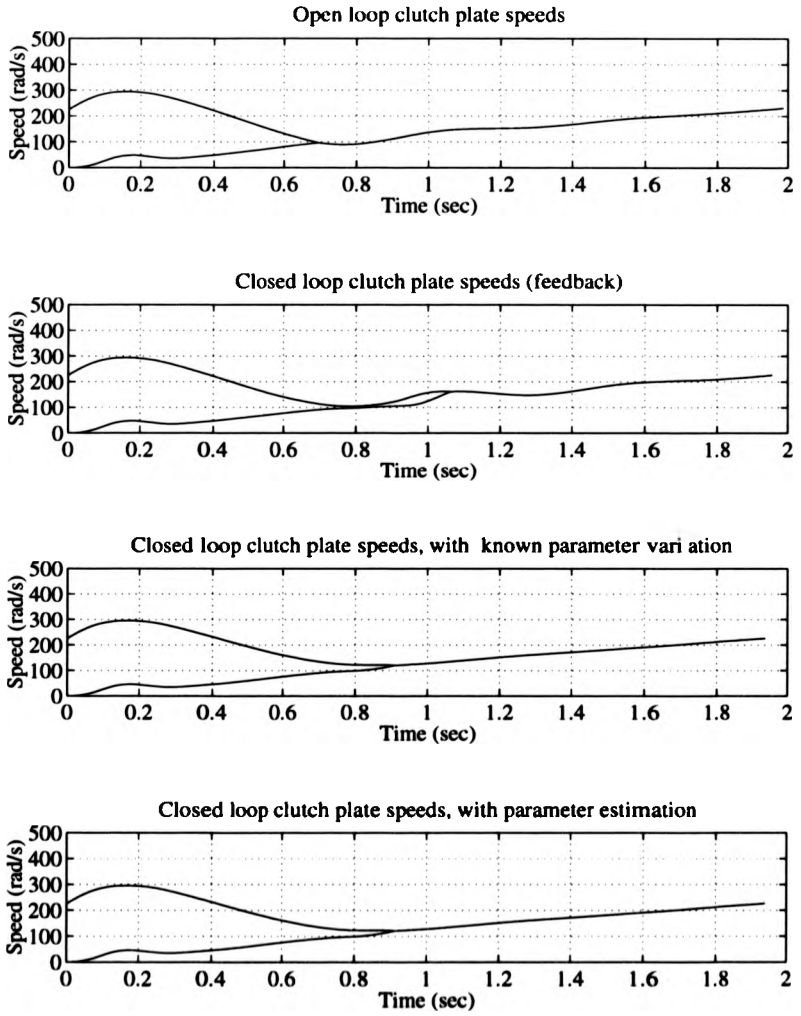


Figure 5.17: Feedback control of an engine producing less torque than expected, using the four parameters estimated

