Multi-Objective Scheduling and Control of a Nonlinear Automotive Powertrain

Multi-Objective Scheduling and Control of a Nonlinear Automotive Powertrain

K. S. Garbett

A Thesis Submitted to the University of Warwick for the degree of Doctor of Philosophy

Department of Engineering

University of Warwick

September 1991

Summary

The automotive industry is faced with the challenge of ever-increasing emission legislation. This study demonstrates the effective use of nonlinear techniques in automotive control for the problem of fuel and emission minimisation. A review of previous work highlights the inadequacy of traditional optimisation formulations. The conflicting requirements of both low fuel and emissions is a design problem for which compromise and trade-offs are unavoidable. This study attacks the problem through powertrain scheduling, an approach ideally suited to both S.I. and diesel engines, and demonstrates how the novel application of multi-objective optimisation methods provides a solution more akin to the real physical problem.

The modern control theory approach presented is a three stage process : formulation of the mathematical model, including the essential dynamics, constraints, and objectives of the physical problem; optimisation of the control strategy with respect to the relevant performance criteria; and synthesis of the optimal control design. The optimisation model is finite-dimensional and nonlinear, the use of which demands a knowledge of nonlinear systems and available methods. These are classified. Results for single and multi-objective optimisations are compared and fully demonstrate the advantages of the latter for the scheduling problem. Optimal schedules are generated and from them, implementable rule-based control laws are derived. Performance, in terms of the ability to track a legislative test cycle and to retain the optimal design specification, is demonstrated through dynamic simulation, as is their driveability and robustness.

This study specifically considers a diesel-engined vehicle incorporating a CVT. The methods are widely applicable however, to other engine and transmissions types, and to other automotive control problems.

Contents

List of Tables

List of Figures

Acknowledgement

There are many people to whom grateful thanks are due. Firstly, I am indebted to my supervisor, Dr. Peter Jones, for his continual encouragement and advice throughout this project, and for his guidance in writing this thesis.

I would like also to thank SERC and Lucas Automotive for the funding of this CASE project. In particular, I would like to thank Dr. P. Extance, for his encouragement, and Dr. P. Scotson and Mr. P. Mason for their technical support and advice, and for provision of essential engine data.

Grateful thanks are due to Prof. P. Flemming, and Dr. A. Grace, previously of the University College of North Wales, for the provision of the MATLAB Optimisation Toolbox, and to Mr. T. Crummey, of University College of North Wales, for his backup help. I hope that my return comments have been useful.

Finally, I would like to give a special thankyou to Mr. D. J. Kerbyson for his love and understanding and to thank him for the many hours of proofreading he has had to endure.

Chapter 1

Introduction

The established methods of control theory have had a pervasive impact on modern technology, [1]. The application of these well-established methods has expanded hand in hand with the growth in computer technology, together with increasing scope for the use of modern control concepts. Motivated by the need for control where the limitations of classical control theory are now recognised, especially within nonlinear dynamic systems, new approaches are being sought within the industry as a whole .

The automotive industry is already a comprehensive user of modern control systems. Originating in the U.S.A. in the early 1970's, automotive control applications started with the introduction of microprocessor based engine control to combat the conflicting demands of high fuel economy and low exhaust emissions, [2]. By the mid 1980's control application had widened to the successful implementation of anti-lock braking systems, (A.B.S.). Based on feedback principles, and using control microprocessors to detect tyre slip, they are now on the verge of becoming a standard feature in today's vehicles. The next generation of car users will wit-

ness a widespread introduction of control microprocessors, into both passenger and commercial vehicles, the resulting features of which are likely to be : integrated engine and transmission control to improve fuel economy and driveability; automatic suspension control to improve both ride and handling; traction control combining powertrain control and anti-slip braking; and variable assist steering systems, giving firm or lose steering dependent on the need of the driver, [3], [4].

The introduction of electronics has also facilitated the growing trend towards the total vehicle control concept - a network of control systems throughout the vehicle, communicating via a Control Area Network (CAN), and possibly under the supervisory control of a central microprocessor, [4]. This progress towards integration of automotive vehicle subsystem controls requires that the interaction between the various subsystems be taken into account at the design stage. To achieve this a systematic approach to vehicle control is needed in order to cope with the wealth of complex interactions between the various subsystems, [2].

The work embodied here concentrates on just one area of vehicle control : powertrain control, which in essence controls the flow of energy along the driveline through the modulation of torque. The powertrain encompasses the engine, transmission and road wheel subsystems, as shown in Figure 1.1. Each unit is extremely complicated in its own right and all have produced intense research areas in recent years. A great deal of the work has been done using linear approximation and steady state models since both nonlinear and dynamic theory is severely restricted and lacks the generality of linear analysis. The application of electronics, however, has considerably improved processing power, allowing the scope of control application to be broadened to both transmission and traction control, [5]. However, it is not the role of this thesis to discuss the control of either traction, transmission or engine

management in their own rights. Instead, powertrain control here is concerned with nonlinear driveline torque management across the three subsystems.

Under the conflicting constraints of user and legislative requirements, the inherent powertrain control system needs first to determine the control inputs required in order to achieve the desired goals. This is the concept of control scheduling, and reflects the need of the control engineer to spend a considerable amount of time in the analysis as opposed to the control design stage. It is the aim of this thesis to investigate the optimum control scheduling rules for a nonlinear powertrain given the conflicting goals of the system.

1.1 Powertrain Control

The function of the powertrain is to provide the necessary forces for locomotion. Produced by the energy released through the combustion of liquid fuels such as petrol and diesel, the resulting mechanical power is translated from a speed of rotation and torque level, output at the engine flywheel, to a tractive effort at the road wheels via the transmission and driveshaft. The transmission usually consists of a clutch, fluid coupling or torque converter to provide a means of decoupling the engine from the drive wheels during a gear change, or when the vehicle is at rest, together with a gearbox.

The gearbox enables the power produced by the engine to be matched to the varying power demands of the vehicle as its speed and acceleration requirements change. Transmissions can be classified into two major groups : discrete ratio and continuously variable. The former, more conventional gearbox, traditionally provides four or five discrete ratios which are manually or automatically shifted. The

second type offers a continuous variation in gear ratio within the upper and lower bounds of the system and are automatically controlled. The inherent fuel economy improvements which in principle result from the use of continuously variable transmissions, (C.V.T.), make it increasingly attractive, [18].

Based on a perceived vehicle speed and road conditions, the driver will make a demand, which is transferred to the engine as a power demand, Figure 1.2. The outputs from the engine subsystem, as seen in the Figure, are engine speed and torque at the flywheel. These are then modulated so that the demanded power requirements are achieved, as well as other system objectives. These include good driveability - the response and ease of handling of the vehicle, fuel economy and low exhaust emissions. On passing through the transmission and road wheels the selected torque must also produce the desired vehicle speed.

The parameters of the modulation process are thus dictated by the user's requirements and increasingly by worldwide legislation. Drivers have always demanded good fuel economy and performance, but wider political issues, such as the oil crisis of the 1970's and erratic prices changes of the 1990's, renew the interest in producing fuel effcient vehicles. Increasingly important however, has become the need to reduce vehicle emissions. Environmental concern, and increasing consciousness of their damaging effect to health, has meant that emission reduction is at the forefront of design issues.

As early as 1953, it was known that hydrocarbon emissions from automobiles were a significant cause of photochemical smog. Led by legislation in California in the 1960's, other emissions such as nitrous oxides (NO_x), carbon monoxide, particulates and smoke, found dangerous to health, have also been restricted by worldwide legislation. Improved fuel and oil development together with engine component de-

sign has gone some way towards improving emissions, [7], as have the developments of lean burn engines, catalytic converters and exhaust gas recirculation, (E.G.R.). Reductions due primarily to engine design can be seen for a diesel engine in Table 1.1. The figures compare two standard commercial vehicle engines, and the values given are the simulated minimum possible emissive output for each function over a prescribed driving cycle, for equivalent vehicle model parameters. The newer data represents a more modern engine incorporating a turbo-charger and E.G.R. . Also shown are the equivalent legislative limits for the respective test. It is still not possible to meet the stringent limits for nitrous oxides on the newer engine even with improved engine design and exhaust gas recirculation.

Any control applied to the problem is very much dependent on the type of fuel used and the subsequent engine's design. Control of petrol fueled or spark ignition, (S.I.), engines is dependent on the driver varied throttle angle, which determines the air intake, the metering of fuel and the adjustment of the spark timing, Figure 1.3a. The total quantity of air and fuel supplied affects the level of torque produced and the ratio of air to fuel (A/F) influences the smooth running, the fuel economy and the exhaust pollutant formation. Figure 1.4 provides a quantitative description of the effect of A/F on fuel economy and exhaust emission levels in a typical S.I. engine, [8]. The trade-offs between the objectives can quite clearly be seen and precise control of A/F ratio is necessary for optimum performance of the engine. For a S.I. engine this can be achieved by running the engine lean or through exhaust gas after-treatment devices, or a combination of both, [9].

The control variables of the S.I. engine are replaced in a diesel engine by the driver controlled fuel metering, the excess air factor controlled by exhaust gas recirculation, and by injection timing, Figure 1.3b. Since a diesel engine must run on excess air,

to avoid soot and hydrocarbon production, and the engine condition range in which it must run for good efficiency is far narrower, [8], engine control is more restricted. The previous methods of lean burn and exhaust gas after-treatment, which operate around stoichiometry, are not appropriate and control is forced away from the engine to the more general control of the powertrain.

The legislation which restricts the emissive output of the vehicle is based on standard driving cycles. Three general configurations exist, the European E.C.E., American F.T.P. and the Japanese cycle. Each attempts to reflect the habitual driving styles of the respective continents. The tests generally cover the whole operating range of the vehicle which is important since some operating conditions are more onerous than others. For instance, for a diesel engine, a greater part of the nitrous oxides are formed in the high load/high speed area, whereas the high speed/light load operations are deemed to enhance emissions of hydrocarbons and lube oil particulates, [6, 7]. Acceleration phases normally cause increased emission of black smoke, whilst motoring phases are thought to be an emissive source of unburnt lubricant, [6]. The minimisation of particulate output is especially challenging for the automotive industry which faces reductions of around 60% of current American levels by 1994 legislation, [6].

Previous attempts at minimising fuel and emission outputs have followed two lines of approach. The first and most predominant, engine calibration, attempts to control the combustion process over each engine cycle to obtain the desired emission and fuel levels. Calibration is a term used in this context to describe the derivation of laws to govern the action of the engine controllers, based on the outputs of fuel, air and spark advance. As such, it is an approach which can only be applied to S.I. engines and theoretical treatments of this and other internal parameter

manipulations have been numerous in recent years. These are summarised fully in Tennant and Blumberg, [11, 12], with modelling approaches summarised in Powell, [14], and discussed more fully in Chapter 2. The treatments involve a mix of on-line and modelling techniques together with the application of control. Traditionally adjusted by mechanical carburettors or distributors which change only in response to engine conditions, digital electronic controllers now allow the optimisation of the running state to be adjusted so that any particular speed and torque output can select any spark advance or air-fuel ratio.

Optimal transmission scheduling, the second approach is available to both S.I. and diesel-engined vehicles, and involves the modulation of torque across the power-train in order to meet the demands of low fuel, low emissions and good driveability. This has traditionally involved treating the drive cycle as a number of steady state operating points, [13] and using inverse kinematic models to drive the vehicle up the 'economy line' – the locus of optimal steady state speed-load points. Analyses have been undertaken in both the discrete gearbox domain, [13, 16, 17] and for that of the continuously variable transmission, C.V.T., [18, 19, 21], although many of the principles are interchangeable. Discrete transmission control can further be classified according to whether gear scheduling and or gear execution is controlled, [2].

Both lines of approach have concentrated on steady state formulations, due until recently to the lack of the necessarily complex models or to the lack of sufficient processing power. Only a few studies have been undertaken in infinite-dimensional calibration or scheduling but the advances in technology now mean that such transient control can be achieved more easily. Infinite-dimensional control is a term used throughout to convey that the system is modelled as a continuous

time-dependent process rather than approximated to a discrete number of either vehicle speed/acceleration or engine speed/load demand points. The omission of such transient effects has consequences on the accuracy and applicability of the results but the role of finite-dimensional studies in the control process is such that it lays the foundation and input to more complex dynamical control.

Both approaches to the emission control problem have covered the two central themes of control. The first is concerned with closed loop control and feedback principles. For engine control this includes the feedback control of idle speed, after-exhaust emissions and spark advance, and includes intelligent control methods such as adaptive and predictive control, [15]. The second approach is open loop control formulated by optimal control principles and generally makes use of search methods and Lagrange Multiplier theory.

The overall goal of either control approach is to provide the best fuel economy and emissions output together with acceptable driveability. It is the optimal control approach however, that is pertinent to powertrain control scheduling, and is the approach taken here. That is, to transfer the demand of the driver into a computed gear ratio and engine torque which when achieved by the actuator, will result in the desired vehicle speed. This modulation of the power demand, termed powertrain scheduling, is, as stated, an approach applicable to both S.I. and diesel-engined vehicles.

The first step in the formulation is the production of mathematical rules for gearshifts, through the optimisation process, which are then implemented as schedules in a control microprocessor. Thus, formulation and not implementation is the major concern here although some discussion will be given to the general implementation of control schedules onto vehicles.

Previous attempts at applying optimal scheduling to powertrain control have almost exclusively consisted of finite-dimensional approximations at discrete speed/load points. The finite-dimensional formulation presented here, will be seen to produce not only acceptable vehicle control but also provide the fundamental building block to more dynamic optimisation studies. Use of this finite-dimensional information allows a more formal approach to the dynamic problem, [22, 23].

The most important flaw of previous work however, is the treatment of the control objectives, specifically that of the fuel and emissions. All previous work has applied the minimisation process to one of the objective functions only, writing the remaining objectives as constraints on the optimisation. As discussed, a trade-off is unavoidable between the full set of objectives, and this cannot be fully exploited by previous formulations. As such, analysis of the trade-offs has to be done manually for the full extent of the compromise space to become apparent. The true optimisation environment is one of minimising many contending objectives simultaneously. This can be done by making use of well established multiple-objective optimisation procedures, [24], which are used extensively in manufacturing and socio-economic spheres. Application to control has until now been limited to compensator design, [82]; applied to the fuel and emission minimisation process of optimal scheduling, these methods can be encompassed into an interactive computer tool capable of producing trade-off solutions, the exact nature of which are dependent on the specifications of the engineer and legislative restrictions.

A novel approach to fuel and emission minimisation is thus presented here. Taking a powertrain control approach, not only provides a solution to emission reduction in diesel engines, where other control is unsuitable, but is also equally applicable to petrol-engined vehicles. The scheduling which it involves can be optimised to

produce the best fuel and emission production given the demands of the driver. If implemented in a finite-dimensional regime, the results not only stand up for themselves in terms of implementable gear schedules, but also lead into infinite-dimensional formulations which include time continuous transient effects, especially important for C.V.T. vehicles. Most importantly however, the inherent objectives of fuel and emission minimisation, given that the driver demand must be satisfied, can be formulated as a multiple-objective optimisation problem, the result of which is a trade-off solution to gear scheduling without time consuming manual analysis.

1.2 Aims and Scope

The general aim of this study is to apply nonlinear techniques to automotive control problems, specifically the problem of fuel and emission minimisation. This is undertaken through powertrain scheduling and control, using nonlinear optimisation techniques and the novel application of multi-objective optimisation methods. Steady state optimisation of the powertrain and dynamic simulation of such will allow rule-based control laws to be formulated.

The study here is specifically applied to a diesel-engined vehicle containing a C.V.T. . The powertrain scheduling approach applied to diesel engines is itself novel, but is not restricted to this alone, and as will be discussed can be used for other engine and transmission types.

A discussion of previous automotive control studies, particularly that of fuel and emission minimisation, is presented in Chapter 2. Models used in the study here are also introduced. This includes the steady state model used for optimisation purposes and a more fully dynamic model used for simulation and control analysis.

The powertrain is inherently nonlinear and unsuitable to linear analysis. Chapter 3 therefore discusses the possible elements of a nonlinear system, classifies them, and introduces some of the techniques available for their analysis. Particular emphasis is given to optimisation methods, the basic techniques of which are discussed in some detail. The discussion is restricted to finite-dimensional optimisation.

Chapter 4 develops the optimisation formulation for the scheduling problem, using methods discussed in Chapter 3. The environment is single-objective optimisation, each fuel or emission function minimised separately. This, together with the use of constraints on the other functions, is the classical approach to powertrain scheduling. Here, it will be used as a first pass solution to gain knowledge of the underlying complex interactions of the system and to give motivation to the use of multi-objective optimisation methods.

The application of multi-objective methods to the scheduling problem is discussed in Chapter 5. The three basic techniques : Weighted Sum, ε-Constraint, and Goal Attainment, are used for the fuel and emission minimisation problem, and the applicability of each discussed. Particular emphasis is given to the Goal Attainment method, which is used in Chapter 6 to develop the scheduling problem and to formulate a control scheme which is able to both dynamically track the test cycle and reproduce the optimality of the resultant schedules. Dynamic simulation is used to measure how far these have been achieved. Finally, the driveability of the control scheme is discussed together with simulation studies to test its sensitivity to varying test cycles.

Chapter 7 draws some conclusions from the work and makes some suggestions for further study.

(g)	Old Data	New data	Limit
fuel	112.16	109.22	-
part	0.20	0.26	0.338
NOx	2.70	1.5	1.28
HC	0.65	0.27	1.0

Table 1.1: The Effect of Engine Design Improvement in Diesel Engines

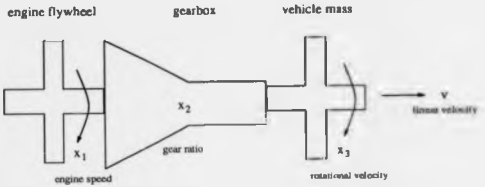


Figure 1.1: Schematic of the Powertrain

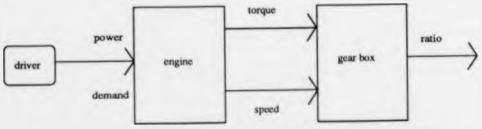


Figure 1.2: The Driver's Demand as Seen by the Powertrain

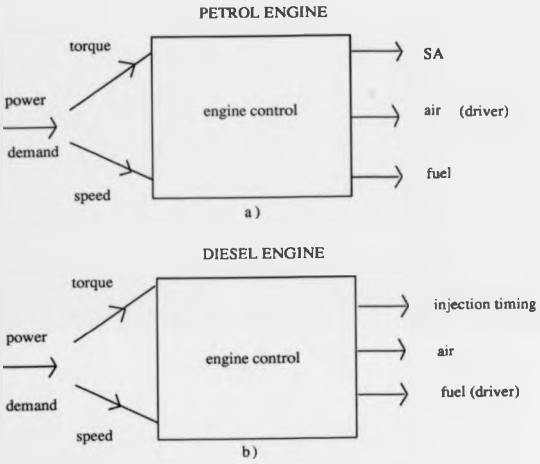


Figure 1.3: Engine Control For Fuel and Emissions

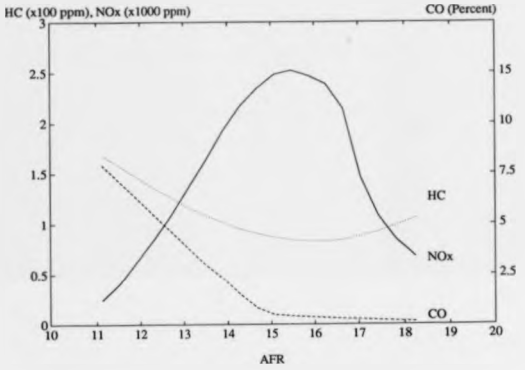


Figure 1.4: The Effect of A/F on Fuel and Emissions

Chapter 2

Modelling and Control of Automotive Powertrains

2.1 Introduction

There are many incentives for the use of modelling and simulation in engineering development work; not least of which is the reduction in time and cost of such work and the assessment and attainment of more optimal design configurations, [13].