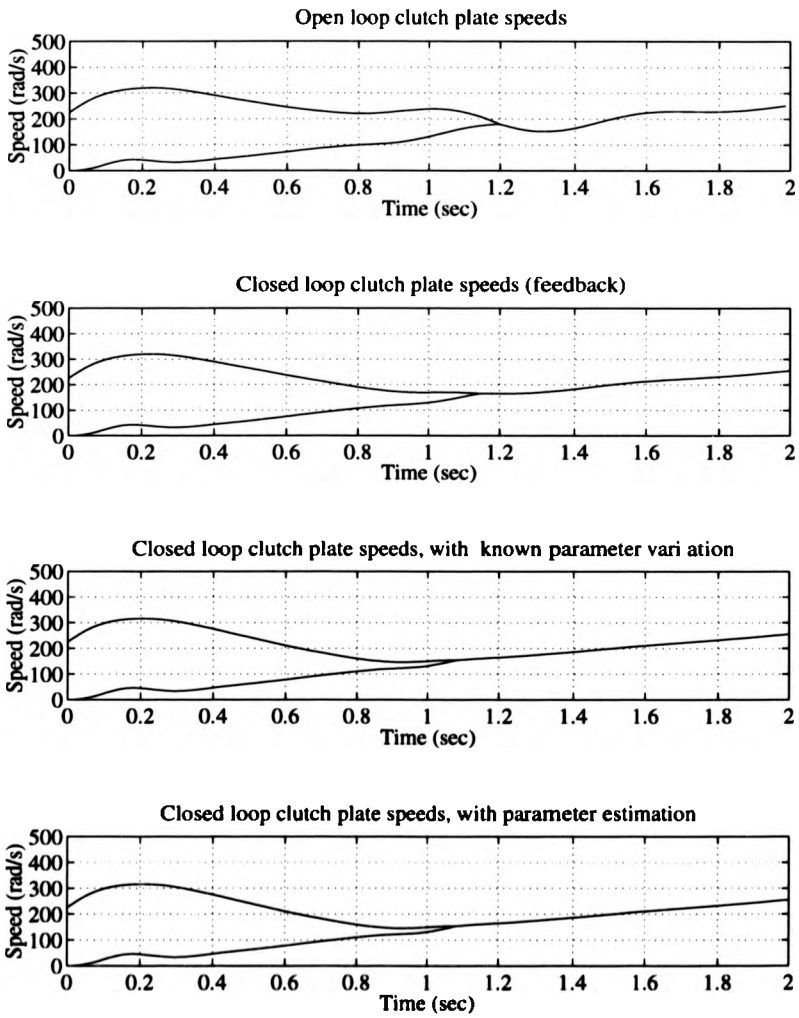


Figure 5.18: Feedback control with a worn clutch, using the four parameters estimated

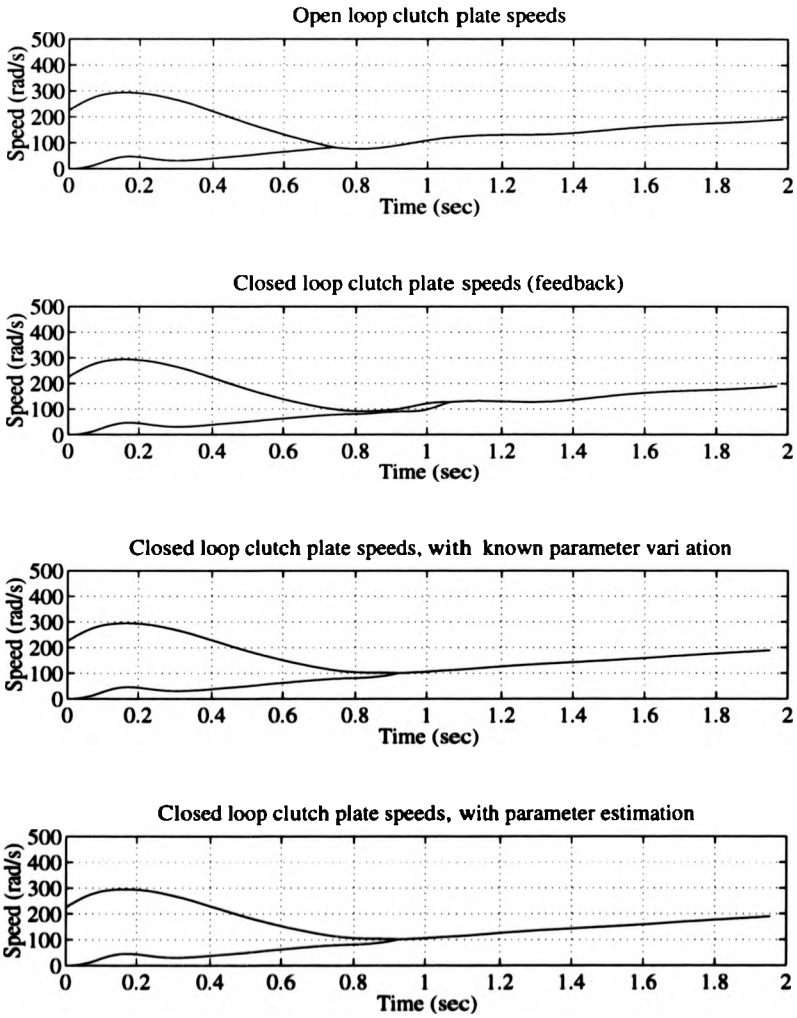


Figure 5.19: Feedback control with a worn engine and a fully laden vehicle, using the four parameters estimated

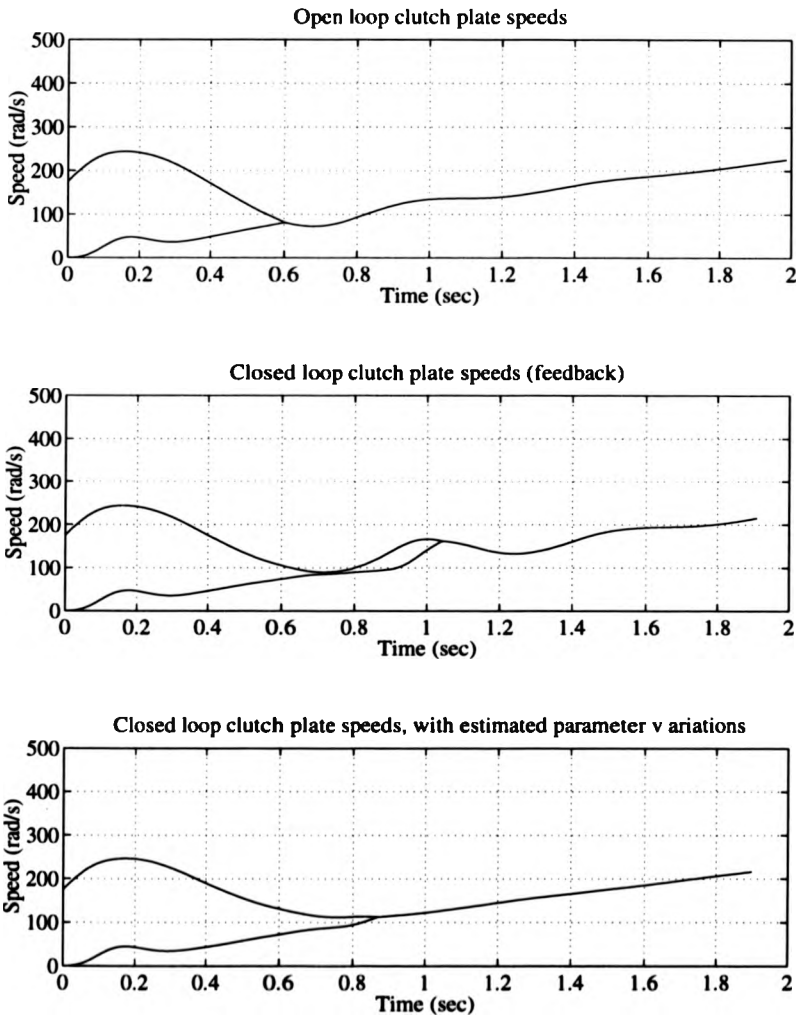


Figure 5.20: Feedback control of an engine producing less torque than expected, using the four parameters estimated and with an initial flywheel perturbation

Variation			Actual variation	Estimated variation using		
Description	Symbol	Units		no initial estimate	actual variation	estimated variation
Worn engine	-	-	-0.1000	-0.0996	-0.1083	-0.1083
Worn clutch	μ	-	-0.1000	-0.0955	-0.1158	-0.1151
Fully laden vehicle	J_4	kgm^2	2.0000	1.6866	1.6845	1.6822
Hill start	a	Nm	100.7190	100.0245	98.4789	98.4604

Table 5.1: Estimates of parameter perturbations, with just one parameter estimated