The choice of mathematical model for a given system is crucial to the resultant control analysis and design. It should reflect the important characteristics of the system while neglecting those over-complications which will only hinder the control analysis and design process. Most successful applications of control theory are based on simple mathematical models, and the models presented here will only reflect the relevant components of the powertrain and their interactions, not their utmost detail.

This chapter gives an overview of powertrain modelling and control approaches to date, and will introduce the models to which the optimisation techniques are applied.

Previous modelling approaches can be classified into engine modelling; transmission modelling, often as a part of complete powertrain models; and modelling of other subsystem units, such as decoupling devices. Approaches to spark ignition, (S.I.), and diesel-engined vehicles will be discussed, together with the two facets of control : optimal scheduling and feedback.

The primary aim of the models presented here, is to formulate the powertrain control problem of transferring the demand of a test cycle, in the form of velocity and acceleration trajectories, into an engine torque, gear ratio profile. The logical flow of the model must thus be to take a velocity and acceleration input at the tyre/road interface, to transfer this back through the gearbox, and then to establish the engine power demand. The formulation here thus closely follows the inverse modelling approach taken by Blumberg, [13], although the exact powertain configurations differ.

Three models are presented in Section 2.3, which will appear throughout later analysis and reflect the level of information required for a specific formulation. The first, a finite-dimensional model, is based on the driver's instantaneous velocity and acceleration demand, for which an engine torque and gear ratio must be selected. This model forms the heart of the optimisation studies. The second, a more fully dynamic formulation, incorporates flywheel and C.V.T. actuation dynamics. The resultant equations are time-dependent with vector-valued variables over an interval [0,T]. This is used solely as an introduction to the final more complex dynamic model. The final model is used for both illustrative purposes in Section 2.4 and for control implementation purposes in Chapter 6, where driveability issues will become important. It is slightly more detailed than the previous model by the inclusion of driveline compliance, downstream of the gearbox. In Section 2.4 the

degree of nonlinearity inherent in the driveline will be shown through simulation, by observing the response of the vehicle to a step input in gear ratio.

All of the models presented are for a diesel-engined vehicle incorporating a C.V.T., with model and constraint data obtained from engine test bed regression analyses. Some vehicle data was unobtainable for the particular test vehicle and estimates based on previous studies have been used. Although the type of engine-vehicle configuration will have implications on the resultant control solutions, this does not restrict subsequent analyses. The models are equally applicable to other engine and transmission types, given the relevant data. The consequence of including a diesel engine or continuously variable transmission are expanded upon in Section 2.3.1, where the aim and scope of the models is also discussed.

Each model is constrained by physically based bounds on the state and control variables, which in turn act as constraints on the optimisation of the control objectives. That is, each vehicle model has an associated engine map which defines the fuel and emission flow over the full operating range. These fuel and emission flow maps are nonlinear polynomials in engine torque and speed, having been regressed from engine test bed data. They define the cost function which the model equations act to constrain in an optimisation environment. The extent of their nonlinearity, and the consequences for optimisation, is discussed separately in Section 2.3.2, since the objective function formulation is unrelated to the level of complexity of the powertrain model.

The powertrain model equations are also nonlinear functions of state and control variables. The extent of their nonlinearity is discussed in Section 2.4 where simulation of the compliant model is undertaken and will demonstrate the response of the C.V.T. to a step reduction in demanded gear ratio. This nonlinear nature of

both objective and constraint equations has profound consequences on the optimisation formulation, as does the recognition that there is more than one optimisation objective.

2.2 Approaches to Powertrain Modelling

Automotive powertrains can be conveniently considered to consist of several interacting subsystems, *i.e.* engine, transmission, road wheels, vehicle, control system and driver, the interaction of the models of which can be seen in Figure 2.1. The approach is based upon the interaction of discrete subsystem models, derived from regression data or physical laws, linked to a controller which interacts with driver and vehicle. In general the individual subsytems are complex and result in a system which is inherently nonlinear. The analytical treatments of each have been covered to a greater and lesser extent in the literature under the different guises of performance, economy and safety. Many powertrain simulations exist which are chiefly geared to the improvement of fuel economy and emission output, either by engine management or powertrain control. Specific vehicle configurations, and general purpose models, have been developed in both the finite and infinite-dimensional regimes, although the latter are far rarer. Other classifications, [25], distinguish between the causal and non-causal (inverse model) type simulations.

Most engine transmission studies are based upon a specific driving cycle. Introduced to standardise legislation, the three basic driving cycles are each tailored to the particular driving habits and legislative need of the respective continents. The European and Japanese test cycles are considerably simpler than the American Federal Test Procedure (F.T.P.) cycles. The cycles, shown in Figure 2.2, specify the

vehicle speed, acceleration demand and gear shift points at discrete time intervals, nominally one second, with the dynamic nature of the test varying between cycles. The ECE-15 cycle shown, used within most European countries, and used in this work, consists of clear stages of idle, acceleration, deceleration, and cruise.

Previous modelling efforts, have attacked all areas of the powertrain but modelling of the engine has perhaps had the most frequent appearance in the literature. Control was applied first to this area in the early 1970's, [26] and it was not until the mid 1980's that control of other parts of the powertrain started to catch up. The following discussion of powertrain modelling reflects this, and the fact that although every effort has been made to include the most important contributions, the inclusion and discussion of all papers is impossible. Only those pertinent to the optimisation and control of fuel and emissions are discussed in detail, and the reader is referred to such review papers as Powell, [14], and Costa, [2] for detail of other criteria.

2.2.1 Engine Modelling and Control

The use of engine models has the potential to reduce the need to repeatedly test a large number of engine map points on engine dynamometers, reducing both time and expense. During their evolution there have been many different formulations, [14]. Some are based entirely on measurements of the relevant engine outputs for varying levels of control inputs. These are referred to as 'input-output' models and are usually directed towards steady state control analysis and design. Others are purely based on physical principles, although even these need some degree of input-output information. These may be steady state or dynamic in nature and although they had their genesis in the 1960's have not come into widespread use until the last

decade, [14]. Some authors prefer not to use models at all and work exclusively with a mixture of engine and computer hardware, in an effort to capture fully the transient nature, [39]. However, the flexibility, repeatability, and availability of engine models means that they will always be a strong feature of engine control design.

Engine models vary in detail dependent on their ultimate usage, and with the effective cost of obtaining them. Static input-output models, together with regression techniques, have been used widely for calibration due to the ease and relative cheapness of obtaining them. Dynamic versions of the same procedure tend to cross the bounds between these and physically based models with system identification relying on the derivation of differential equations from physical principles and observation of input-output data, [9]. The latter are becoming increasingly popular with the rapid growth in microprocessor control. Both types have been used in engine design studies, fuel and emission control and torque transient analysis, but dynamic physical models tend still to be used more to further the understanding of the transient nature of the engine than for control implementation studies. Much of the early work on dynamic internal combustion engine models was applied to diesel engines, since most are turbo-charged creating a much slower set of dynamic response characteristics, [14]. Work has included modelling of the dynamic behaviour of the engine speed/load characteristics using sampled data techniques, [30], modelling of the combustion process and the effect of modelling complexity, [29], evaluation of a dynamic engine model and its use in the production of a linear transfer function for control studies, [31], and the use of state-space sampled-data analysis for control design, [32]. Most fuel and emission modelling studies however, have concentrated on S.I. engines and the use of static engine speed-torque maps together with regression techniques. However, with the microprocessor control of vehicles now well

established, dynamic models are becoming increasingly popular, [10, 14, 27].

The control of spark-ignited petrol engines typically embodies the calibration of ignition timing, air-fuel ratio (A/F), and Exhaust Gas Recirculation (E.G.R.). This has been attempted in both an open-loop manner, [33], [34], and through feedback principles of closed-loop control, [11], [28]. The former involves steady state inverse model representations to map the characteristics of the driving schedule to the power requirements at the engine. In fact, there has been a great deal of interest in optimisation approaches to engine control, lead by an initial assessment by Prabhabkar et al. (1975), [33]. Applying a warmed-up engine control problem in conjunction with a vehicle model, Prabhabkar used a gradient search technique to determine the engine calibrations. Blumberg (1975), [13], used an inverse vehicle model to discretise the engine map into a limited number of regions for projection of fuel and economy emissions, although no calibration was attempted. Richavy et al. (1977), [34], applied a linear programming approach to a discrete approximation of the driving cycle and Cassidy (1977), [35], developed on-line methods for engine calibration at selected speed-load points, without the need for a mathematical model. Rao (1977), [36], used nonlinear programming methods to optimise controls whereas a dynamic programming approach was taken by Auiler (1977), [37], to allocate the contribution to emission output of the speed-load points on a test schedule for maximum fuel economy. Dohner (1978,1981), [39, 38], attacked emission constrained optimisation by applying the Maximum Principle to a terminal control problem. Without using engines models he obtained gradient information by perturbing the controls one at a time and running the engine over the entire F.T.P. cycle. Matsomuto (1978), [40], discussed the the development of powertrain simulation, optimisation and calibration implementation, using direct search, on an actual vehicle and more

recently, Tennant et al. (1983), [12], used a test bed derived steady state engine model together with Lagrange multiplier theory to optimise fuel economy subject to driveability and emission constraints.

Feedback and modern control concepts have also been applied to engine control in recent years. Blumberg (1981), [11], used feedback of E.G.R. to reduce the emissions of nitrous oxides on a vehicle having a three way catalytic converter. Adaptive open and closed-loop control schemes have been developed by Kiencke (1988), [15], for A/F and knock control and has combined partially dynamic models with static control maps.

In contrast, diesel engine control is not amenable to engine calibration methods. The nature of the engine is such that only fuel quantity and injection timing are available as control variables. That is, the time length of the injector pulse and the start of injection timing. Air quantity can be altered by E.G.R. in turbo-charged vehicles but these are only found on large commercial vehicles and almost exclusively in the U.S.A.. Further to this, the engine operating conditions of a diesel engine are narrower in range for fuel efficiency and smoke production to be within customer acceptable bounds, even before legislative bounds are considered, [8]. The engine must be run on excess air and thus catalytic converters, which must operate around stoichiometry, cannot be used either. Thus modelling and control of diesel engines has tended to concentrate on engine design investigation, [7], [41], [42], rather than calibration.

2.2.2 Transmission Modelling and Control

Powertrain control, involving shift execution and the scheduling of gears, has been much less widely tackled in the literature. Gear train scheduling, as with the former approach, is most often posed as either a fuel and emissions problem or under performance criteria. Shift execution control is usually undertaken as a means to improve driveability and performance.

For fuel and emission reduction through gear scheduling, treatments have again concentrated on the use of steady state models. For discrete ratios Thring (1981), [16], looked at the effect on fuel economy of changing the number of gears, while Wong (1979), [17], looked at shift scheduling for maximising fuel usage by optimum power matching. Both used analysis rather than optimisation techniques.

Much of the optimisation work has been undertaken using vehicles incorporating C.V.T.'s. Christenson et al., [18], looked at the fuel saving potential of a C.V.T. through inverse model analysis similar to Blumberg's, [13]. Radtke et at, [21], used 5^{th} order regression equations with a static powertrain model to calculate the maximum fuel economy improvement that could be achieved with a C.V.T. . Using a combination of both engine calibration variables and gear ratio, Lagrange Multipliers were used to formulate the problem as a constrained optimisation. This meant that emissions objectives were formulated as constraints rather than optimised variables, and as such no trade-off was made between the conflicting requirements. Kuzak et al., [19], attempted the same fuel economy, emission constrained formulation but with dynamic programming as the search method. Employing a second by second optimisation on engine calibration variables and shift points, Kuzak reduced the infinite shift selection possibilities for a C.V.T. vehicle down to a reduced number of choices through a sequential decision making process. This reduced the usual

lengthy process of dynamic programming by eliminating non-optimal shift patterns. However, the method is still computationally burdensome and the nature of the search is such that only discrete points on the objective function space are tested. Although, it did incorporate dynamic effects such as shift frequency and driveability, the analysis still did not exploit the trade-offs between fuel and emissions.

Another important aspect is that neither Radtke or Kuzak explicitly considered the manipulation of engine torque and speed in their studies. Neither formulation allowed the engine torque to act as an optimised variable. Whilst Radtke's formulation did calculate the optimal fuel economy, it indicates the optimal ratio selection only as a function of driveshaft speed and load and only at a limited number of points. Kuzak developed vehicle implementable shift strategies based on piecewise linear approximations as a function of driveshaft speed and throttle angle, making the ratio selection a complex function of two variables.

Full dynamic optimisation of fuel economy has been undertaken by Jones (1991), [22], using the Maximum Principle, for a vehicle incorporating a C.V.T., using a two state, two control variable model. Infinite-dimensional optimisation was used to minimise fuel economy over the whole driving cycle in one pass of the optimisation, whilst also tracking the test cycle. Improvements to this initial formulation were made by Garbett et al., [23], using steady state time history solutions as suboptimal approximations to the dynamic problem.

Practical shift scheduling of C.V.T.'s has traditionally been undertaken by driving the vehicle up the 'economy line' - the locus of steady state engine speed-torque solution points, [43]. This constrains the vehicle to optimum constant power points where no acceleration demand is considered. Specific-fuel consumption and not absolute fuel consumption has been traditionally minimised, which has the disadvantage

of constraining the analysis to the positive region of the speed-torque map.

Diesel-engined vehicle transmission control to reduce fuel and emission output, has until now escaped any major treatment in the literature. This may be because the diesel engine is, by nature, intrinsically cleaner than the petrol engine. However, with legislation tightening on the commercial market at as quick a pace as for the passenger car market, and with current methods unable to fully solve the particulate emission problem, [6], powertrain scheduling approaches are more valuable now than ever before.

2.2.3 Other Vehicle Modelling and Control Issues

The design and development of interactive models of vehicle subsystems also requires the inclusion of torque converter or clutch models. The driveability problems associated with unwelcome interactions between subsystems are concealed to a great extent by the torque converter in an automatic vehicle and controlled by the clutch in a manually shifted vehicle.

Traditionally modelled as input-output systems, developments are progressing towards nonlinear models such as those developed by Kotwicki, [45], for torque converters, and Moskwa and Hedrick, [46] for clutches. The control of torque transients and the linked subject of shunt control can be handled to a large extent by smoother clutch action. This can be approached through the use of nonlinear models and rule-based control, [44], [49], or through advanced nonlinear techniques such as sliding mode control, [46], [47].

2.3 The Powertrain Model

2.3.1 Aim and Scope

The models presented in the following sections are used in subsequent chapters to simulate and optimise the vehicle fuel and emissions output over a specified driving cycle. As such, they need to translate the cycle demanded vehicle speed and acceleration into an engine torque, speed and gear ratio profile. They thus follow the inverse model methodology of Blumberg, [13]. The first is a finite-dimensional formulation using a discrete approximation to the test cycle at one second intervals, the inputs of which are therefore scalar. The second model is a more complicated infinite-dimensional formulation where the variables are now continuous vectors over time. Finally, a compliant version of the same model is presented.

The models have been kept simple for implementational purposes but are flexible enough for other dynamic considerations to be added if necessary. Engine block dynamics which have a frequency of approximately $7Hz$ are not included, since this is less important in a continuously variable transmission where shift changes are smoother. Dynamic engine effects associated with the combustion process and the intake of air and fuel, together with the cyclic production of torque are ignored in this model since these too are outside the frequency range of interest, and are more applicable to engine than transmission control. Also, the engine torque is assumed to be directly available as a control variable, together with vehicle braking force and C.V.T. ratio demand.

The models are based on physical principles whilst the objective and constraint functions are derived from regression of test bed data. The specific configuration is that of a diesel-engined vehicle with a C.V.T., in a medium sized commercial vehi-

cle. Unless otherwise stated, vehicle model parameters, not derived from regression analysis, are based on previous modelling studies, [22].

The Diesel Engine

Diesel engines due to their increased efficiency but increased noise and weight problems have traditionally been exploited in the light to heavy commercial vehicle market. Unlike petrol engines, which are most efficiently run at stoichiometry or lean of it, diesel engines must be run with an excess of air and at higher temperatures. Failure to do so not only causes inefficient running of the engine but also causes increased output of soot and hydrocarbons (H.C.), [8]. For a properly controlled engine, carbon monoxide, and gaseous hydrocarbon output is relatively low. However, one of the major disadvantages is the relatively high level of particulate emission, not only responsible for the blackening of buildings, but also found to contain carcinogenic agents, [48].

The monopoly of the S.I. engine in the overall market has meant that a plethora of techniques have been developed to reduce emission output without a reduction in fuel economy. Since most of these techniques are based on the two basic variables of air supply and temperature they do not easily transfer across to diesel engines. As such, the commercial vehicle market is in a quandary as to how to achieve the ever increasing emissions legislation, and especially the stringent particulate measures due to be introduced by American legislation in 1994, [6].

The industrial nature of this project has influenced the type of data available and as such diesel engine data has been used throughout. This makes little difference to the techniques used, which are applicable to any engine design which can be characterised as engine speed and torque pairs. However, any resultant control

cle. Unless otherwise stated, vehicle model parameters, not derived from regression analysis, are based on previous modelling studies, [22].

The Diesel Engine

Diesel engines due to their increased efficiency but increased noise and weight problems have traditionally been exploited in the light to heavy commercial vehicle market. Unlike petrol engines, which are most efficiently run at stoichiometry or lean of it, diesel engines must be run with an excess of air and at higher temperatures. Failure to do so not only causes inefficient running of the engine but also causes increased output of soot and hydrocarbons (H.C.), [8]. For a properly controlled engine, carbon monoxide, and gaseous hydrocarbon output is relatively low. However, one of the major disadvantages is the relatively high level of particulate emission, not only responsible for the blackening of buildings, but also found to contain carcinogenic agents, [48].

The monopoly of the S.I. engine in the overall market has meant that a plethora of techniques have been developed to reduce emission output without a reduction in fuel economy. Since most of these techniques are based on the two basic variables of air supply and temperature they do not easily transfer across to diesel engines. As such, the commercial vehicle market is in a quandary as to how to achieve the ever increasing emissions legislation, and especially the stringent particulate measures due to be introduced by American legislation in 1994, [6].