Description	Worn engine	Worn clutch	Fully laden vehicle	Hill start
Symbol	-	μ	J_4	a
Units	-	-	kgm^2	Nm
True variation	-0.1000	0.0000	0.0000	0.0000
Estimated variations with feedback	-0.1015	0.0006	-0.2599	7.2413
Estimated variations with feedback using true variations	-0.1029	0.0084	0.0996	-4.9124
Estimated variations with feedback using first estimate	-0.1032	0.0086	0.1063	-5.1487

Table 5.2: Estimates of the engine torque perturbation, with four parameters estimated

Description	Worn engine	Worn clutch	Fully laden vehicle	Hill start
Symbol	-	μ	J_4	a
Units	-	-	kgm^2	Nm
True variation	0.0000	-0.1000	0.0000	0.0000
Estimated variations with feedback	0.0176	-0.0847	0.2937	-0.5063
Estimated variations with feedback using true variations	0.0069	-0.1104	-0.1778	8.0496
Estimated variations with feedback using first estimate	0.0063	-0.1106	-0.1890	8.2833

Table 5.3: Estimates of the clutch wear perturbation, with four parameters estimated

Description	Worn engine	Worn clutch	Fully laden vehicle	Hill start
Symbol	-	μ	J_4	a
Units	-	-	kgm^2	Nm
True variation	-0.1000	0.0000	2.0000	0.0000
Estimated variations with feedback	-0.0982	0.0007	1.2517	9.0650
Estimated variations with feedback using true variations	-0.0994	0.0072	1.5756	-3.2837
Estimated variations with feedback using first estimate	-0.1008	0.0068	1.5779	-3.9517

Table 5.4: Estimates of the engine wear and vehicle mass perturbation, with four parameters estimated

Description	Worn engine	Worn clutch	Fully laden vehicle	Hill start
Symbol	-	μ	J_4	a
Units	-	-	kgm^2	Nm
True variation	-0.1000	0.0000	0.0000	0.0000
Estimated variations with feedback	-0.1099	0.0047	-0.7995	23.0561

Table 5.5: Estimates of the engine torque perturbation, with four parameters estimated and with an initial engine flywheel perturbation

Chapter 6

Conclusions

Most of the aims of the thesis have been completed successfully. In particular:

1. a formal mathematical investigation of clutch engagement has been conducted;
2. the notion of good clutch engagement has been quantified;
3. a procedure to calculate control strategies that result in good clutch engagement has been constructed;
4. and feedback control strategies, that it is felt might be needed in the implementation of clutch control, which cope with perturbations of the model states and parameters, have been developed.

The result, is the design of a clutch engagement controller, which under simulation, successfully engages the clutch for a variety of undesirable situations. In this thesis, the worse case problem of engagement from rest, and quite often with just one control active has been concentrated on, to identify, as fully as is possible, the limitations of any control strategy. However, the theory is applicable to any clutch engagement problem, such as clutch engagement during gear change.

It is is felt that there are two uses of this research in the practical implementation of a clutch engagement controller.

The first use of this research, is to use the designed control strategies, as a design aid in designing simple clutch engagement control laws. These controls laws might consist of simple rules, with techniques such as fuzzy logic, describing how to implement such rule based control strategies. For instance, the open loop control calculation can be used to identify characteristics, which any rule based algorithm must ensure, such as the reduction in clutch torque capacity at the point of clutch lock up. The feedback control strategy and model parameter adaptation can also be used to design simple control laws. For instance, the simulations suggest that changes in the flywheel speed are best coped with by varying the engine torque, especially when the likelihood of stalling is high. Simulations, might also enable rules to be developed suggesting how best to cope with parameter changes such as clutch wear.

The second use of this research is to take the designed control strategies forward to direct implementation. The possibility of implementing the control strategies was envisaged in there development and has influenced their design. This direct implementation can be carried out in three progressive stages:

- the implementation of the open loop controls,
- the implementation of the feedback control strategy,
- and the implementation of an adaptive or self-learning clutch controller,

with the requirements of each step now outlined.

As mentioned, good open loop control strategies can be calculated for a variety of different engagement problems. Assuming that each engagement problem can be described by known variables, such as the initial flywheel speed in the case of engagement from rest, then a series of open loop controls can be calculated, describing how best to engage the clutch for each different engagement problem. By varying these variables describing the engagement problem, open loop controls

can be found for a class of clutch engagement problem. Any element of the class can then be described from these calculated open loop controls, possible by using interpolation or by regressing some function through the calculated controls. There is a slight difficulty, due to the non-uniqueness of the solution, with the solution dependent on the initial control strategy used in the algorithm to calculate the open loop controls. This can to a certain extent be overcome by using the same initial control strategy for all of the calculated open loop control strategies. The data required to evaluate any open loop control strategy in an engagement class, whether variables describing a function regressed through the data or the raw data itself, is then required to be stored in some memory on board the vehicle. This then allows the required open loop control strategy to be located from this memory by an automotive controller, if it is assumed that the variables required to describe the engagement problem are known. For instance, in the case of engagement from rest, the initial engine flywheel speed can be used to determine the open loop controls. This approach is also valid for variations in model parameters, with open loop controls being calculated for different values of model parameters. The difficulty with this is that the actual model parameters must be known and the the size of the data required to be stored increases exponentially with the number of variables which the open loop controls are dependent on. Furthermore, the simulations suggest that if incorrect values of the variables describing the engagement problem are used to calculate the open loop controls then poor clutch engagement might result. There is also the problem that the simplistic model used in the calculation of the open loop controls, might fail to fully represent the dynamics of clutch engagement.

The implementation of the feedback control strategy is similar. Again several open loop control strategies can be calculated, describing a class of engagement problems. However, this time, due to the ability of the feedback algorithm to

cope with certain perturbations, the number of open loop controls calculated for each class of engagement problem can be dramatically reduced. It even might be possible to successfully describe a class of engagement problems by just one typical engagement problem. For each set of open loop controls, the required feedback matrices can then be calculated, with both the open loops controls and feedback matrices for any given engagement problem, again determined by techniques such as interpolation or regression. Again data describing the feedback matrices and open loop controls needs to be stored in memory. Again, variables describing the problem, if present, are required, along with full model state information. Techniques such as Kalman filtering exist for calculating such state information from system measurements. However, these techniques do require the system to be observable. In the case of clutch engagement, as the dynamics either side of the powertrain are decoupled whilst the clutch is slipping, this means that at least two measurements of the powertrain dynamics from either side of the clutch will be required. The memory requirement are also much higher for the feedback control strategy. If n_T is the number of time points used to store the controls in s, n is the number of states, q is the number constraints and m is the number of controls, then the storage requirement of a single open loop control is $m \times n_T$ floating point numbers, whilst the storage of the feedback matrices is $m \times n_T \times (n+q+1)$ floating point numbers, assuming that their is no choice of Lagrange multipliers made. If the cases of different Lagrange multiplier sets are considered, then this increases by a further multiple of 2^{q_1}, where there are q_1 inequality constraints, along with the addition of matrices estimating the constraint variations and performance measure variations.

The implementation, of an adaptive or self-learning controller, requires the implementation of the feedback control strategy described in section 5.2, along with an algorithm for estimating the parameters variations, possibly as described

in section 5.4. The implementation of the feedback control is as previously described with the an increase in size of the stored matrices, due to additional terms relating to alterations due to parameter variations. In particular, the storage requirements of each feedback matrix increases to $m \times n_T \times (n + q + 1 + p)$ floating point numbers, where p is the number of model parameters adapted for.

On reflection, the use of this research as a design aid is more appropriate. The reasoning behind this is that in any clutch engagement implementation, the conceived benefits must out weigh the costs of the implementation. With the direct implementation, requirements such as full state estimation and large memory units might make the implementation too expensive. However, in the future, if more control devises are implemented on automotive vehicles, these difficulties may reduce, with components becoming cheaper and with the required powertrain states being calculated for other control systems.