The industrial nature of this project has influenced the type of data available and as such diesel engine data has been used throughout. This makes little difference to the techniques used, which are applicable to any engine design which can be characterised as engine speed and torque pairs. However, any resultant control

strategies will differ to some extent.

The C.V.T.

Continuously variable transmissions were used for many years in the machine tool industry and as variable speed transmissions in many small vehicles such as tractors. The automobile industry's interest in them started in the 1960's with periodically renewed interest since then. Most recently two major car manufacturers, Ford and Fiat, have introduced C.V.T.'s into their smaller saloon cars, with other manufacturers expected to follow.

The C.V.T. has been used in the models here not only because it is interesting in its own right but, because of its very nature, the differential equations which describe it are continuous. This makes both the optimisation formulation and control analysis easier mathematically, and in essence will give the 'optimal' gear scheduling solution to which manual gearboxes can only approximate. If a manual gearbox solution is required this approximation theory would have to be carried out.

The exact mechanics of how they work is not needed for the work presented here, but an insight into their nature is worthwhile. The type modelled here is a Perbury transmission which utilises the varying contact of cylindrical rollers between rotating discs to accomplish the continuous ratio variation. Unlike manual transmissions, which typically have four or five discrete gear ratios for passenger cars and as many as fifteen ratios for heavy goods vehicles, the C.V.T. is characterised as having an infinite range of ratios r within certain bounds, r_{min} and r_{max}. The gear ratio is defined here as the ratio of gearbox output speed to gearbox input speed, and its

variation can be defined mathematically as,

$$r \in [r_1, r_2, ...r_n], \quad n = 1, ..5, .. \tag{2.1}$$

and

$$r \in [r_{min}, r_{max}] \tag{2.2}$$

for discrete and continuous gearboxes respectively. The minimum ratio achievable is dependent on the C.V.T. used. Many incorporate torque converters or clutches which decouple the engine from the vehicle. The Perbury transmission however, can accommodate a zero minimum gear ratio, or 'geared neutral'.

The result of such a wide range of ratios is that the vehicle is able to hug the maximum power curve at all times. With a discrete ratio gearbox the vehicle falls away from the maximum power curve at the top end of each gear ratio selection, Figure 2.3. This means that for a C.V.T. vehicle a constant engine speed can be maintained under acceleration, and can be chosen such that fuel economy or emission output is much lower.

The benefits credited to the C.V.T. thus include improved fuel economy and acceleration performance, more flexible engine and vehicle control strategies and weight-size reductions to the vehicle, [50]. Under optimisation, they produce truly 'optimal' control strategies compared with discrete gear sets which can only give approximate or 'suboptimal' solutions to the continuous trajectories.

2.3.2 Regression of Engine Data

The powertrain models presented in this chapter are based on the state equations of physical laws. However, the model constraints and accompanying fuel and emission

functions are based on engine test bed data for which regression functions must be defined.

The engine data used throughout is for a 6cyl. turbo-charged diesel engine, nominally used on a light to medium commercial vehicle, with engine capacity of 2500cc. and E.G.R. . The test-bed procedure involved the steady state mapping of fuel and emission output at each engine speed/torque point at varying fuelling levels.

Engine Characteristics

The first information which can be obtained from the test bed data is an indication of the characteristics of the engine in the form of a peak torque and overrun curve. The peak torque curve, relating to maximum fuelling, is obtained by taking the upper torque values at each speed point. These are then regressed as a second order polynomial in engine speed using least-square fit techniques. The resultant coefficients are shown in Table 2.1 and the peak torque curve which they define shown in Figure 2.4. The peak torque curve consists of a second order polynomial which is defined over most of the operating range together with two 'lug-down' speed areas which define the torque output in the low and high engine speed regions. Each region was regressed separately, with coefficients a_1, a_2, a_3, defining the central area and d_1, d_2, e_1, e_2, defining the first and second straight line portions respectively. The exact equation and its role as a bound on the engine torque control variable is discussed further in Section 2.3.3 together with other physical bounds. Overrun torque data, regressed in the same way, is also shown in Figure 2.4, and is a first order function of engine speed, with coefficients b_1, b_1 as defined in Table 2.1. Together with engine idle and maximum speeds, taken as the minimum and maximum engine

test points respectively, the peak and overrun torque curves define the available engine power at any point in time.

Fuel and Emission Characteristics

Having characterised the engine's power constraints, the fuel and emission output flow maps can be characterised from the same data. This is achieved using MATLAB facilities in the form of user-written m-files, [51]. The programs define an orthogonal polynomial fitting routine able to produce functions of the form,

$$f = a_1 + a_2x_1 + a_3u_1 + a_4{x_1}^2 + a_5{u_1}^2 + a_6{x_1}^3 + a_7{x_1}^2u_1 + \ldots + a_m{u_1}^n \qquad (2.3)$$

for engine speed x_1 and engine torque u_1, [77]. The degree of polynomial, n, produced is defined by an input variable. Statistical analysis was undertaken to determine the significance of increasing the order of fit, and determination of the best fit to the data. All functions were regressed to an accuracy of at least two decimal places, and the resultant flow functions are shown in Figure 2.5 ... 2.9, with the order of regression given on the plot. The regression coefficients are shown in Table 2.2

The character of the plot is very important with regards optimisation since it is the fuel and emission flow functions which define the cost function in later analyses. The degree of nonlinearity and extent to which the function is convex will determine the type of optimisation procedure used, how well it converges and whether the ultimate solution is the global minimum. It is very important therefore to have a good graphical knowledge of the objective functions being minimised.

The fuel function is a fourth order function and is seen to be convex within

the bounds of the engine operating range, Figure 2.5, with the global minimum being the only minimum within this range. The NO_x, and H.C flow functions were also regressed to fourth order but can be seen to be far more complex, Figures 2.7, 2.8. Both are non-convex functions and have more than one minimum within the operating range. Particulate and smoke output flow have been regressed as fifth order functions, Figures 2.6, 2.9 and again are nonconvex over a wide area of the operating range.

One problem created by the nonlinear nature of some of the functions, is the accuracy of the regression in the low flow areas. For fuel and smoke functions only, in order to obtain regressions of a high enough accuracy over the major part of the operating range, negative flow areas can result. This is a modelling anomaly, and of course cannot take place in practice. However, steps can be taken, as will be discussed in Chapter 4 to overcome this for optimisation purposes.

All the figures show functions of absolute flow in g/s rather than the more traditional specific flow values in g/KWh. This is due to the formulation of the optimisation problem, as will be discussed in Chapter 4. It allows the functions to be defined over the whole operating range, whereas specific-flow is undefined at zero torque.

Vehicle Characteristics

The vehicle model used has a mass M kg, engine and gearbox inertias I and $I_g kgm^2$, wheel radius, r_w m, and final drive ratio, r_d , as shown in Table 2.1. The vehicle mass is an estimate based on the actual test-bed vehicle used. A graphical formulation of the road load for the test-bed vehicle was provided and, using regression techniques, a second order polynomial obtained, as illustrated in Figure 2.10. The interaction

of this function with the rest of the vehicle model is discussed in Section 2.3.3.

2.3.3 Finite-Dimensional Model

Engine Power and Model Constraints

The finite-dimensional model discussed here, is described as such because of its role in subsequent optimisation formulations. The model is used to determine the relationship between the instantaneous driver demand, assumed constant over a sufficiently small time interval δt, nominally one second, and the power demand at the engine needed to achieve it. Thus, time is not explicitly involved in the equations.

The available power from the engine can be characterised by a peak torque and overrun curve as in Figure 2.11, and expressed mathematically as,

$$u_{1_{min}}(x_1) \leq u_1 \leq u_{1_{max}}(x_1), \tag{2.4}$$

where engine torque, u_1, is a function of engine speed, such that

$$u_{1_{max}}(x_1) = \begin{cases} a_1 x_1^2 + a_2 x_1 + a_3 & if\ x_{l_1} \leq x_1 < x_{l_2} \\ d_1 x_1 + d_2 & if\ x_{1_{min}} < x_1 < x_{l_1} \\ e_1 x_1 + e_2 & \text{if } x_{l_2} \leq x_1 \leq x_{1_{max}} \end{cases} \tag{2.5}$$

where x_{l_1} and x_{l_2} are where the respective peak torque curve changes occur, and where

$$u_{1_{min}}(x_1) = b_1 x_1 + b_2\,. \tag{2.6}$$

The available engine power is further constrained by physical bounds on the engine

speed x_1,

$$x_{1_{min}} \leq x_1 \leq x_{1_{max}} \tag{2.7}$$

Model regression parameters, as discussed in Section 2.3.2, and constraint values are shown in Table 2.1, together with gear ratio bounds as expressed by,

$$x_{2_{min}} \leq x_2 \leq x_{2_{max}} \tag{2.8}$$

where the gear ratio is defined as output over input shaft speeds. The application of a braking force can be modelled by a variable u_3 which is again dependent on physical limits such that,

$$0 \leq u_3 \leq \mu_s Mg \tag{2.9}$$

where μ_s in this case is the tyre/road adhesion coefficient and M and g are the vehicle mass and gravitational acceleration in kg and m/s^2 respectively. The exact value of the maximum braking is shown in Table 2.1, for an approximate adhesion coefficient of 0.4 corresponding to a wet road. The maximum brake force limit used here is thus somewhat lower than that possible for a dry road with good tyres when μ_s is around 0.8.

The Vehicle Model

The powertrain model can be shown schematically as in Figure 2.12, where the external forces acting on the vehicle are the reaction to the tractive force generated at the tyre/road interface and the road load and brake retarding forces. The road load term is the sum of the rolling friction and aerodynamic drag forces and is a

function of the vehicle speed,

$$R_l(v) = A + Bv + Cv^2, \tag{2.10}$$

where v is the vehicle speed in m/s and A, B, and C are regression parameters as shown in Table 2.1. Gravitational effects are ignored in this model. The torque output at the gearbox, u_g, is thus dependent on the road load, any braking force applied, and the force to accelerate the vehicle at a demanded acceleration, v, reflected through the wheel and final drive ratio,

$$u_g = \mu R_l + \mu u_3 + \mu M v, \tag{2.11}$$

for brake force u_3 ($Newtons$), and μ, a combination of wheel radius and final drive ratio, both shown separately in Table 2.1. The gearbox output torque must be balanced at the input to the gearbox by the torque at the flywheel u_1 (Nm), hence

$$u_1 = \mu x_2(M\dot{v} + R_l + u_3), \tag{2.12}$$

where x_2 is the selected gear ratio. Translating the linear motion of the vehicle into the rotational speed of the engine x_1 (rad/s) gives,

$$v = \mu x_1 x_2, \tag{2.13}$$

and equation 2.12 can now be expressed as,

$$u_1 = \frac{v(M\dot{v} + R_l + u_3)}{x_1}. \tag{2.14}$$

Since the road load, and tractive effort to produce the required acceleration are known, given an instantaneous demand in vehicle speed and acceleration, the development of the required engine power becomes a non-unique choice of engine torque and engine speed. The problem is thus one of optimisation - choosing an engine operating region and thus gear ratio which satisfy the power requirements and subsequently minimise a predefined cost function. The model is thus ideally suited to the finite-dimensional optimisation of a vehicle tracking a specific test cycle. For control purposes engine torque u_1, gear ratio x_2 and brake torque u_3 are available to modify the power output of the engine and the resultant tractive effort at the road surface. The specific formulation of this is discussed fully in Chapter 4.

2.3.4 Infinite-Dimensional Vehicle Model

Non-Compliant Model

The dynamic powertrain model is basically of the same nature as the steady state model just discussed, and is still modelled schematically as in Figure 2.12. The significant difference is that the models, rather than containing variables assumed constant over the interval of time δt, now vary with time. The dynamics of the gearbox and flywheel are thus no longer negligible and are included in the model.

The powertrain now consists of the nonlinear state variables, $x_1(t)$, $x_2(t)$, representing engine speed and gear ratio, and the control variables, $u_1(t)$, $u_2(t)$, and $u_3(t)$, of engine torque, gear ratio and brake force demand, with constraints as in Section 2.3.3, but with gear ratio demand and not the gear ratio now bounded. The gear ratio selection is no longer considered instantaneous and is modelled as a first

order lag, with unit gain, such that,

$$\dot{x_2}(t) = \frac{(-x_2(t) + u_2(t))}{T_c}, \tag{2.15}$$

i.e. the difference in demanded and actual gear ratio at time t, for a time constant of T_c. Conservation of energy across the powertrain leads to the state equation,

$$\begin{aligned}(I + (\mu x_2(t))^2 M)\dot{x_1}(t) &= u_1(t) - x_2(t)x_1(t)\dot{x_2}(t)M\mu^2 \\ &\quad -\mu x_2(t)(A + Bv(t) + Cv(t)^2) - \mu x_2(t)u_3(t). \end{aligned} \tag{2.16}$$

The engine torque is thus balanced by the road load and brake forces translated through the gearbox, and by the momentum change within the rotational elements of the C.V.T., denoted by the term,

$$x_2(t)x_1(t)\dot{x_2}(t)M\mu^2, \tag{2.17}$$

due to the continuous ratio variation. The efficiency of the CVT is assumed to be ideal due to the lack of appropriate data.

An inverse dynamic model can also be constructed, as in the previous section, where the velocity and acceleration demand is known. The model can then act to constrain the cost function

$$\min_{u_1} \int_0^T f(x_1, u_1)dt \tag{2.18}$$

under dynamic optimisation, where $f(x_1, u_1)$ is a fuel or emission function.

2.3.5 Dynamic Compliant Model

If a mechanical element or structure experiences a steady deformation when acted upon by steady state forces it is said to exhibit compliance. The model here includes compliance in the driveline downstream of the gearbox. It is effectively modelled as a spring and damper system and shown schematically in Figure 2.13. In practice however, the compliance is made up of both compliance in the halfshaft, and deformation in the tyre as a function of vehicle speed. A more representative model is that shown in Figure 2.14, with the tyre characteristic incorporated as a damper at the wheel and with the deformation in the halfshafts acting as a spring downstream of the gearbox. For the purposes here however, the first representation is adequate since this study is primarily interested in transmission and not traction control. Experimentally, the formulations are very similar.

Incorporation of compliance effects requires two additional state variables together with the gearbox inertia. The variables x_1 and x_2 are as before, but x_3 is now the driveline output speed and x_4 is the rotational deflection taking place along the driveline. The additional inertia of the gearbox is I_g, aggregated at the output shaft.

The bounds on state and control variables are as discussed in Section 2.3.4 and the C.V.T. actuation dynamics are still represented as in equation 2.15. The dynamics of the transmission can now be represented by

$$[I + I_g x_2(t)^2]\dot{x}_1(t) = u_1(t) - x_2(t)x_1(t)x_2(t)I_g - x_2(t)[kx_4(t) + d(x_1(t)x_2(t) - x_3(t))], \tag{2.19}$$

and downstream of the gearbox,

$$\dot{x}_4(t) = x_1(t)x_2(t) - x_3(t), \tag{2.20}$$

for damping rate d and stiffness in the driveline, k. The spring and damping factors have been adjusted from previous experimental data, [49] for the specific vehicle mass. Equation 2.19 again represents the conservation of energy across the power-train and balances the engine torque against momentum changes across the gearbox represented by,

$$x_2(t)x_1(t)\dot{x}_2(t)I_g, \tag{2.21}$$

and the aggregate driveline torque at the C.V.T. output shaft, represented by,

$$kx_4(t) + d(x_1(t)x_2(t) - x_3(t)) \tag{2.22}$$

The rotational motion of the vehicle driveline is represented as

$$\mu^2 M\dot{x}_3(t) = (kx_4(t)+d(x_1(t)x_2(t)-x_3(t)))-\mu(A+Bv(t)+Cv(t)^2)-\mu u_3(t), \tag{2.23}$$

for brake force u_3, vehicle mass M, and road load as before. The result is four nonlinear coupled state equations with three control inputs and velocity now represented by,

$$v = \mu x_3. \tag{2.24}$$

For optimisation rather than simulation purposes, an inverse dynamic compliant model can be constructed. As in the non-compliant case, knowledge of vehicle speed and acceleration allows the solution of the dynamic equations, which in turn act to

constrain a minimisation integral, typically

$$\min_{u_1,u_2} \int_0^T \lambda f(x_1,u_1) + (1-\lambda) g(x_1,v,x_2) dt. \tag{2.25}$$

The cost function consists of a weighted sum of a fuel or emission function together, in this case, with a driveability term which is a function of engine speed, vehicle speed, and gear ratio. Dynamic optimisation is not considered in this thesis. However, there is much scope for its application to the fuel and emission minimisation problem.

2.4 Dynamic Simulation

Dynamic simulation allows the use of nonlinear dynamic models to predict the effect of changes in system parameters, control strategy or just to investigate transient performance. This section uses dynamic simulation to illustrate the nonlinear effects inherent in an automotive powertrain.

The dynamic powertrain model of Section 2.3.5, having four nonlinear coupled equations of four states and three control variables, will be used, to show the dynamic effects of the generation of power within the engine; the delivery of power to the drive wheels via the transmission; and the resulting effect on the longitudinal motion of the vehicle. This will be achieved through the use of the general purpose high level simulation language ACSL, [52]. The powertrain model of Section 2.3.5 was known to have a small dynamic frequency range which primarily included an approximate component of $2Hz$ for the driveline and a slower frequency component due to the vehicle mass. This range of frequency has consequences on the choice of integration routine used within the simulation, and the Adams-Moulton variable step algorithm

has to be used for computational efficiency and accuracy, [53].

The simulation experiment involves investigating the response of the C.V.T. powertrain to a step reduction in demanded gear ratio u_2. The C.V.T. actuation time is around 2 seconds with the vehicle initially travelling at a steady speed of $6.16m/s$, approximately 25 miles/hour, and with initial gear ratio 0.8. This approximates to a manual gear ratio of 3. At time 2 seconds a demanded drop in gear ratio is made to 0.4, around second gear. Braking force u_3 is set to zero throughout.

The Figures 2.15 to 2.19 show the nonlinear response of the powertrain. Figure 2.17 displays the resulting C.V.T. ratio response which achieves its new value after 2s. The resulting response in vehicle speed is shown in Figure 2.15. The vehicle speed is seen to drop momentarily before rising slowly to a new value of $6.35m/s$ after 15 seconds. The engine speed, in Figure 2.17, shows the engine speed rising rapidly, then stabilizing to a slow climb. Figure 2.19 shows the driveline torque which oscillates before settling to its new value of $9.3Nm$. The compliance has the effect of producing a deflection in the driveline which can be seen in Figure 2.18. The effect is seen to settle after 4s to a new steady state deflection value.

Given that engine torque, or equivalent throttle position, is held constant throughout the simulation the rapid reduction in C.V.T. ratio after 2s decelerates the vehicle slightly at the expense of accelerating the engine flywheel. This is characteristic of a C.V.T. vehicle which is able to snatch energy from either vehicle or engine and place it somewhere else. This momentum transfer is contained within the nonlinear term

$$x_1 x_2 x_2 I_g \tag{2.26}$$

and in this case is seen to be giving energy to the flywheel. The result is that the engine, now seeing a lighter load on it due to the reduced gear ratio, and with the

torque level held constant starts to accelerate.