There is one other important use of this research. The optimisation approach used to design clutch enagement controllers, is a general non-linear technique applicable to other other control problems. In applying this theory to the clutch engagement problem a number of technical difficulties have been solved, such as problems with the free terminal time in the optimal control problem and the calculation of the feedback control strategy with variable Lagrange multipliers. Other problems might have similar, if not identical problems, solutions to which might be provided by this research.

Bibliography

[1] Simpson R J. A history of electric control. In *Proceedings of the second European Control Conference*, volume 3, pages 1246–1250, 1993.

[2] Bennett S. The pid controller: the history of an idea. In *Proceedings of the second European Control Conference*, volume 3, pages 1235–1239, 1993.

[3] Franklin G F, Powell J D, and Emami-Naeini A. *Feedback Control of Dynamical Systems: second edition.* Addison-Wesley, 1991.

[4] Nyquist H. Regenation theory. *Bell Sys Tech J*, 11:126–147, 1932.

[5] Evans W R. Graphical analysis of control systems. *Trans AIEE*, 67:547–551, 1948.

[6] Lyapunov A M. *On the General Problem of the Stability of Motion.* PhD thesis, Khartov Mathematical Society, 1892.

[7] Pontryagin L S and et al. *The Mathematical Theory of Optimal Processes.* Wiley, 1962.

[8] Bellman R. *Dynamic Programming.* Princeton University Press, 1957.

[9] Kalman R E. A new approach to linear filtering and prediction problems. *Trans ASME J Basic Eng*, 82:34–45, 1960.

[10] Athans M and Falb P L. *Optimal Control.* Mcgraw-Hill, 1966.

[11] K. J. Åstrom and B. Wittenmark. *Adaptive Control.* Addison-Wesley, 1989.

[12] J. A. Farrell and W. L. Baker. An introduction to learning control systems - workshop notes. In *The 1993 IEEE International Symposium on Intelligent Control, Chicago, USA.*, 1993.

[13] Tsypkin Y Z. *Adaptation and Learning in Automatic Systems.* Academic Press: Mathematics in Science and Engineering Volume 73, 1971.

[14] Antsaklis P J, Passino K M, and Wang S J. An introduction to autonomous control systems. *IEEE Control Systems Magazine*, June:5–13, 1991.

[15] C. C. Lee. Fuzzy logic in control systems: Fuzzy logic controller - Parts I & II. *IEEE Transactions on Systems, Man and Cybernetics.*, 20(2):404–453, 1990.

[16] Farrall S D and Jones R P. Energy management in an automotive electric/heat engine hybrid powertrain using fuzzy decision making. In *Proceedings of the 8th IEEE International Symposium on Intelligent Control*, pages 463–468, 1993.

[17] Antsaklis P J. Neural networks in control systems. *IEEE Control Systems Magazine*, April:3–5, 1990.

[18] Antsaklis P J. Neural networks in control systems. *IEEE Control Systems Magazine*, April:8–10, 1992.

[19] Efstathiou J. Expert systems, fuzzy sets and rule based control explained at last. *Transactions on Instrumentation Measurement and Control*, 10(4):198–206, 1988.

[20] Kharitonov V L. Asymtotic stability of an equilibrium position of a family of linear differential equations. *Differential Uravren*, 14:2086–2088, 1978.

[21] Goodall D P and Ryan E P. Feedback controlled differential inclusions and stabilization of uncertain dynamical systems. *Siam J. of Control and Optimization*, 26(6):1431–1441, 1988.

[22] Doyle J C, Francis B A, and Tannenbaum A R. *Feedback Control Theory.* MacMillan, 1992.

[23] Francis B A. *A course in H_∞ control theory.* Springer-Verlag, Lecture Notes in Control and Information Sciences Vol 88, 1987.

[24] Doyle J C, Glover K, Khargonekar P P, and Francis B A. State-space solutions to standard H_2 and H_∞ control problems. *IEEE Transactions on Automatic Control*, 34(8):831–847, 1989.