The decrease in gear ratio produces a rapid change in driveline deflection, which winds up in the negative direction. Nonlinear oscillations result and as the engine speed starts to rise the the driveline deflection settles to a constant positive value. The driveline torque shows a similar response.

The C.V.T. and driveline compliance thus show the nonlinear delivery of power to the road wheels and the resulting effect on vehicle longitudinal motion. As we have already seen the engine itself possesses a nonlinear torque characteristic, has a nonlinear fuel flow usage and emission output characteristics, and experiences external forces on the vehicle which are nonlinear in vehicle speed. The result is a powertrain model which is nonlinear in its basic form, its constraints and its optimisation cost functions.

2.5 Summary

A brief introduction has been given, within this chapter, to the previous research efforts in modelling and control of automotive powertrains. Engine, transmission and decoupling devices were considered with recognition that most of the work has been done with the engine subsystem. The three main models used within this work for optimisation, simulation, and control purposes, were also introduced in detail. These included the powertrain model itself, the state and control bounds and the related fuel and emission functions. It was also shown how some of this data was obtained from regression analysis. Finally, a simulation was undertaken of the dynamic compliant model to aid understanding of its inherent nature.

The development of the finite-dimensional model equations, where only the in-

stantaneous driver demand is considered, showed how a non-unique choice of engine operating region and gear ratio selection could be made. This poses an optimisation problem, such that the consequence of the choice can be monitored in order to minimise, for example, fuel and emissions. This is the essence of the optimisation problem considered in later chapters.

The model, shown to give a nonlinear response to a step input, is also constrained by nonlinear state and control bounds. Further, the fuel and emission functions linked to each model are nonlinear. In optimisation terms therefore, a nonlinear programming problem with nonlinear constraints is being considered. However, with more than one objective, i.e fuel and emission reduction, the problem is also one of multiple decision theory, discussed more fully in Chapter 5. Nonlinear theory is much less widely available than its linear counterpart, the methods that are available usually being specific to a particular type of nonlinearity. These methods of analysis and control are discussed in the next chapter where the important tools of optimisation will become apparent.

A	50	d_1	5.50
B	0.34	d_2	-449.15
C	0.0055	e_1	-5.0
T_c	0.5	e_2	2316.2
I	0.159	b_1	-0.045
I_g	0.05	b_2	1.68
M	1800	$x_{1_{max}}$	460.77
r_w	0.28	$x_{1_{min}}$	89.0
r_d	0.275	$x_{2_{max}}$	2.0
μ	0.077	$x_{2_{min}}$	0.0
a_1	-0.0029	$u_{2_{max}}$	2.0
a_2	1.613	$u_{2_{min}}$	0.0
a_3	-3.093	$u_{3_{max}}$	6500
k_c	318	$u_{3_{min}}$	0.0
d_c	4.69		

Table 2.1: Model Parameter Coefficients

Fuel	Particulate	NO_x	HC	Smoke
2.0021	0.0044	0.0314	0.0040	0.0031
1.8141	0.0036	0.0377	0.0011	0.0015
1.6529	-0.0054	0.1002	0.0004	-0.0046
0.4515	0.008	0.0364	-0.0097	0.0077
1.4432	-0.010	0.0811	0.0051	-0.0094
-0.0389	-0.0067	0.0764	0.0012	-0.0073
-0.0457	0.011	0.0024	-0.0001	0.0179
0.3395	-0.0015	0.0104	0.0010	-0.0091
-0.3782	-0.008	0.0151	0.0024	-0.0063
0.1405	0.0182	-0.0299	-0.0017	0.0166
0.0014	-0.0017	-0.0128	0.0107	-0.0012
-0.4437	-0.0024	-0.0269	-0.0111	-0.0031
0.3182	-0.008	-0.0061	0.0055	0.0097
-0.2996	0.0188	-0.0056	-0.0019	0.0175
0.0823	0.0078	-0.0296	0.0010	0.0069
	-0.0071			-0.0129
	-0.0013			0.0096
	0.0006			-0.0056
	0.0149			0.0157
	0.0101			0.0092
	-0.0113			-0.0093

Table 2.2: Fuel and Emission Regression Parameters

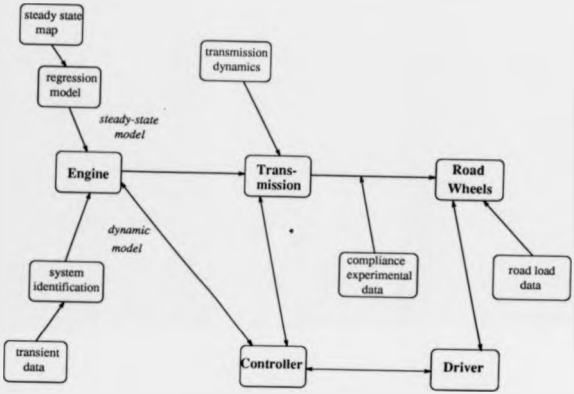


Figure 2.1: The Modelling Environment

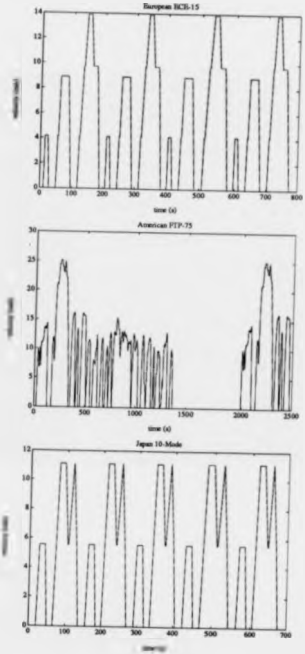


Figure 2.2: Worldwide Legislative Test Cycles

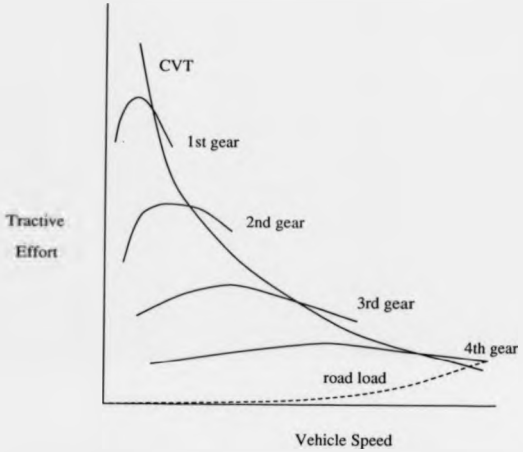


Figure 2.3: Tractive Effort Characteristics of Traditional Gearbox and CVT

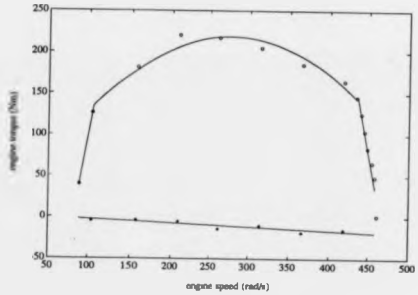


Figure 2.4: Regression of Engine Characteristics

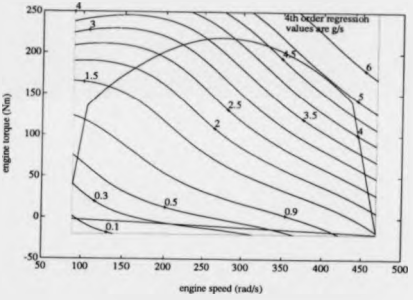


Figure 2.5: Fuel Contour Map

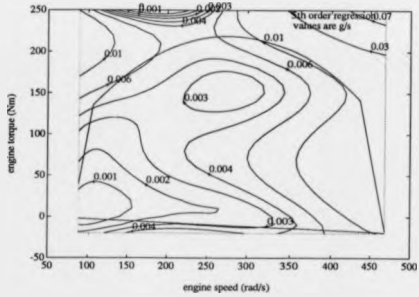


Figure 2.6: Particulate Contour Map

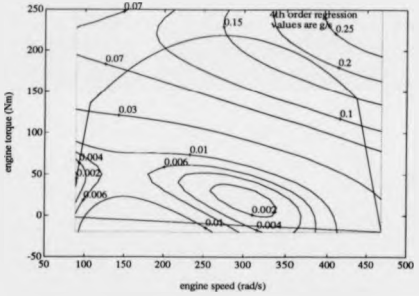


Figure 2.7: NO_x Contour Map

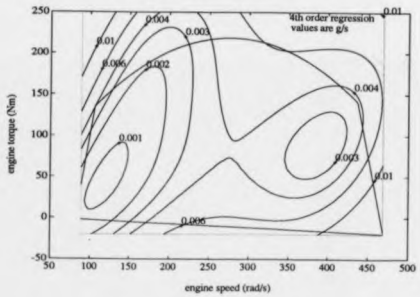


Figure 2.8: H.C Contour map

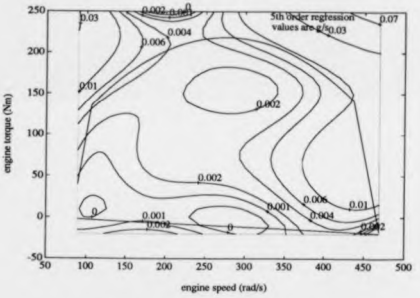


Figure 2.9: Smoke Contour Map

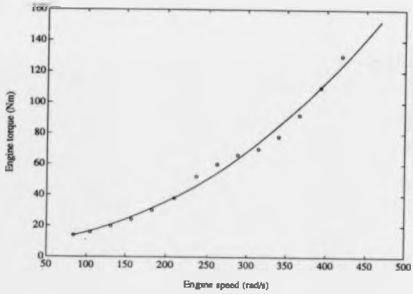


Figure 2.10: Rolling Resistence of Vehicle

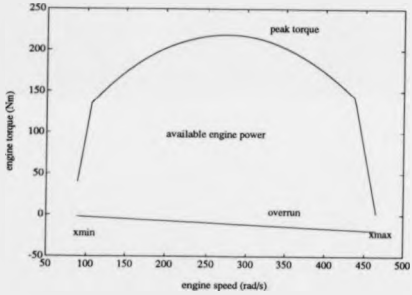


Figure 2.11: Engine Characteristics

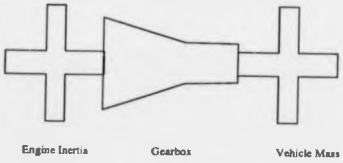


Figure 2.12: Schematic of Powertrain Model

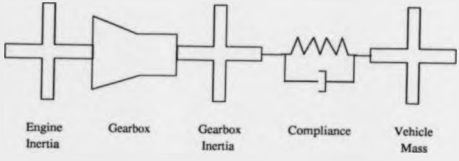


Figure 2.13: Schematic of Compliant Powertrain Model

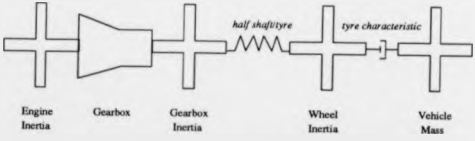


Figure 2.14: A More Representative Model For Compliant Powertrain

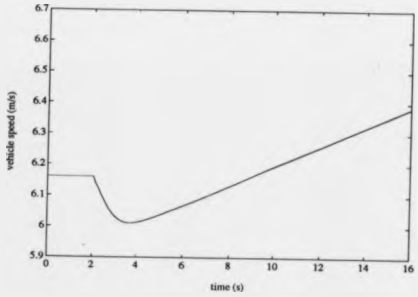


Figure 2.15: Vehicle Speed Response

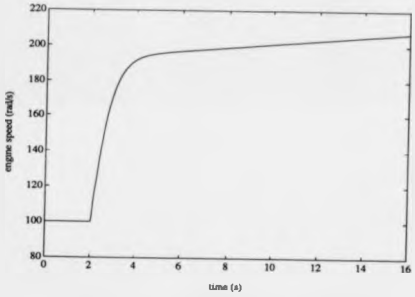


Figure 2.16: Engine Speed Response

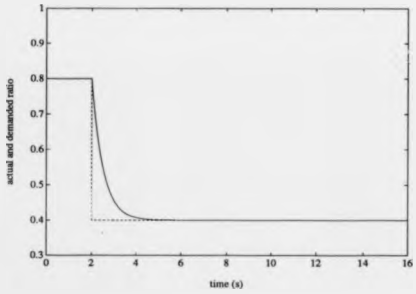


Figure 2.17: Gear Demand Speed Response

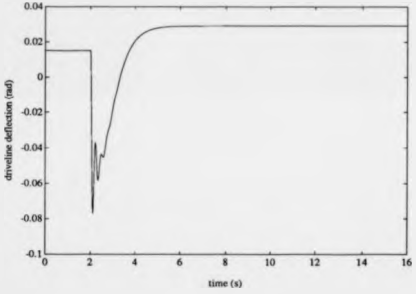


Figure 2.18: Driveline Deflection Response

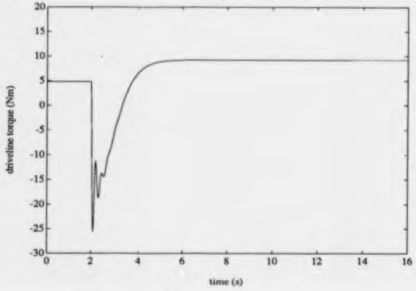


Figure 2.19: Driveline Torque Response

Chapter 3

An Overview of Nonlinear Control Theory

3.1 Introduction

The general area of nonlinear theory has grown extensively since the 1940s and early 1950s. At this time a considerable amount of literature was published on the effects of nonlinearities in position control systems - a direct result of the wartime requirements for gun and antennae position control systems. In fact, until the end of World War II developments in control theory had proceeded along purely linear lines of approach, [57]

As far as automotive control is concerned, that too has developed along linear lines for the most part of its history. Many of the elements of a powertrain are nonlinear and in order to apply classical control theory, approximate linear models have been used. The use of microcomputers to control todays advanced automotive systems such as the engine, transmission, suspension, and brakes means that nonlinear

plant characteristics, time-varying parameters and fast dynamics can increasingly be considered in the control system design. As a result, recent years have seen an increased interest in the development of nonlinear subsystem models. The increasing complexity which this brings necessitates the use of a whole wealth of nonlinear and modern control theory ideas. It is the aim of this chapter to give an overview of these methods, particularly those relevant to automotive powertrain control.

Very few nonlinear methods have been used to overcome and control the nonlinear nature of the vehicle. Subsequent analysis needs to include the stability of the resultant control scheme, obviously important in terms of safety, and the performance of the control, i.e it must satisfy both the driver's and legislative requirements. In the past, linearisation of the subsystem under consideration and use of the vast theory of linear methods has been the practice, but it is becoming increasingly important to make use of nonlinear methods.

Control of the clutch for example has been attempted using a special branch of nonlinear theory termed Sliding Mode theory, not discussed in great detail here, but has resulted in controllers which aim to improve the efficiency and response of the clutch actuation, essentially using phase plane techniques [46].

Nonlinear optimisation has been used extensively for many years in the control of engines and to a lesser extent powertrains. Optimising the performance of the control allows for better fuel economy and reduced emissions, as was discussed in Chapter 2.

It is thus important for the control engineer to appreciate the nonlinear methods which lie open to him. The motivation for studying nonlinear dynamic systems thus lies partly in their practical importance and partly in their fascination and richly varied behaviour. Nonlinear control systems are difficult to analyse and subsequent

research has often been fragmented. Since a nonlinear system is 'unpredictable', observation has to be carried out over many types and magnitudes of input in order to discover its many facets. Before one can attempt to control, one must understand the nonlinear behaviour and the problems which it causes. Some of the distinctive properties of nonlinearities are discussed in Section 3.2. The result has been that research has tended to be carried out only on specific nonlinear elements rather than generic systems. The literature can become rather overwhelming and a need was felt for some type of classification system. This grouping of various classes of nonlinearities is discussed in Section 3.3.

The techniques of analysis have also developed piecemeal, with many only applicable to certain types of nonlinearities. This is especially true in control analysis where the nonlinearities are often introduced into a 'well behaved' system. In general, nonlinear techniques have evolved along two main lines - exact and approximate. Furthermore, due to the differing preferences of engineers and mathematicians, graphical as well as analytical techniques have evolved. This is true for both the solution of nonlinear equations and for the control analysis of nonlinear systems.

The main interest with regard to the work discussed here, is the latter, for which techniques can be split into two main areas: optimisation of the control inputs to achieve the best possible control outputs; analysis to determine the stability, response and performance of a given nonlinear control system. The first can be thought of as 'how well do we want to do' and the second as 'how well can we do it'. Each has its own techniques. The first uses nonlinear programming, the second nonlinear control analysis techniques. Both are discussed in Section 3.4 and separately in Sections 3.4.2, 3.4.3. Nonlinear programming techniques are relevant to the control scheduling problem investigated in this thesis and thus are discussed

in greater detail. The subject of mathematical programming however, is vast and can only be touched upon lightly in this context.

3.2 Behaviour of Nonlinear Systems

It has been said [54] that, "whereas linearity is a specification which therefore limits the field of activity, nonlinearity is a 'nonspecification' such that the field is unbounded, linearity being only a particular part of it". In view of this, it is useful to first define linearity. An equation is said to be linear if the dependent variables or their derivatives appear only to the first power. If orders higher than the first appear or if variables appear as products of one another, or with their derivatives, the equations are nonlinear. A linear system will therefore obey the principle of superposition, and a nonlinear one may thus be described as one which does not. That is, the response to a particular input cannot be implied from the known response to some other input. As a result of this important and fundamental difficulty, no uniform approach to nonlinear analysis has been possible.

The physical behaviour which the engineer is accustomed to observing in linear systems is subject to relatively simple laws. If however, the system has nonlinear parameters to a significant degree the behaviour is essentially 'incomprehensible' from the viewpoint of linear analysis. Sometimes this behaviour can be turned to good account [55], or at the very least sufficiently approximated by linear theory.

The first characteristic of nonlinear systems stems from the fact that nonlinear systems may produce harmonics or subharmonics from a single input frequency. The application of the sum of two sine waves of different frequencies, in the nonlinear case, results in a response that may be an elaborate combination of the two. In

a situation like this, it is impossible to justify the concept of 'frequency response' without drastic modification of its definition and interpretation.

Even for a nonlinear system that does not obviously show output waveform deformation, the phenomenon of jump resonance may occur. When a saturating amplifier is incorporated into the servomechanism of Figure 3.1 the inertia of the motor will tend to smooth the clipped peaks. The servomechanism output may then resemble closely a pure sine wave but when an attempt is made at measuring its harmonic response discontinuities become apparent. The jumps which may occur in amplitude ratio and phase angle can be seen in Figure 3.2, where the dashed lines represent an unstable condition that cannot be observed in practice. The frequency at which they take place is not only dependent on the history of the test but also on the direction of approach.

Although jump resonance is normally undesirable in a control system, situations where it has been used to advantage are described in the literature, [55]. In the example above, the phenomenon can be avoided if the inputs are kept small. It is only when the system has large inputs that the phenomenon may occur, and the stability of the system is changed. Thus it is often true that nonlinear characteristics are inimical to the system's stability, which is in the hands of the initial conditions and inputs rather than the underlying system.

The nonlinear approach to rest may also be non-standard. Exhibited typically by systems with Coulomb friction, this is a further characteristic peculiar to nonlinear systems. For linear systems the motion can only increase or decrease along a exponential envelope.

Limit cycles, the name given to nonlinear oscillation of fixed frequency and amplitude, are the most frequently encountered behaviour peculiar to nonlinear systems.

The limit cycle may itself be stable or may fall into this condition if the amplitude of oscillation approaches that of the limit cycle, independent of the initial condition. In control systems, limit cycles are undesirable and will cause 'chattering', although this type of behaviour can be manipulated in sliding mode control systems, [56]. They can also be made use of in rule-based controllers such as those in anti-lock braking systems where control is applied at the peak of the μ-slip curve which characterises the tyre adhesion properties [75], [76].

Further to this, a nonlinear system may exhibit more than one equilibrium point, the position or quantity of which is not necessarily dependent on the system. The current state of a nonlinear system determines which local minimum it converges to. Unlike linear systems, nonlinearities force global and local investigations to be made. The stability of a nonlinear system in the neighbourhood of an equilibrium point does not necessarily imply global properties. There may indeed be many equilibria, some stable others not; in which case there will only be a limited region of convergence around any equilibrium point which is locally asymptotically stable. In addition, the type of behaviour whether stable, unstable oscillatory, or chaotic may depend critically on the input applied.

This is very important for optimisation studies where it is important to know how many equilibria are present and what the relationship between initial conditions and convergence is. In optimisation studies, nonlinearities will become apparent through non-repeatability of the solution from different initial conditions and through the investigation of the function space. In control systems they will become apparent through performance degradation, limit cycle excitation, or loss of stability at large inputs. In either case, it is a warning to the control engineer that more elaborate methods than those supplied by linear theory are needed.

3.3 Classification of Nonlinearities

The variety of all possible nonlinearities is essentially infinite, but nevertheless it is worthwhile to classify them into some general categories, with features which preclude or permit the application of various nonlinear methods. Classifications have been made in the past, [59], [60], but no clear diagram of division has emerged from the texts. In the face of such variety, there is a need for a simple plan of the most common possibilities. As a result, the following classification was undertaken on a simple systems level and includes such adjectives as single-valued and symmetric, both referring to character and useful in determining the method of analysis, or the use of adjectives such as intentional and separable which determine the nature of the nonlinearity's existence within the system.

Whenever a quantitative study of a physical system is undertaken it is necessary to describe the system in mathematical terms. As a starting point, the general nonlinear second order equation used in this section is,

$$\ddot{x} + \omega^2 x + f(x, \dot{x}, t) = G(t), \tag{3.1}$$

where ω is a constant, x is a dependent variable, t is time, and $G(t)$ is the forcing function which may be dependent on time. The term $f(x, \dot{x}, t)$ is a nonlinear function. Discussion is thus restricted to second-order systems. This equation was chosen since it is the most general equation given in nonlinear control theory text books, and concerns lumped parameter, deterministic systems only, which are the consideration of this thesis.

The classification, taken in a top down approach, can be most clearly illustrated by Figure 3.3. The starting point is to look at the system as a whole and then

to work through to the more detailed specific nonlinear elements. It is important to state that the hierarchy shown was chosen since it seemed to yield the most straightforward answer to classification, although many more representations are possible. The following section headings refer to levels in Figure 3.3.

3.3.1 Separable and Nonseparable Parts

Looking at a nonlinear system as a whole, two main categories become apparent, as shown in Figure 3.4. The first, 'separable', represents the common control situation in which models of devices may be obtained in one or more separable linear and nonlinear parts. In this case, the nonlinear elements are usually static functions of one or more input variables, such as relay, saturation, or hysteresis characteristics. For example, a power amplifier may be modelled as a saturation nonlinearity followed by a linear transfer function representing the dynamic behaviour, as in Figure 3.1. In this case the feedback gain H is typically unity.

In more complicated cases such as in Figure 3.4(b), the separation into nonlinear or linear parts may be less obvious if not impossible. In such cases, the nonlinearity may be in the plant, the feedback loop, i.e. a nonlinear feedback gain H, or both.

At this stage therefore, it is the existence of the nonlinearity within a control system that is being examined and not its mathematical detail. Thus, terms may be coined from control theory to describe the two classes. The first category discussed may be described as an 'open loop' configuration where the nonlinear element is in series with the linear element. The second category may be described as 'closed loop' since the nonlinearity is compact and may lie in the forward or feedback loops.

3.3.2 Intentional vs. Inherent

The nature in which the nonlinearity came to be apparent in the control system, leads to the terms 'intentional' and 'inherent'. Nonlinear functions may arise in a dynamic model either because they are intrinsic to the nature of the system, or because they were introduced, deliberately or not, by the inclusion of another device into the control system. To clarify the situation, inherent nonlinearities are those that are part of the physical model before any control is applied to the system. These can be termed 'non-control' nonlinearities and separable examples include backlash, saturation, hysteresis and nonlinear friction. Nonseparable nonlinearities are generally inherent to the physical laws governing the model dynamics. Intentional or design consequent nonlinearities include those introduced into the system for a purpose : nonlinear damping, nonlinear gain; and those that are the consequence of adding control devices : relays, limiters, and rule based controllers. These types of 'control nonlinearities' may be removed from the system by removal of the device.

In the context of the work presented here, the static and dynamic models presented in Chapter 2 have nonlinear equations which are 'inherent' to the modelling of the system dynamics, and are modelled in 'non-separable' form. Thus emphasis is given in the following levels of classification to heading which are immediately applicable to this type of system. Generally, the same classifications apply in all branches of the classification tree, and where differences occur these are made apparent.

3.3.3 Autonomous and Nonautonomous Systems

Examining the nonlinear function, rather than the control system as a whole, leads to a discussion of the differential equations which describe the system. The general

differential equation, 3.1, involves the variable t of time in the nonlinear expression $f(x, \dot{x}, t)$. The forcing function G is also a function of time. Equations explicitly involving time are described as nonautonomous. This is often however an ambiguous term. Here, nonautonomous equations are those that are not time-invariant, *i.e* time appears explicitly in the nonlinear function, the forcing function or both. An example of an equation nonautonomous in the nonlinear function would be,

$$\ddot{x} + (a + q\cos 2t)x^2 = 0, \tag{3.2}$$

for constants a and q. Of the forcing function type, an example is,

$$\ddot{x} + \omega^2 x + hx^3 = g\cos\omega t, \tag{3.3}$$

known as Duffing's equation, [54]. It represents either the mechanical system of a mass on a nonlinear spring, or an electrical system with a series combination of capacitor and inductor, one of which is nonlinear. In this case $(g\cos\omega t)$ is the driving voltage.

If time does not explicitly occur in the general system equations, *i.e.* only appears in the derivative, then the equations are described as autonomous. The differential equations of Section 2.3.4 were examples of autonomous equations. The general equation thus becomes,

$$\ddot{x} + \omega^2 x + f(x, \dot{x}) = 0\,. \tag{3.4}$$

Since the forcing function is independent of time, only internal inputs or constant input reference signals are present.

3.3.4 Piecewise Linear Functions

Considering only the nonlinear function $f(x, \dot{x}, t)$, its structure can be defined. Two classes exist : piecewise linear functions (present naturally or through the modelling of the function) and phase-plane continuous functions *i.e* those to which small and large scale linearisation can be applied.

A piecewise linear function consists of a set of linear relations which are valid in different regions, *i.e*

$$\begin{aligned} \dot{x} &= Ax_i + Bu_i \\ x_i &\in X_i, \quad \bigcup_{i=1}^{n} X_i = R^n \\ u_i &\in U \end{aligned} \tag{3.5}$$

where U is the set of all admissible control functions. A typical phase plane description of a piecewise linear system is shown in Figure 3.5. The boundaries between the different linear regions are hyperplanes in R^n. However, although the individual regions are linear, the functions may be discontinuous across the boundaries and thus nonanalytic at these points. As a whole, the system is still nonlinear but since the dynamic equations become linear and solvable in each particular region nonlinearities are often written in this way. Approximations of this type are used widely in control engineering to model actuating devices such as valves, motors or clutches, [49], [61].

The following headings are more appropriate to separable nonlinearities which can be isolated from the main plant, and for which graphical interpretations are easily obtained.

3.3.5 Continuity

The continuity of a nonlinear function can again be divided into continuity in state and state derivative. Piecewise linear functions may be continuous in state at the boundary but are unlikely to be continuous in state derivative. Phase-plane continuous equations on the other hand must be continuous in state by definition, so that linearisation can be applied. Thus, 'strong' and 'weak' nonlinearities can be defined: the former are discontinuous in state and state derivative, the latter only in derivative. This can have major implications on the applicable method of analysis.

3.3.6 Symmetry, Multi-Valuedness and Dead-Zones

Symmetry examines the graph of the nonlinearity $f(x, \dot{x}, t)$ as a whole, in order to determine whether it is odd or even, symmetric or asymmetric. The reason for inclusion of this class is that nonlinear analysis techniques are often developed solely for odd or even functions.

In the multi-valued case more than one output value occurs for each input level, hysteresis being a well known example. Systems such as these are often said to have 'memory', and although this does not introduce any essentially new difficulty it does make the state space formulation more difficult, since a prescription for switching between the various branches must be defined.

Many systems exhibit discontinuities produced by the actuator needing a certain level of input before it will respond. Such deadzones are often modelled as piecewise-linear systems, the deadzone forming the boundary. Modelling of clutch actuation is an example of this and is often handled through the use of variable structure systems, [46]. Many nonlinear analysis techniques cannot cope with deadzone nonlinearities.

3.4 Nonlinear Methods

The classification just described is sufficient to give a clear understanding of the scope and important qualities of nonlinear systems. Its aim was to separate the infinite population of all nonlinearities into some manageable categories. This not only aids understanding and places in context those of particular interest, but also helps define those techniques which are applicable to the problem in hand.

For linear constant parameter systems powerful methods are available to determine the characteristics, and the practical application of these techniques has been the paramount factor in the enormous growth of automatic control theory. Unfortunately, few general methods of analysis comparable to operational calculus have been developed to treat systems with nonlinearities. Those that are available are generally specific to a certain type of nonlinearity.

Methods of nonlinear control theory can be separated into those which aim to solve the differential equations, either approximately or exactly; those which determine the behaviour of the system relative to various control inputs; those which are interested in the stability of the control system; and those which aim to optimise the control inputs relative to predefined specifications. The first three categories, together with linearisation procedures are discussed briefly in Sections 3.4.1, 3.4.2, although not all are isolated categories and some intersection will undoubtedly appear. The most important set of techniques in the context of this thesis are the optimisation techniques of Section 3.4.3.

In recent years, attention has focused on optimising the behaviour of nonlinear systems. Limitations of classical control theory with respect to nonlinear systems led to the development of modern control engineering, which embodies control op-

timisation, [73]. The techniques of nonlinear optimisation have emerged to predict how the system parameters should be adjusted in order to bring the system to an optimum. The search for the control which attains the desired objective whilst minimising (or maximising) a defined system criterion constitutes the fundamental problem of optimisation theory. If carried out well the system is more economic, will have improved design, and the designer will have a better understanding of parameter interaction effects.

Automotive powertrains, being inherently nonlinear, lend themselves well to optimisation techniques for minimum fuel and emissions. In optimising, one gains not a controller design, but a means to knowledge of the complex underlying system, its capabilities and pitfalls and an idea of the form the control system should take. As a result, the control engineer knows what the achievable target is and the parameters needed to reach it. With this knowledge, one can then set about designing a suitable controller, testing its stability, and its performance.

Section 3.4.3 is an introduction to the methods of computing solutions to nonlinear optimisation problems. Emphasis is given to the basic methods which so often form a part of the more complex numerical algorithms available. This section will also introduce techniques used in the remaining chapters and justify their particular use. As with other nonlinear approaches, the number of available methods is vast, and only the most prominent methods can be explained here. The particular method chosen will again depend on the type of nonlinearity, the formulation of the problem, and personal choice.

It should be recognised that this section represents the distillation of work done by a great many investigators over a long period of time. The discussion presented shows the effect of continued revision and retains what is thought to be the most

common techniques from engineering and mathematics.

3.4.1 Solution of Differential Equations

The mathematical investigation of the majority of physical systems leads naturally to the solution of differential equations. However, exact analytical solutions to nonlinear differential equations are notoriously difficult to find. Only in a few cases can exact solutions be found in terms of known functions. The obvious approach has thus been to approximate and to linearise the differential equations in some form. In view of the great simplicity and tractability of linear opposed to nonlinear equations, it has been desirable to use linear approximations which can be used to provide a good first estimate of the system.

Linearisation can take three forms : small scale linearisation, large scale linearisation, and piecewise linearisation. The first linearises about one or more operating points, the second over the whole function range for a specific input to the system, and piecewise linearisation produces a set of linear equations over separate regions. If linearisation is chosen, care must be taken with regards the analytic nature of the equations. Systems can be defined as nearly linear or quasi-linear if there are no discontinuities in state or gradient, or strongly nonlinear if discontinuities in state appear. Nonlinearities are nonanalytic at discontinuities and by definition linearisation cannot be achieved at these points.

Methods of linearisation include Taylor series expansion, Equivalent Linearisation of Krylov and Bogoliubov, [57], the Describing Function Method, and a similar Harmonic Linearisation method developed by Popov and Pal'tov, [58]. Other methods such as the Perturbation Method, [57], give an approximate solution in the form of a power series for a pair of first order differential equations. The requirement of

the method that the equations are analytic however is quite severe and in such cases the method of Successive Approximations, [65], can be adapted for calculation of the solution to the differential equations. A comprehensive treatment is given in Janáč, [65]. The latter technique is based on an operational algorithm of the first order which leads to the systematic determination of linearised parameters, and is applicable to nonautonomous and autonomous, nonanalytic and analytic equations. Equivalent Linearisation developed by Krylov and Bogoliubov determines a technique for replacing a second order nonlinear differential equation with an equivalent linear one. Equivalent linearisation is the forerunner of Harmonic Linearisation where the price paid for obtaining linear differential equations is that the coefficients are not constant but are variables changing with amplitude and time. This can in fact be shown to be a special case of the dynamic describing function method of Gelb [63]. The Describing Function technique, [61], is aimed at the analysis of control systems and is the technique most used in industrial design. The fundamental idea is to represent the nonlinearity by a gain which is optimum in some sense for each 'magnitude' of input. The gain is not only dependent on the nonlinear characteristic but also on the input signal. Although an approximate method, it does have a wide range of applicability, and various attempts have been made to make it more rigorous, [61, 62].

For systems which are nonanalytic and even for those which are, simulation tools and the inherent graphical facilities they contain, can also play an invaluable role in the analysis and solution of nonlinear differential equations.

3.4.2 Stability

Techniques for the assessment of stability generally split into time and frequency domain methods. Time domain techniques generally investigate the behaviour of the system in the phase plane. Frequency domain techniques examine the behaviour of the system under test signals of varying frequency. Both give topological representations of the transient performance of the system and a large part of the efficacy of the methods are attributable to the easily envisaged graphical constructions. Although control system design specifications often relate to system time responses there are relatively few analytical techniques in the time domain for nonlinear systems. The most commonly used techniques are generally graphical in nature and mainly in the frequency domain.

In the face of discontinuities, graphical methods of solution have usually resulted. In general, graphical methods highlight changes in parameters more immediately. Most are similar to numerical methods with some geometrical construction eliminating part of the computation and as such can be useful as an explanatory tool when a nonlinear equation is first attacked.

Stability of a nonlinear system in the phase plane can be accomplished by investigation with Lyaponov's 2nd Method, [54]. The application of Lyaponov's Method requires the finding of an energy function $V(x)$. If the energy function of the system is always decreasing the system is stable. The problem is the finding of the function $V(x)$. Two systematic approaches are described in [61], but there are other problems such as not being able to treat discontinuities. This rejects both piecewise linear and strongly nonlinear systems. Secondly, it is applicable only to second order systems but is redeemed by the fact that it guarantees stability in the large.

In the frequency domain, Popov, lead by Lyaponov's work and others, showed

that the problem of stability could be treated by well-known linear frequency methods provided that the nonlinearity was bounded and not of the self-excited type. This leads to a graphical construction on the Nyquist plot for systems which are continuous single-valued and do not explicitly depend on time, although forced time varying input signals are allowed. The Describing Function technique, in the same way, lends itself to stability analysis and the detection of limit cycles.

Drawbacks of the previous two approaches can to some extent be overcome through simulation [64]. Controllers are added to nonlinear system simulations and parameters adjusted until closed-loop simulations suggest (but do not guarantee) stability and adequate performance. Methods which reduce the amount of simulation needed and provide formal methods are available [64].

3.4.3 Optimisation Methods

Problem Definition

The generalised optimum control problem may be defined as,

$$\min_{x \in R} F(x) \tag{3.6}$$

$$\begin{aligned} \text{subject to } g_i(x) &= 0, \quad i = 1, 2, \ldots, m. \\ g_i(x) &\geq 0, \quad i = m+1, \ldots, m. \end{aligned}$$

for the nonlinear objective function F and linear or nonlinear real-valued constraint functions $g_i(x)$. Within this section, all theoretical results are deduced with reference to the general nonlinear formulation, equation set 3.6, and for minimisation

problems only, since maximisation of a function $F(x)$ is equivalent to the minimisation of a function $-F(x)$. The discussion is also restricted to scalar variables x, and thus finite dimensional optimisation, although the optimisation problem can be extended in functional space to infinite dimensional dynamic problems. The wealth of optimisation methods is vast and, to limit the discussion, line search or univariate optimisation techniques are not discussed. The problem formulation can be seen more easily through graphical constructions as shown in Figure 3.6. Both one and two dimensional optimisation problems are shown, with the latter also including a constraint function. The essence of the problem is to move over the objective function map in search of the minimum function point.

3.4.4 Finite Dimensional Optimisation Methods

The distinction between finite and infinite dimensional equations relates to the system equations. The equations seen in Section 2.3.3 did not explicitly involve time, being representative of an instantaneous point in time only, and therefore were finite dimensional. Thus, in the vast majority of cases finite dimensional optimisation, (F.D.O.), is concerned with static systems. Conversely infinite dimensional optimisation, (I.D.O.) usually refers to dynamic systems in the continuous time domain, *i.e* the infinite dimensions of a scalar-valued control variable $u(t)$ are the values at each instant of time within a bounded region. This observation is upheld in the following chapters, although this is not the rule in general, and examples exist where the converse is true, [73].