[25] Mcfarlane D C and Glover K. *Robust Controller Design using Normalised Coprime Factor Plant Descriptions.* Spriger-Verlag, Lecture Notes in Control and Information Sciences Vol 138, 1990.

[26] Packard A and Doyle J. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[27] Dorato P, Tempo R, and Muscato G. Bibliography on robust control. *Automatica*, 29(1):201–213, 1993.

[28] Dorato P. *Robust Control.* IEEE press, 1987.

[29] Robert Bosch GmbH. *Automotive Handbook.* Robert Bosch, 1986.

[30] A. S. Cherry. *An Investigation of Multibody System Modelling and Control Analysis Techniques for the Development of Advanced Suspension Systems in Passenger Cars.* PhD thesis, University of Warick., 1992.

[31] Sano S. Furukawa Y, and Shiraishi S. Four wheel steering with rear wheel steer angle controlled as a function of steering wheel angle. *SAE 860625*, 1986.

[32] Ackermann J, Sienel W, and Steinhauser R. Robust automatic steering of a bus. In *Proceedings of the second European Control Conference*, volume 3, pages 1534–1539, 1993.

[33] D. Hrovat and W. F. Powers. Powertain computer control systems. In *Selected papers from the 10th Triennial World Congress of the International Ferderationn of Automatic Control, Munich, Germany.*, volume 3, pages 213–219. Pergammon Press, 27-31st July 1987.

[34] Main J J. New developments in powertrain control. In *Fifth International Conference on Automotive Electronics*, pages 135–141, 1985.

[35] U. Kiencke. A view of automotive control systems. *IEEE Control Systems Magazine*, Vol. 8(No. 4):pp 11 – 19, August 1988.

[36] Dobner A R. Optimal control solution of the automotive emission - constrained minimum fuel problem. *Automatica*, 17(3):441–458, 1981.

[37] Patil P B. Computer controlled shifting of an automatic transmission. In *Fifth International Conference on Automotive Electronics*, pages 111–120, 1985.

[38] Y. Hojo, K. Iwatsuki, H. Oba, and K. Ishikawa. Toyota five-speed automotic transmission with application of modern control theory. *SAE 920610*, 1992.

[39] Garbett K. S. *Multi-Objective Scheduling and Control of a Nonlinear Automotive Powertrain.* PhD thesis, University of Warwick, 1991.

[40] Murakami K and Maeda M. Automobile speed control system using fuzzy logic controller. In M. Sugeno, editor, *Industrial Applications of Fuzzy Control.*, pages 105–123. Elsevier Science Publishers B. V., 1985.

[41] A. N. Costa and R. P. Jones. Motion management for automotive vehicles. In *Proceedings of the IEE International Conference Control '91, Edinburgh, UK*, volume Vol 2., pages pp 932 – 937, 1991.

[42] Tanaka H and Ishihara T. Electronically controlled fully automatic transmission for commercial vehicles. In *Proceedings of the XVIIth FISITA Congress*, page 1463, 1978.

[43] Tsangarides M C and Tobler W E. Dynamic behavior of a torque converter with centrifugal clutch. *SAE 850461*, 1985.

[44] Suga M, Niikura Y, Murasugi T, Saitoh K, and Takase S. The control of the lockup clutch used in the microprocessor controlled automatic transmission. In *Fifth International Conference on Automotive Electronics*, pages 107–110, 1985.

[45] Millward J P. Automatic control of heavy goods vehicles dry plate clutches. In *Fifth International Conference on Automotive Electronics*, pages 183–188, 1985.

[46] Shimazu T, Yanagihara S, and Tanaka H. Directly control pneumatic clutch for heavy duty vehicles. In *Proceedings of the 1986 American Control Conference*, volume 1, pages 252–257, 1986.

[47] Cho D and Hedrick J K. Sliding mode control of automotive powertrains with uncertain actuators. In *Proceedings of the 1989 American Control Conference*, volume 2, pages 1053–1058, 1989.