Two underlying optimisation schemes are apparent for F.D.O. . Referring to the set of equations 3.6, and given that x is an n-dimensional scalar-valued vector of independent variables, the following results are necessary and sufficient conditions

for a minimum [73],

$$\frac{dF}{dx} = 0 \Leftarrow F \text{ is a minumum at } x \tag{3.7}$$

$$\frac{dF}{dx} = 0 \;\&\; \frac{d^2F}{dx^2} \geq 0 \Rightarrow F \text{ is a minimum at } x \tag{3.8}$$

A method which finds the minimum by solution of the stationary conditions is termed an indirect method of minimisation. This calculus approach is useful if the equations can be solved directly but is impractical and often impossible for nonlinear models. These methods are exact and if they fail, direct or iterative methods must be employed. Approximate techniques or direct methods iteratively search the function space for a minimum and the number of numerical techniques available is extensive. However, direct methods do fall naturally into two classes which although not completely separable do form different approaches to the problem. Search methods employ function evaluations only, whereas gradient methods require additional gradient information.

3.4.5 Search Methods

Search methods apply to both single and multivariable optimisation and attempt to reduce the value of the objective function by the use of tests near to an estimate of the solution. These then determine a search direction in which the minimum is expected to lie. Methods include the vast theory of Simplex Methods, [68], for which the Nelder-Mead algorithm [67] is an example, together with 'one at a time' searches and pattern searches [66]. In advanced methods the direction of search, rather than being kept constant, is changed to improve efficiency, *e.g* conjugate direction methods, [66], and Rosenbrock's 'rotating coordinate' method, [66]. Effective use

of these methods is limited to well-behaved functions in low dimensions. However the methods are invaluable when gradient information is impossible or difficult to compute.

3.4.6 Gradient Methods

General experience has shown that gradient methods are superior to search methods if the functions have continuous derivatives which can be calculated analytically [70]. The best search direction is evaluated from a gradient calculation at a point rather than a function evaluation. Based on the Taylor expansion,

$$F(x + \delta x) \approx F(x) + g^T \delta x + \frac{1}{2} \delta x^T H \delta x. \tag{3.9}$$

two classes exist. Optimisation methods which neglect the final term, using only the gradient vector g, are termed first order methods. Second order derivatives are then assumed small enough to be neglected. If an optimisation process utilises these second derivatives, it is termed a second order method.

First Order Methods

The first and most elementary of direct first order methods is Steepest Descent. The underlying idea is to move through the function contours via the steepest path, taking small steps of magnitude K in a direction given by $-\frac{g}{\|g\|}$, as shown in Figure 3.7. For a sufficiently small step, the movement is guaranteed steepest descent. Hence, if K is kept small a reduction in objective function is ensured. The Taylor expansion, equation 3.9, may become invalid for larger steps, as higher order terms become significant. A large step may also overshoot the minimum, but on the

other hand an unnecessarily small K will result in slow convergence. In either case, convergence to the minimum tends to be slow since as the minimum is approached the derivatives $\frac{dF}{dx}$ contained in the update procedure,

$$x^{i+1} = x^i - K\frac{dF}{dx}, \tag{3.10}$$

tend to zero. Thus, progress to the minimum from a point some distance away will start rapidly, dependent on proper scaling, but will slow down as the minimum is approached.

Improved Gradient Search: Second Order Methods and Improved Search

Improved gradient methods are generally of two types: those that improve the search by using a higher order approximation (generally known as Newton Methods), and those which improve the search direction formulation in some way.

Second order methods rely on the provision of the Hessian matrix H, of second order partial derivatives. Advanced methods make an approximation to the Hessian rather than calculate it directly (Quasi-Newton Methods [70]), but once the approximation has been made the methods reduce to the fundamental technique of Newton's Method. If a stationary point x_s is a minimum of the quadratic

$$\phi(\delta x) = g^T\delta x + \frac{1}{2}\delta x^T H \delta x \tag{3.11}$$

the stationary point satisfies,

$$H\delta x = -g, \tag{3.12}$$

and defines Newton's direction of search, [66]. Many improvements and refinements

have been made to Newton's Method and have been very successful for use with quadratic functions when the approximation to H at any point x is equal to the Hessian matrix H_{min} evaluated at the minimum. For quadratic functions the approximation is exact but for nonlinear functions the method is less attractive, and direct use of (3.12) is limited since the Hessian matrix H must be computed and inverted at every step of the iteration. Furthermore, using the fact that the Jacobian vector of a function with continuous partial derivatives is zero at the minimum leads to ,

$$\delta x^T H_{min} \delta x \geq 0, \tag{3.13}$$

for any $\delta x \neq 0$, i.e. H_{min} is positive definite. However, H is not necessarily positive definite when evaluated at a point other than the minimum. So the process may not converge. This situation is most likely to occur at some distance from the minimum. At a point close to the minimum, all sufficiently differentiable functions tend to become quadratic as third order terms in the Taylor expansion about the minimum become negligible. Complicated methods are available for manipulation of the Hessian to ensure that it is positive definite [70], but again this increases the computation time.

In practice, second order Newton Methods are preceded by first order methods until the minimum is approached, at which point a powerful method which ensures a positive definite H is used. Second order Newton methods are powerful computationally and as such are used in commercial packages to a large extent. The MATLAB optimisation toolbox used in later chapters uses Quasi-Newton methods for its constrained and unconstrained optimisation procedures, together with updating procedures to ensure positive-definiteness. However, on an implementational level the method is complex and not the ideal solution when an improved solution

procedure is needed quickly and easily.

A specialist case can arise when the objective function is the sum of squares of functions. In this case, provided the derivatives are available, the classical method due to Gauss-Newton is usually recommended. For a greater explanation and modifications made by Levenberg and others see Gill [70].

Improved search direction methods can be of both first and second order but the former shall be concentrated on here. The Steepest Descent method relied on a search direction which moved through the contours taking the steepest path at all times. As explained, this can be slow but it is simple in its implementation. A method which retains the simplicity of the Steepest Descent method, while considerably improving the speed of convergence, is the method of Conjugate Gradients. It is easily applicable to nonlinear functions and allows minimisation of $F(x)$ without the need for storage of matrices or computation of second order derivatives. Not only is the method essential when matrix factorisation is not viable - the matrices are too large or too dense, but it provides the simplest implementational changes from Steepest Descent at the minimum of effort.

The basic scheme of Conjugate Gradients (a sister method based on search strategies is available but not discussed here, see [70]) is concerned with producing n search directions which form conjugate, linearly independent vectors. Considering equation 3.9, the search direction u_1 is conjugate with respect to search direction u_2 if,

$$u_1^T H u_2 = 0, \tag{3.14}$$

where H is the positive definite Hessian matrix. Using non-interactive search vectors in this way means that the minimum of the nonlinear function can be reached in a reduced number of searches or iterations.

Several minimising procedures, for functions with continuous first and second partial derivatives, depend on the above property; the methods differ only in the way the conjugate vectors are generated and the type of gradient information used. A conjugate gradient method which employs second order gradients was derived by Davidon in 1959 [72] and improved by Fletcher and Powell in 1963 [69]. However, whilst acknowledging the power of the second order methods in which the Hessian matrix yields the information of the curvature of $F(x)$, Fletcher and Reeves developed an alternative procedure, applicable to nonlinear functions and based on conjugate gradients of first order only. This improves the economy in terms of matrix storage and calculation, both of which are significant for large problems. One iteration proceeds as follows,

Assume x^i and a search direction s^i (initially $s^0 = -g^0$). Obtain α^i such that $F(x^i + \alpha s^i)$ is a minimum with respect to the scalar α which can be shown to be positive. Set

$$x^{i+1} = x^i + \alpha s^i \tag{3.15}$$

and evaluate the gradient of F at x^{i+1} , namely g^{i+1}. Evaluate the scalar,

$$\beta^i = \frac{(g^{i+1})^T g^{i+1}}{(g^i)^T g^i} \tag{3.16}$$

such that

$$s^{i+1} = -g^{i+1} + \beta^i s^i \tag{3.17}$$

and return to the minimisation of $F(x^{i+1} + \alpha s^{i+1})$.

The result is that all search directions are mutually conjugate and that each new direction is a linear combination of previous search directions plus a gradient

evaluation at the present point. A strategy known as restarting or resetting can also be employed whereby the search direction is set to the steepest descent path after every n iterations. This ensures that linearly independent directions, which span the n-dimensional space, are produced throughout the minimisation. It thus avoids the case when the direction of the minimum remains unchanged giving $(x^{i+1} - x^i)$ no component in this direction. For nonlinear problems the step size α must be determined iteratively rather than in closed form and line search or multiple guess techniques must be used [66], albeit at the expense of increased computation.

The Fletcher-Reeves algorithm is thus easily implementable, with the minimum of change from what is usually the first-guess procedure of Steepest Descent. In terms of convergence rate, savings as much as one third reduction in the number of iterations can be made compared to Steepest Descent, [73], with little increased computational or storage cost. For these reasons the Fletcher-Reeves algorithm is seen as one of the more powerful methods of nonlinear optimisation and as such is implemented in the optimisation formulations of Chapters 4 and 5.

3.4.7 Constrained Optimisation

Figure 3.6 showed the effect of a constraint equation on an optimisation solution. In general, constraints can either take the form of bounds on the state and control optimisation variables, or as constraint equations inherent to the optimisation problem.

Constraint Equations

Nonlinear and linear constraint equations, independent of the objective function can be considered in two forms :

$$\text{i)} \quad g_i(x) = 0 \quad \text{equality constraint} \tag{3.18}$$

$$\text{ii)} \quad g_i(x) \geq 0 \quad \text{inequality constraint.} \tag{3.19}$$

Several basic approaches, incorporating them into the optimisation procedures already discussed, can be applied to the handling of constraints. Some of these are discussed below but since the field of constrained optimisation is a large and growing one, the treatment given here can only serve to introduce the basic ideas and point out some of the difficulties.

For problems with linear, quadratic, or convex objective functions and linear constraints there is a wealth of efficient procedures [70], many of which make use of the special structure of the problem. Accordingly, algorithms with nonlinear objective functions and constraints are emphasised here and it is these methods which will be discussed, although for greater implementational detail the reader should again consult Gill [70].

The most obvious way of dealing with a constraint equation is to convert the constrained optimisation problem into a new unconstrained one. The most simple solution essentially involves variable elimination. However, multivariable nonlinear functions for which this is explicitly possible are rare. The classical method of removing equality constraints has thus been through the use of Lagrange Multiplier theory. For equations of the form 3.18, the method involves finding undetermined scalars, or Lagrange Multipliers, which ensure that the constraints are satisfied at the

minimum. These are appended to the objective function to form a new minimisation problem,

$$\min H(x) = \min(F(x) + \lambda^T g(x)) \tag{3.20}$$

for Lagrange Multipliers calculated by,

$$\lambda^T = -\left[\frac{\delta F}{\delta x}\right]\left[\frac{\delta g}{\delta x}\right]^{-1}, \tag{3.21}$$

for constraint equations $g(x)$ and Lagrange Multipliers λ. The optimisation methods of Section 3.4.4 can be applied to this new objective function $H(x)$ with suitable calculation of the λ_i.

Given that conversion of inequality constraints to equality constraints, through the use of slack variables or transformations, is difficult the second major approach is either to modify an unconstrained technique to take account of the extra conditions, or to use a method designed specifically to deal with constrained optimisation problems. This first option can be done in two ways, by either taking explicit account of the constraints at each step, or by taking implicit account of the constraints through the use of global penalty functions. Those that take explicit account of the constraint at each step include methods such as direct search methods, *e.g.* modified pattern search and rotating co-ordinates, and small step gradient methods, which continually modify the search direction away from the constraint boundary and into the feasible region, *i.e.* act as reflecting barriers. Implicit methods turn the constrained optimisation process into a unconstrained one by adding sequences of penalty functions in such a way that the optimum of the new objective function tends to the constrained minimum. Methods belonging to this class are generally termed sequential unconstrained methods [66], and include exterior penalty function

methods which move into the feasible region, and interior penalty function methods which constrain movement within the feasible region. The first is generally used for equality constraints, the second for inequality constraints, although mixed penalty functions also exist which handle both simultaneously.

The second set of methods, which specifically consider the constrained optimisation problem, can be divided into two major approaches. The first, Large Step Gradient methods have two principal variants : Gradient Projection and Reduced Gradient methods [71]. Both guide the optimisation along the nonlinear constraint surface, first finding a direction of search from the current point, and second making a step to a feasible point as nearest, in a suitable sense, to the first infeasible point generated. The feasible region can be determined by active set strategies, [66]. These determine an index set which reference those constraints which are binding at a current point. Assuming an initial feasible point, operations are carried out such that the current point at the end of each iteration remains feasible. Rules are thus defined which determine which constraints should be added or deleted to the active set at each point, [66].

The alternative approach to Large Step Gradient methods is to use modified or augmented Lagrangian methods. In essence, these combine the global penalty function approach with Large Step Gradient methods. These adapted Lagrange methods either use an approximation of the objective function or approximate the Lagrange Multipliers to obtain a global penalty function.

In choosing an optimisation procedure, Lagrange Multiplier methods give a more rigid and well tried framework for implementation than the Penalty function methods, although the computational cost may be higher for more complex problems since computation of derivatives is involved. Of reflecting barrier methods, these

become inefficient when the constraints are complicated since the search direction must be continuously modified to find its way around the constraint surface. Direct search methods are preferable to small step methods since the function gradients of the small step methods may become unreliable in the constraint region. Global penalty function methods face the problem of infinite weightings as the constraints boundaries are approached and in producing a new cost function may produce peculiarities not previously there. These latter methods are more suitable for general applications, with explicit movement along the constraint boundary being avoided. The disadvantage is that the Hessian may become ill-conditioned in the process. Constrained problem-specific methods are more complicated in their implementation - needing several steps to produce a feasible point at each iteration, or need an approximation of some form. Together with the active set strategies, where care must be taken over the choice of rules used, they provide efficient methods of solution but at increased implementational cost.

Variable Bounds

For a nonlinear optimisation problem where the variables have to lie within an acceptable range, *i.e*

$$x_{\min} \leq x \leq x_{\max} \tag{3.22}$$

it can be shown [74], that a simple clipping or multi-guess approach can be applied to state-independent variables. Where the variables are state-dependent a straight forward approach is to reduce the the search direction step size and recalculate [66]. Although simple this can lead to a large number of subiterations and premature exit from the convergence test. If the initial guess to the solution is made away from the constraint boundary and the problem is well scaled this method should produce

efficient results. Otherwise slack variables must be applied to convert the bounds into equality constraints. Another method which allows the incorporation of bounds is the modified Rotating Co-Ordinate Method [66]. However, this does not allow the incorporation of equality constraints and the modifications made are not applicable to other methods of minimisation.

3.5 Computational Difficulties

Up to now there has been no mention of the inherent nature of the objective function, only that it is nonlinear. This itself suggests that the function is complex in some way, but the degree of complexity and its geometrical shape can affect the optimisation results. A classic result of nonlinearity is nonunimodality of the function, as shown in Figure 3.8. This particular function is the NO_x flow function for an automotive engine as a function of engine speed and engine torque. Nonunimodality thus infers that there is more than one minimum and that the minimum found by the optimiser may not necessarily be the global one. Few methods guarantee finding the global minimum for functions with multiple minima and so care must be taken in trusting the solution found. The presence of nonunimodality can be found, if not recognised earlier, by testing the optimisation with different starting points. If different answers are yielded when starting from different areas on the function map then a clear indicator of nonconvexity is given.

In one and two dimensional problems this can be checked by graphical inspection of the function. However, in more complex cases, especially when dealing with high order functions it may be less obvious. To some extent the choice of tolerance parameters and step size α can overcome the problem [70].

Other implementational problems which may become apparent are often due to shallowness of the function or the presence of a long narrow valley. The first are especially onerous for gradient methods and the second can cause problems of slow convergence for both gradient and search methods, . The accuracy of the solution and the rate of convergence can be improved by scaling the objective function and the supporting criteria. In such cases however, the optimiser may stop a fair distance from the true minimum and it is up to the user to determine the tolerance criterion for termination and whether this lack of sensitivity is important.

3.6 Summary

This chapter has given a brief overview of nonlinear theory and nonlinear methods in an attempt to convey the practicality of treating a problem as nonlinear rather than reverting to a linearisation process.

The exact nature of nonlinearities were discussed in Sections 3.2 and 3.3 and the resultant techniques in Section 3.4. It was stated that the control engineer is interested in stability, response and performance of the resultant control system and that optimisation of the system can provide a better control scenario. The tools needed for optimisation of these functions were discussed in some detail for both constrained and unconstrained problems. It is these methods on which the remainder of this thesis is based and how the optimisation formulation and implementation for control scheduling is performed will be expanded upon in the next chapter for a single objective function.