[48] Aussedat F and Maurel G. Innovative concept for high-speed I/O processor dedicated to engine control system. In *Proceedings of the IMechE Seventh International Conference: Automotive Electronics*, pages 229–235, 1989.

[49] Intel Corporation. *Intel Automotive Handbook.*

[50] Multiplexing standards - far from standardization, December 1990.

[51] Phail F H and Arnett D J. In-vehicle networking - serial communication requirements and directions. *SAE 860390*, 1986.

[52] Jones R. P. Powertrain dynamics and control. In Smith J. R., editor, *Mathematics in the Automotive Industry*, pages 63–79. Oxford University Press, 1992.

[53] Donald J. Dobner. Dynamic engine models for control development - part 1: Non-linear and linear model formulation. In *Technological Advances in Vehicle Design: Application of Control Theory in the Automotive Industry.* Int J of Vehicle Design, Special Publication SP4, 1983.

[54] Baruah P. C. A simulation model for transient operation of spark-ignition engines. *SAE 900682*, 1990.

[55] A. M. Foss, R. P. G. Heath, P. Heyworth, J. A. Cook, and J. McLean. Thermodynamic simulation of a turbocharged spark ignition engine for electronic control development. In *Proceedings of Institute of Mechanical Engineers Seventh International Conference. Automotive Electronics. London, UK. 9-13 Oct. 1989.*, pages 195–202, 1989.

[56] Powell J.D. A review of ic engine models for control system design. In *Selected papers from the 10th Triennial World Congress of the International*

Federationn of Automatic Control, Munich, Germany., volume 3, pages 233–240. Pergammon Press, 27-31st July 1987.

[57] R. P. Jones and others. Modelling and simulation of an automotive powertrain incorporating a perbury continuously variable transmission. In *Institute of Mechanical Engineers Conference on Integrated Engine/Transmission Systems.*, pages 41 – 49, Bath, U. K., 1986.

[58] Ford research and engineering centre, May 1991. Private communication with P. R. Crossley.

[59] Szadkowski A and Morford B. Clutch engagement simulation: Engagement without throttle. In *Transmission and Driveline Symposium: Components, Gears and CAE (SP-905)*, pages 103–117, 1992.

[60] Tanaka H. and Ishihara T. Electro-pneumatically operated proportional clutch for semi-automatic power transmissions. *JSAE Review*, Vol. 5:31–36, 1981.

[61] Yu P. L. Cone convexity, cone extreme points, and non dominated solutions in decision problems with multiobjectives. *Journal of Optimization Theory and Applications*, Vol. 14:319–377, 1974.

[62] Matthews J C and Jones R P. Optimisation based strategies for the control of clutch engagement in an automotive vehicle. *I.M.A. International Conference on Control: Modelling, Computation, Information*, 1992. Manchester, 2nd-4th September.

[63] Athans M and Falb P L. *Optimal Control.* McGraw-Hill, 1966.

[64] Polak E. *Computational Methods in Optimization.* Academic Press: Mathematics in Science and Engineering Volume 77, 1971.

[65] Sage A P and White C C. *Optimum systems control: second edition.* Prentice-Hall, 1977.

[66] Luenberger D G. *Optimisation by Vector Space Methods.* Wiley, 1969.

[67] J. C. Matthews and R. P. Jones. Optimal feedback control of clutch engagement in an automotive vehicle. In *Proceedings of the second European Control Conference, ECC'93, Groningen, The Netherlands.*, volume 2, pages 986–991, 1993.

[68] Hull D G. On the variational process in optimal control theory. *Journal of Optimization Theory and Applications*, 67(3):447–462, 1990.

[69] Kugelmann B and Pesch H J. New general guidance method in constrained optimal control. *Journal of Optimization Theory and Applications*, 67(3):421–446, 1990.

[70] Golub G H and Van Loan C F. *Matrix Computations.* North Oxford Academic, 1983.

[71] Maybeck Peter S. *Stochastic Models, Estimation, and Control (Volume 2).* Academic Press, Mathematics in Science and Engineering Volume 141-2, 1982.