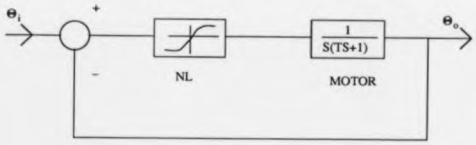


Figure 3.1: Nonlinear Servomechanism

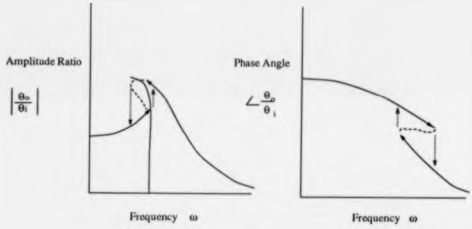


Figure 3.2: Jump Resonance

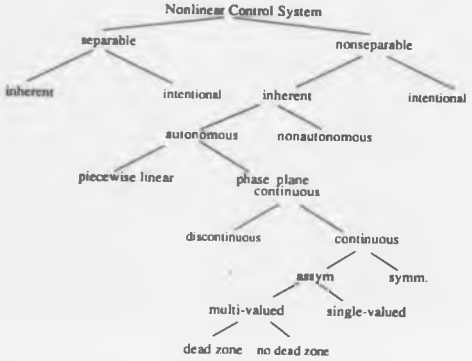
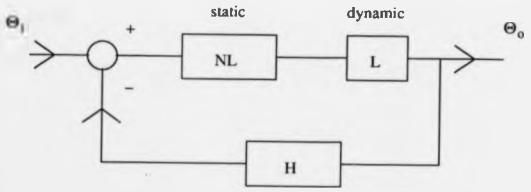
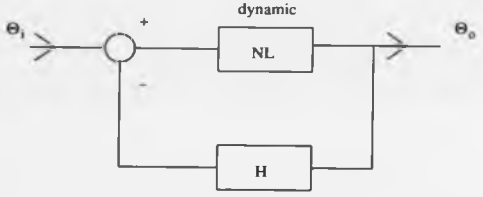


Figure 3.3: A Classification of Nonlinearities



a)



b)

Figure 3.4: Separable and Nonseparable Nonlinearities

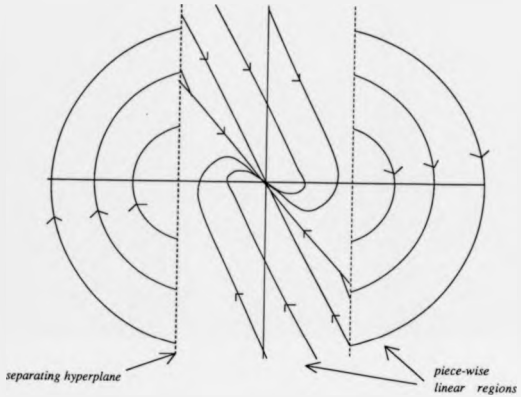


Figure 3.5: Schematic of a Piecewise Linear Phase Plane

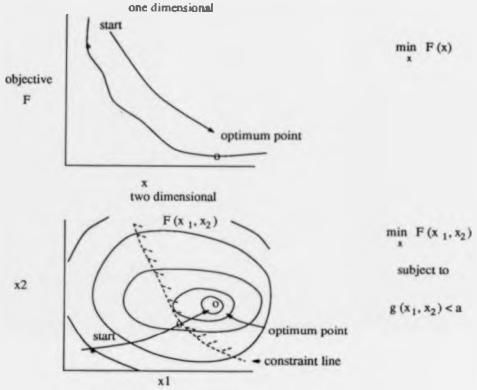


Figure 3.6: Single Objective Optimisation

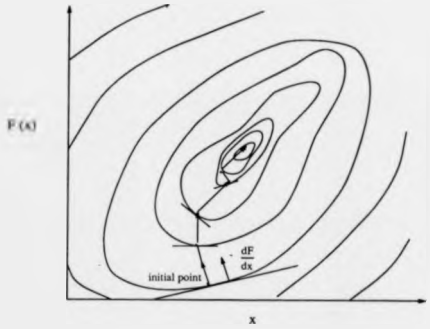


Figure 3.7: Steepest Descent Optimisation

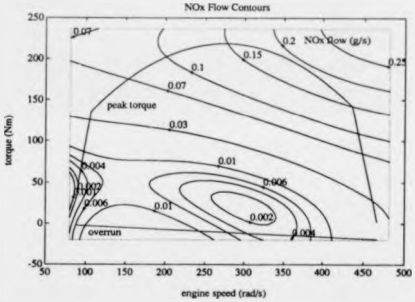


Figure 3.8: Nonlinear Function Showing Non-Convexity

Chapter 4

Single-Objective Optimisation of Powertrain Scheduling

4.1 Introduction

Minimisation of fuel economy and exhaust emissions is an important aspect of today's automotive control problems. It was noted in Chapter 2 that use of nonlinear models to predict and simulate vehicle dynamics can liaise with the powerful methods of nonlinear programming to produce a means of achieving these requirements. This chapter uses the finite-dimensional models of Chapter 2 and the optimisation theory of Chapter 3 to formulate the fuel and emission minimisation problem as a single-objective, nonlinear programming problem. This chapter thus presents the formulation and results of independent minimisations of fuel, particulate, NO_x, HC and smoke functions, to provide a basis for the more advanced optimisation studies of the following chapters.

A solution to the problem of reducing the fuel and emission output from a pow-

ertrain, can be accomplished by taking each objective separately and optimising through the manipulation of the available variables of engine torque and brake force. This traditional approach involves minimisation of one function subject to added constraints on the output of the others. This constrained optimisation approach is discussed in Chapter 5. For now, the interest lies in unconstrained optimisation of the separate fuel and emission functions. This chapter will explain the scheduling concept and how the available information is used to produce the desired results. This includes the formulation of the instantaneous driver demand, the powertrain model and its bounds, the constraints these impose on the optimisation process and a definition of the objective functions. The model, the origins of which were discussed in Chapter 2, will be discussed in an optimisation and control context.

Section 4.2 develops the concept of powertrain scheduling and breaks the process down into a number of stages. These include the formulation of the power demand from the test cycle, Section 4.2.1, the role of the power demand in the optimisation formulation, Section 4.2.2, and the associated functions of fuel and emissions, Section 4.2.3. A graphical interpretation of the optimisation process is given in section 4.3 before the implementation and results of the optimisation are discussed in Sections 4.4 and 4.5 respectively.

4.2 Powertrain Scheduling

The concept of powertrain scheduling can be most easily explained through Figure 4.1. A driver sat at the steering wheel has control over both the gear ratio selection (for a manual gearbox), the time of its application, and the amount of pressure applied to the accelerator and brake pedals. For simplicity, assume no braking is needed, and that the driver can demand a gear ratio and throttle angle. The latter

is directly related to the amount of torque produced at the flywheel and thus the power available from the engine. The gear ratio determines how much of that torque is transferred to the drive wheels in order to realise the driver's perceived velocity demand.

For a manually shifted gearbox therefore, the scheduler is quite simply the driver. He decides when to change gear, what gear to select and the angle of throttle. This is the situation as seen in Figure 4.1. The role of the electronic scheduler is to decide whether these demands are correct for minimum output of fuel and emissions, and to manipulate them if not. An onboard controller can be used to implement this decision process.

For a vehicle with an automatic gearbox the decision of which gear and when to shift is directly made by the onboard controller, into which is incorporated some scheduling knowledge. Optimisation of the schedules is thus not only a key part of the more futuristic 'drive-by-wire' vehicle control schemes, where the scheduler takes away some of the driver's control, but it is also important in today's automatic gearbox control. Therefore, any advance in the techniques of fuel and emission optimisation is of use both in current vehicles and the vehicles of the future.

The information required to accomplish the minimisation of fuel and emissions is based on the formulation of schedules. These define the area of the engine map in which the engine should operate and the level of gear ratio and brake force which should be implemented. In essence, they define instructions for the choice of gear ratio, engine torque and brake force at any point in time based on a demanded velocity and acceleration. Control over the throttle through electrical linkage is already being introduced into vehicles in Europe [78]; electronic control of the gear ratio and brake force is a fundamental part of automatic gearboxs and anti-slip

traction systems. The next logical step is to filter the demand made by the driver through to a scheduler which will modulate it to comply with a minimum fuel and emissions regime.

How the scheduling process fits into the overall control scheme is shown in Figure 4.2. The outputs of engine torque, gear ratio and brake torque are actuated by a controller which monitors the levels achieved and keeps them at their optimum value. The velocity and acceleration produced is then sensed at the road/tyre surface. If only the forward loop of the control process is implemented, an open loop control scheme results. A decision on the optimum control variables is based solely on the current demanded input from the driver. In practice, the actual vehicle velocity and acceleration sensed at the road/tyre surface may not be the demanded values, and some feedback around the controller would be needed as shown. For a more advanced control scheme, the velocity and acceleration would be fed back in an outer control loop to the scheduler input to provide information on the status of the vehicle at the end of the previous time interval. This, when compared with the current instantaneous driver demand, gives information on the nature of the vehicles motion, i.e predicts what the demand may be over the next few time intervals. Using past and present information the scheduler has knowledge of the vehicle's intended motion and is thus able to determine an action consistent with this.

This thesis deals only with the decision process of the scheduler, and not the modulation of its input. In effect, the information provided by the scheduler is open loop, that is it is based on the current input and not the history of previous demands. This decision process is the fundamental building block of the more complete control scheme shown. The information it provides is vital for progression to more advanced control schemes which have as an integral part prediction, and estimation facilities.

4.2.1 The Power Demand

Referring back to Figure 4.1, the scheduling process and thus the optimisation formulation follow a logical path. The tractive effort demanded at the road/tyre surface is derived from the driver's perceived velocity and acceleration demand. The driver chooses a gear ratio and torque level to accomplish this, which the scheduler must decide to accept or not. The correct gear ratio and torque demands are then given to the engine and gearbox controller respectively. The process of defining the optimum schedule can be shown more clearly in Figure 4.3. The process is the same as for a driver controlled vehicle, except that for theoretical studies the power demand is extracted not from an actual driver but from a legislatively defined test cycle. The specific test cycle chosen here for optimisation purposes is the European ECE-15 cycle, Figure 4.4 [8]. It is somewhat simpler than its American and Japanese counterparts, consisting of three phases of motion separated by zero velocity (idling) conditions. Overall, the cycle consists of moderate accelerations, decelerations and periods of cruise. Optimisation is only carried out over the non-zero speed portions, the remaining sections are assigned nominal values. In practice, the idling conditions would be taken care of by an idle-speed controller, incorporated into the overall vehicle controller, outside the scope of this project.

The test cycle defines a velocity and acceleration demand for the vehicle at an instantaneous point in time. Interpolation of a set of optimised points along the cycle thus produces a continuous schedule for analysis. For the ECE-15 test cycle used here, this amounts to one hundred and thirty two non-zero velocity/acceleration data pairs for optimisation. The remainder are assumed to be in a nominal idle speed, zero torque and zero gear ratio condition. The validity of this approach, which interpolates a set of finite-dimensional optimisation solutions over a transient

cycle, is discussed more fully in Chapter 6, where comparisons of optimum schedules with dynamic simulations of such are discussed.

As shown by Figure 4.3, the power demand is determined from the test cycle and provides initial values of the control and state variables. How these are defined is dependent on the formulation of the constraint equations. It is clear that in developing a formulation to compute the engine speed/torque requirements for a specific drive cycle, the flow of information is counter to the flow of energy. That is, calculation begins at the wheels and proceeds towards the engine, as shown by the dashed line in Figure 4.5. This is known as the 'inverse model approach' of Blumberg [13]. Given a velocity and acceleration demand, v and $\dot{v}$, the torque required at the driveline u_s, Figure 4.5, is known,

$$u_s = \mu(A + Bv + Cv^2) + M\dot{v} + u_3), \tag{4.1}$$

where the first term determines the road load R_l, as defined in Chapter 2, the second, the tractive force needed to accelerate the mass M of the vehicle, and the final term the initial braking force at the road/tyre surface. The total resistive force is translated through the road wheels and final drive ratio ($\mu = r_w r_d$) to produce the driveline torque u_s. The engine torque required at the flywheel to balance the driveline torque, and the associated engine speed is then defined by the equations

$$u_1 = x_2 u_s \tag{4.2}$$

$$x_1 = \frac{v}{\mu x_2} \tag{4.3}$$

Given an initial choice of engine torque u_1 and brake force u_3 the gear ratio and thus

engine speed are specifically defined. The power demand at the flywheel is then,

$$P = x_1 u_1. \tag{4.4}$$

This initial calculation of control and state variables is passed to the optimisation process which then calculates the optimal control variables.

4.2.2 The Optimisation Problem and its Formulation

Having understood the process of scheduling, the next step is to formulate the minimisation of fuel and emissions as an optimisation of the scheduling process. Given that the inherent problem is to choose acceptable engine torque and brake torque levels such that the power demanded by the driver is achieved, one must decide what information is needed in order to accomplish this.

Figure 4.3 showed the individual elements of the formulation, the first being the test cycle and the power demand, already discussed. The demand made is then translated into constraint equations on the optimisation process, and the control and state variables defined. These variables, having physical relevance, have limiting values which must be taken into account. Finally, the fuel and emissions functions must be defined mathematically.

The powertrain model is as discussed in Section 2.3.3 and the constraint equations formulated as in equations 4.2 and 4.3. An additional constraint on the optimisation process is that the brake force is applied only when the acceleration is non-zero and negative. This precludes the application of brake torque during cruise and acceleration phases. The control and state variables are bounded as in Section 2.3.3.

For each constant power demand therefore there is a optimal engine torque/speed point giving a minimum specific fuel consumption (SFC). Specific fuel consumption defines the fuel use, normalised by the power output of the engine, and has traditionally been used to meter fuel economy. The locus of all optimum SFC points on the engine map constitutes the 'economy line', Figure 4.6. Optimum fuel economy is then achieved by driving the vehicle 'up' the 'economy line', [43]. However, this formulation has two fundamental flaws. Firstly, using SFC as a measure of fuel economy limits the analysis to the positive torque region of the engine map, since in overrun conditions SFC is singular at zero engine torque. Furthermore, using a constant power demand based only on steady state constant velocity experiments takes no account of the transient nature of the vehicle and the necessity to accelerate.

A solution to the initial problem is to consider instead the minimisation of absolute fuel flow for a vehicle. Consider a vehicle travelling at a constant velocity v. Subject to equation 4.3, for a suitably small time interval δt, the engine torque needed to overcome resistances is now,

$$u_1 = \mu x_2 (R_l + u_3) \tag{4.5}$$

subject to the limits on engine speed, torque, gear ratio and brake force. From the traditional economy line approach, for a $\mathrm{SFC}(x_1, u_1)$ at a fixed $(x_1^* = x_1)$, there exists a unique value,

$$E(x_1^*) = \min_{u_1} SFC(x_1^*, u_1) = \min_{u_1} \frac{f(x_1^*, u_1)}{x_1^* u_1} \tag{4.6}$$

where $f(x_1^*, u_1)$ is fuel flow rate as defined in Section 2.3.2. The locus of all such points $(x_1, E(x_1))$ such that $x_1 \in [x_{1min}, x_{1max}]$, is the traditional economy line. For

a constant velocity, x_1^* can be replaced by 4.3 and manipulation of 4.5 leads to

$$\begin{aligned} E(x_1^*) &= \min_{u_1} \left[\frac{f(x_1^*, u_1)}{v(R_l + u_3)} \right] \\ &= \min_{u_1 | v_{constant}} \left[\frac{f(x_1^*, u_1)}{constant} \right] \\ &= \min_{u_1 | v_{constant}} f(x_1^*, u_1). \end{aligned} \tag{4.7}$$

Minimising the SFC is thus equivalent to minimising the absolute fuel flow for a vehicle constrained to lie on constant speed curves. This removes the problem of singularities at zero torque.

Power requirements associated with acceleration demands of the powertrain and vehicle inertias are not considered by this formulation. For a vehicle travelling with velocity v, at time t, the torque required to overcome the road load resistance, the brake force applied and to produce an acceleration $\dot{v}$, is defined by 4.2 with 4.3, holding for sufficiently small time intervals δt. The economy line specification now becomes,

$$\min_{u_1} \left[\frac{f(x_1^*, u_1)}{v(M\dot{v} + R_l + u_3)} \right] \tag{4.8}$$

where the denominator represents the power required to overcome resistive forces and to accelerate the vehicle inertias. For a constant acceleration demand, fixed vehicle speed, and steady brake force the denominator becomes a constant and again the formulation is equivalent to minimising absolute fuel flow.

The consequence of this interpretation is that not only can the whole of the engine map region be considered, and thus overrun conditions, but the fuel or emission optimisations will also consider dynamic operation of the vehicle to some extent. The level of approximation within the model will have important consequences on

the optimisation solution. It can be shown that the level of dynamic consideration given to the model moves the economy line over the engine map [23]. The less the approximation and the more refined the model, the more applicable the analysis to dynamic vehicle operation.

4.2.3 Fuel and Emission Functions Defined

Formulation of the absolute fuel and emission flow rate maps was discussed in Section 2.3.2. It was seen that the flow rates could be written as functions of the engine torque and engine speed. Based on engine test-bed steady state data, regression techniques were used to produce polynomial descriptions of each function. These are repeated here for convenience in Figures 4.7 to 4.11. The task is thus to find the minimum point on each surface independently, subject to the constraints already discussed.

The functions shown are based on experimental data and the use of regression techniques [77]. Thus error can be introduced through either or both of these processes. There was no control over the level of error introduced during data collection but error was controlled, using statistical analysis, during the regression process. The origin of errors within the data collection, modelling and optimisation formulations is discussed more fully in Section 4.7.

4.3 A Graphical Interpretation

Thus far the optimisation functions and the constraint equations have been defined analytically, together with the limiting bounds on the variables. Formulating this in a graphical context considerably aids the understanding of the process, and will

be expanded upon here before any results are given.

Consider minimisation of the fuel function. For control of engine torque u_1 alone *i.e.* brake torque is assumed zero, the state variables x_1 and x_2 of engine speed and gear ratio respectively are defined at specific velocity and acceleration points. The optimisation process then condenses to a two-dimensional graph, as shown in Figure 4.12. This shows an engine map as a function of engine speed and torque, onto which fuel flow contours have been projected. Also shown is a single constraint line for a specific velocity/acceleration point on the test cycle. The equation describing this line is a combination of equations 4.2 and 4.3. The role of the optimiser is to move from any initial point to the minimum point on that line. If no boundary condition is hit, the movement continues until the gradient of the constraint line lies tangential to a fuel contour. If, as shown, a boundary condition is reached before this occurs the optimiser can go no further and stops.

In this two dimensional case the optimisation thus reverts to a line search. When the braking variable is also optimised however, the process changes to that of a surface search in three dimensions. Writing the constraint equations as a single function of both engine torque and brake force produces the surface in Figure 4.13 for a specific point on the test cycle.

4.4 Optimisation Implementation

For numerical solution, the problem of minimising the required output through the manipulation of the control variables, must be stated in mathematical terms. Whilst minimising the required function, several constraints are imposed on the optimisation process. The vehicle must achieve a required speed and an instantaneous

acceleration demand, over a small time interval. The control and state variables must also lie within feasible physical bounds. The problem is thus a nonlinear minimisation subject to equality constraints which define the motion of the vehicle, and inequality constraints which define the feasibility of the chosen variables. Hence, given any fuel or emission function, its optimisation can be stated as follows,

$$\min F(x_1, u_1) \tag{4.9}$$

subject to

$$\begin{aligned} c_1 &= \mu x_1 x_2 - v, \\ c_2 &= u_1 - x_2 \mu (R_l + Mv + u_3). \end{aligned} \tag{4.10}$$

and

$$x_{1min} \leq x_1 \leq x_{1max}, \tag{4.11}$$

$$x_{2min} \leq x_2 \leq x_{2max}, \tag{4.12}$$

$$u_{1min} \leq u_1 \leq u_{1max}, \tag{4.13}$$

$$u_{3min} \leq u_3 \leq u_{3max}. \tag{4.14}$$

The function $F(x_1, u_1)$ is a general polynomial defining either a fuel or an emission function. The constraint equations c_1 and c_2 are such that they will be zero at the minimum. All model parameters are for a 2.5l turbo-charged diesel-engined vehicle as described in Chapter 2, with coefficients as in Table 2.1.

Chapter 3 discussed the variety of optimisation methods available for nonlinear programming problems. Despite the wide variety of methods, all formulations gen-

erally adhere to the logical flow of information shown in Figure 4.14. Search routines were discussed in Chapter 3 and the logical flow of such shown in Figure 4.15. Upon choice of initial conditions, the resultant constraint equations are calculated and provision taken to satisfy them. For nonlinear functions this choice is extremely important. A direction of search is then calculated together with a decision on the size of step taken in that direction. The state equations are then recalculated for the new control variables and feasibility and convergence checked. If progress has been made the process is repeated until some convergence or terminal criteria is satisfied.

All single-objective optimisations and associated calculations were undertaken in MATLAB through the use of a package of user-written m-files. A discussion of the MATLAB coding can be found in Appendix A.

Initial conditions

The two constraint equations define four unknown initial conditions. Thus, some manual decision must be made on at least two of these variables. The control variables are engine torque u_1, and brake force u_3 and specifically define x_1 and x_3, the engine speed and gear ratio respectively. An initial choice of u_1 and u_3 will thus define an initial position on the engine map and an initial torque translated to the tyre road interface.

How these variables are chosen can be highly crucial to the success of the optimisation, as discussed in Chapter 3. In order to choose feasible initial conditions, i.e those which obey the constraint equations, both graphical and numerical knowledge is needed.

Graphical knowledge of the function under consideration will highlight bad areas of the map in which to start - saddle points or shallow minima hinder the optimiser's

progress. Knowledge of the feasibility of a choice of control variable and its impact on the state variables requires a numerical routine which conveys the limits of choice to the user. Such a routine has been developed.

Given the required vehicle speed and acceleration demand and the limiting values of the variables, the program, written as a MATLAB m-file, detects the intersection of the constraint curve, as a function of engine torque, with the state variable constraints. The maximum value of engine torque is approximated to a constant upper bound, rather than the nonlinear peak torque curve, for ease. An approximate range of engine torque values producing feasible engine speed and gear ratios is then reported to the user, the intersection of ranges giving the total range of choice, Figure 4.16. On deceleration, this is slightly complicated by the introduction of a brake force variable (brake force is constrained to be zero during acceleration and cruise). The choice can no longer be viewed on a 2-dimensional plot and the feasible region must be searched for in 3 - dimensional space. However, the result of the addition of the brake variable is to move the two dimensional constraint line vertically over the engine map. The engine modelled is such that the grading of the engine is not high enough to achieve the levels of deceleration through engine braking alone. Without braking applied therefore, the constraint lines tend to lie below the peak overrun curve. A routine can be written which detects the minimum value of brake force needed to bring the constraint line into the working range of the engine. This two dimensional representation allows a quick first guess to the problem and, although a simple approximation, can have enormous time saving benefits on a first attack of the problem. Since a choice of control variables which produce infeasible state variables does not guarantee convergence to a minimum within the feasible function space, it is important that the choice can be made easily and effectively.

Constraint Handling

Two different types of constraints must be handled by the optimisation process. The variable bounds defined as inequality constraints must themselves be handled in two different ways, dependent on whether they are state dependent or independent. State independent constraints can quite simply be handled by a process of clipping [74]. Upon calculation of the control variables a routine is called which clips the control variables to their maximum or minimum values if a violation takes place. This method is simple and effective. State dependent constraints are slightly more difficult to handle since manipulation in this way would lead to a change in precalculated control variables.

As mentioned in Chapter 3, state dependent inequality constraints can be handled by multiple guess techniques. This can also be applied to nonconvergence of the routine, i.e the objective function increases on an iteration. If a control variable is chosen which produces either infeasible state variables or an increase in objective function, the optimisation is temporarily halted whilst a step back is taken. A step of reduced size is then taken in the calculated search direction. Here, the reduction is half the original choice. *i.e* ($\frac{\Delta}{2}$).

Equality constraints are handled most straightforwardly through the use of Lagrange Multipliers. The objective function is altered to take account of the constraint equations so that, at the minimum, the constraints are satisfied. The new objective function is a weighted function of cost and constraints such that,

$$H(x_1, u_1) = F(x_1, u_1) + \lambda_1 c_1 + \lambda_2 c_2 \tag{4.15}$$

where λ_1 and λ_2 are the calculated Lagrange Multipliers and c_1 and c_2 are the

constraints equations, 4.10. Calculation of the Lagrange Multipliers through differentiation of the cost function was shown in Chapter 3.

Method of Search - Conjugate Gradients

At the heart of the optimisation routine is the search algorithm. Various approaches to search were discussed in Chapter 3, the conclusion of which was that the use of conjugate gradients, and more specifically the Fletcher-Reeves algorithm [69], offered the simplist and most economic way of producing results at an acceptable convergence rate. It is this routine which is implemented here. Gradient information is readily available which removes the necessity of function search routines. The added complication of matrix inversion and storage needed for second order routines was felt to outweigh the benefits of faster convergence. Thus, a first order conjugate gradient routine as described by equations, 3.15, 3.16, 3.17, is implemented.

The added process of resetting, Section 3.4.6, is also used and shown more clearly in routine 'steda.m', Appendix A. All routines were initially implemented with a steepest descent routine. This was found to be slow and troublesome if initial conditions were close to the minimum. Progression to Fletcher-Reeves provides an acceptable performance at minimum implementational cost.

Parameter Influence

The choice of algorithm dictates the nature of the optimisation parameters which in turn affect the optimisation performance. These choices include the initial step size and its reduction upon infeasibility, the convergence and termination criteria.

The choice of step size will affect the rate of convergence of the algorithm : too large a step and the optimiser may overshoot the minimum and thus need

many subiterations to reduce to one which produces convergence; too small and the number of iterations needed to reach the minimum will increase substantially. Determination of a correct initial step size and step reduction ratio are reliant on experimentation and the user's knowledge of the problem. Suitable values here were found to be a step of $K = 1$ and a reduction ratio of one half.

Termination and convergence criteria also have important consequences on the optimisation. The convergence criterion determines when convergence has been achieved. The termination criteria stops the optimisation process when convergence has not been achieved within a certain number of iterations.

The former can be handled in a number of ways. Convergence criteria can be either sensitive to the movement of the control variables, to the cost function or to the Lagrange Multipiers. For a well scaled problem all three should produce similar results. Here a convergence criterion dependent on the movement of the control variables was chosen, but others were tested and all showed similar results. The specific test is an interpretation of the general convergence test but is based on steepest descent movement. The optimisation proceeds while

$$K \parallel g_c \parallel > tol \parallel u_c \parallel \tag{4.16}$$

for g_c the current gradient evaluation and u_c the current control vector. The algorithm thus checks the current movement of the control variables against a function of its current position. Choice of the tolerance parameter will affect the sensitivity of the optimiser as it approaches the minimum. Too lax a value and the routine will stop way short of the true minimum, this is especially a problem for shallow functions where the change is gradient is minimal. Too high a value and the optimiser

will continue even for very small rewards. This choice is thus critical to the accuracy of the solution and the rate at which it is achieved. The tolerance parameter used here has a value of 1.0e-0.8.

Choice of termination parameter involves a representation of the maximum number of iterations made before convergence is seen not to have been achieved. A value of $100 * n$ is usually seen to be reasonable, where n is the number of control parameters [70]. The number chosen was therefore set at 200 iterations.

Problem Scaling

The varying ranges of state and control variables suggest that the problem should be scaled to produce a uniform function space for optimisation. Thus, control and state variable coefficients, and functions there of, were scaled prior to optimisation. Scaling of the fuel and emission function definitions was also undertaken as part of their regression.

4.5 Optimisation Results

Having formulated the problem and discussed some of the implementational problems, the results of the optimisation will now be discussed. This section introduces results for all five objective functions independently, and discusses some of the inherent features and associated problems of each.

4.5.1 Optimisation of Fuel Flow

Optimisation of fuel economy has perhaps been the most frequently optimised parameter in the literature to date. Every driver is interested in the fuel economy of

his vehicle, for curiosity if not economic reasons.

The fuel flow objective function is a fourth order polynomial in engine speed and torque of the type equation 2.3, with coefficients as in Table 2.2. Finite dimensional optimisation of the 132 non-zero velocity data points of the ECE-15 cycle leads to the solution points as shown in Figure 4.17. This is an engine map for which the fuel flow contours are shown in grams per second.

The unconstrained minimum lies in the low engine speed/high engine torque region relative to the constraint lines, Figure 4.20 and the locus of solution points lies in this region. Most points lie on the first portion of the peak torque curve - further progression of the optimiser is halted by this constraint. In the upper region the solution has moved away from the peak torque boundary into higher engine speed areas. Scattering of the points can occur in this area due to the shallowness of the function. Despite being convex throughout most of the engine map the shallowness of the function with respect to the constraint lines in this region can lead to a lack of sensitivity, Figure 4.20. Increasing the tolerance of the optimisation algorithm can generally cure this, but since movement along the constraint line results in a negligible difference in fuel flow the problem is not too great. The exact optimisation solution may become more dependent on the initial start point but this simply indicates that, at high speed and acceleration points, a wide range of gear ratios and engine torques can be chosen to produce practically the same level of fuel economy.

More information can be gained by looking at the actual schedules themselves, Figure 4.18. The three phases of the test cycle are separated by nominal zero torque, idle engine speed, zero gear ratio portions which are not optimised as already discussed.

The engine speed, as expected, is low throughout most of the cycle, staying mainly at idle speed, but rising to higher values during acceleration manoeuvres. The engine speed increases only slightly during deceleration, if at all. At the same time the gear ratio drops quickly giving some engine braking assistance together with the brake force which is applied. The brake force is applied consistently throughout the deceleration manoeuvres, being released when the vehicle is brought to a halt.

Engine torque along with gear ratio rises during acceleration portions and plateaus at low values during cruise periods. During deceleration, the engine torque moves slightly into overrun. The gear ratios produced are high and well into overdrive, as one would expect for good fuel economy, with the maximum gear ratio being reached during the highest constant velocity phase.

Examining the particular flow graphs, Figure 4.19, shows that fuel usage is at a high during acceleration periods with cruise and deceleration portions using very little fuel in comparison. This would also seem to be the case for the particulate output for the fuel optimum gear-torque points. Particulate output rises rapidly during acceleration, dropping very low during cruise phases, and rising slightly again in deceleration. Nitrogen oxides are consistently low in relative terms except at very high acceleration points. This is consistent with the nature of NO_x output, it being at a premium during high temperatures as experienced during high acceleration/high load manoeuvres. The remaining emissions follow the particulate pattern being high during acceleration and and low during cruise and deceleration. Also of interest to note is that the nominal values chosen for the zero velocity periods is suited to low fuel production.

The absolute values of fuel and emission functions over the cycle (excluding zero speed portions) are detailed in Table 4.1. More will be said of these when the

emission results have been discussed and comparisons can be made.

The overall recommendation for low fuel flow therefore seems to be that engine torque should be kept high whilst high overdrive gear ratios should be used. Braking should not reach more than 1500 Newtons. The effect is to keep engine speed very low, practically at its idle speed value. In practice, this would need some added control to overcome engine roughness at this low speed and to provide acceptable engine performance.

4.5.2 Optimisation of Particulate Flow

The particulate objective function is a fifth order polynomial in engine speed and torque with coefficients as in Table 2.2. The contour map of this flow function is shown in Figure 4.21, together with the optimisation solution points. The optimiser moves through the valley, at the bottom of which are two local minima. The power constraint curves however, never pass through the upper minimum, but move towards it. Again the engine speed torque points lie in the bottom left hand corner of the engine map but this time are not reaching peak torque values. The difference can also be seen clearly in the schedules of Figure 4.22. Acceleration and deceleration torque levels are generally very similar: acceleration torques are a little lower, deceleration torques a little higher, than for the fuel optimum case. The gear ratio levels have reduced slightly during constant velocities and the engine speed reaches higher speeds over the whole cycle. Constant velocity phases generally have an engine speed around $100 rad/s$ compared with $89 rads/s$ for the fuel optimisation. The brake torque pattern is very similar.

Examining the flow graphs, the effect of the higher torque during deceleration has been to bring the particulate flow output down in these regions, Figure 4.23.

The lower torque during acceleration has brought the low velocity - low acceleration output down too. The effect on the other emissions has been to substantially increase the amount of NO_x output over the whole cycle; substantially reduce the amount of HC output during acceleration and deceleration (now lower than its cruise phase output); and to reduce throughout the level of smoke.

The specific cycle figures are shown in Table 4.1, the changes producing a 22% reduction in particulate output at the expense of an increase in fuel, an increase in NO_x, but a decrease in HC and smoke. The results indicate that only slight movement across the engine map plain causes significant changes in the fuel and emissions output. The complex relationships that are becoming apparent therefore are very difficult to detect through manual analysis alone which is where optimisation becomes so important.

4.5.3 Optimisation of Nitrogen Oxides

The NO_x flow contours are shown on the engine map of Figure 4.24, the function being a fourth order polynomial in engine speed and torque. The function is certainly nonconvex, especially in the low torque region where most of the optimisation is carried out due to the range of the constraint lines. Again two local minima are apparent, the left most one being clipped by the first portion of the peak torque curve. However, the saddle point which exists between them is a problem which can hinder the optimisation process. Careful choice of initial conditions can alleviate this problem.

As seen in the diagram, most points have found the global minimum in the high engine speed/low torque region. Some have found it more beneficial however to remain close to the idle speed/peak torque curve. This large movement across

the engine map does make the engine torque and speed profiles more erratic, as shown in Figure 4.25. The engine speed especially is seen to hunt between the two regions and acceptability of this would obviously have to be taken into account under driveability considerations. It may be that in practice some concession on the NO_x output would be made to produce a smoother engine noise, a rough noise obviously uncomfortable for the driver.

In comparison with the previous optimisations the torque levels are very low, the maximum being around 70 Nm. The low torque/high engine speed profiles are accompanied by the general pattern of gear ratio except notably at the highest constant velocity phase section, (140-160s). This is due to the consistently high engine speed, but low torque output, over the period. The higher vehicle speed pushes the constraint line into the the local minimum rather than below it (where most of the cruise constraint lines lie), the result being that it can find a lower NO_x value at higher engine speeds.

The reduction in the overall NO_x output level for its specific optimisation is clearly seen in Figure 4.26. The flow reaches no higher than a maximum of 0.006g/s, compared to 0.02g/s for other optimisations. Reduction is most dramatic during the acceleration and deceleration phases, the high load portions most inimical to NO_x output. The absolute values are shown in Table 4.1. A 46% reduction over best levels as yet seen for NO_x output, is however, accompanied by significant increases in the other functions, fuel particulate and HC approximately doubling in value. This is reflected in the remaining flow functions of Figure 4.26, HC output being particularly bad over all but cruise periods. It is interesting to note that the same nominal resting values which were good for fuel output are now extremely bad for NO_x flow rate.

The compromises for minimisation of NO_x alone therefore are great. It is unlikely that the benefits of reduced NO_x through the use of these schedules would be acceptable to either the manufacturer or the consumer when the cost to both fuel economy and other emissions is so great. Minimising solely for nitrogen oxides destroys the optimisation of other functions. It is in these situations that the multiple conflicting interactions need to be inherent in the optimisation procedure. Acceptable scheduling cannot be accommodated by single-objective optimisation.

4.5.4 Optimisation of Hydrocarbons

An engine map onto which HC flow contours have been projected is shown in Figure 4.27. The function is of the fourth order and, as can be seen, quite complex. There are two local minima within the working range of the engine so it is important to have good graphical knowledge of the function in order to choose the initial conditions.

The single-objective solutions are plotted on the graph and lie in the low engine speed / high torque portion of the constraint equations, although the optimiser has stopped short of the peak torque curve. Despite the second minimum, and the very shallow saddle point which lies between them, the optimiser has been able to find the global minimum of the function. Transferring these points into schedules, Figure 4.28, the position of the minimum is reflected in the engine torque and speed trajectories. The engine speed reaches a maximum of around 140 *rad/s* during the acceleration phases, staying mostly around the 100 *rad/s* level during cruise manoeuvres. Gear ratio is pushed high during acceleration phases although it doesn't reach its maximum bound.

How this affects the fuel and emission output is shown in Figure 4.29. The level of HC flow rate is well down on other optimisations, the overall value being

shown in Table 4.1. The optimised portions of the cycle produce a lower level of HC output rate than the zero-speed portions for which nominal values are chosen. This reflects the inimical nature of idling conditions to HC output and emphasizes the need for control also during these phases. Acceleration and deceleration phases produce the lowest flow rates with cruise periods producing higher outputs. How the optimisation of HC affects the other flow rates can be seen both in Figure 4.29 and Table 4.1. Deceleration phases of the optimised HC schedule have slightly increased the fuel flow. For particulate and smoke also, the output rate has increased during the deceleration phases. This is logical since fuel, particulate, HC and smoke have optimum solution points in the low engine speed, high torque region and will thus be sensitive to movement in this area. Sensitivity is greatest during deceleration phases since brake force is now manipulated also, forcing a search in three dimensions rather than just two

The NO_x output is particularly bad compared to its optimised value. All but one manoeuvre on the test cycle produces output above the nominal zero velocity output value. The optimised NO_x schedule produced values well below this level at all points on the cycle. This adds to the increasing feeling that whatever is good for NO_x is bad for everything else and whatever is bad for NO_x is good for everything else. Most conflict will occur then, when NO_x and other emissions are needed to be optimised simultaneously.

4.5.5 Smoke Optimisation

The last of the functions to be optimised is that of smoke, the contours of which are shown in Figure 4.30, together with the optimal solution points. The smoke function is a fifth order function, the higher order reflecting its complexity. The function has

three individual minima, the third not being reached by any of the constraint equations. The remaining two lie in a very shallow valley running horizontally along the engine map, almost parallel to the line of the constraint equations. This therefore, can cause a lack of sensitivity in the optimiser and produce greater dependence on initial conditions. However, since all initial conditions were chosen to start essentially in the same region of the engine map it would seem that the constraint lines to some extent only run through one of the minima by nature of their shape.

The schedules, shown in Figure 4.31, reflect the jump across the engine map, this being most clearly seen in the engine speed trajectory. Deceleration points are associated with the higher engine speed solutions, which in turn are accompanied by lower torque values and transients in gear ratio. The gear ratio is seen to dip momentarily on deceleration, signaling a change down to decelerate in a overrun or lower engine torque condition.

The output level of smoke, Figure 4.32, reflects the movement of the optimiser across the two very low minima. The modelling error of the smoke function, as discussed in Section 2.3.2, is reflected in the regions where zero smoke output has been detected, these occurring mainly during cruise but also during deceleration phases. It is noticeable again that the nominal idle speed-zero torque unoptimised regions are not particularly good for smoke output, although not as high as the output rate during accelerations.

The effect on other functions is variable. The jumps to higher engine speeds, on deceleration, are extremely bad for HC output rate but advantageous to NO_x since the jump overshoots the saddle point region on the NO_x contour map. Fuel and particulate output rates worsen for higher engine speeds but are of similar measure for solution points in the lower engine speed region. These increases are reflected in

the overall output values, Table 4.1.

4.5.6 Single-Objective Optimisation Solutions : Legislative Limits

All single-objective results discussed separately so far are listed in Table 4.1. However, it is the interaction of these results which is of importance to a designer wishing to extract a schedule which is expedient to all functions. The comparison of schedules on a second by second basis has already been discussed. The total emission and fuel figures per cycle are also important since they give an indication of how far away from the legislative limits the optimiser is moving, and also to compare on a whole cycle basis the merits of each optimisation.

The most immediate trade-off is that between NO_x and the other functions. The minimised NO_x output rate is well below the NO_x output levels of other optimisations. However, it carries with it the burden of increasing fuel by 100% of its minimised value. Even more damaging, particulates, HC and smoke are tripled. A NO_x optimised schedule therefore, would not be a good compromise solution to select, although referring to the last column of Table 4.1, the levels are close to the example legislative limits.

In order to provide a limit for each function, American F.T.P. limits for the E.P.A cycle have been chosen, [8]. These relate to a vehicle of $1800kg$ subject to current commercial vehicle limits. The limits, using conversion factors, are changed from EPA gram/mile values into equivalent European ECE gram/test figures, scaled to one quarter of the cycle. A summary of the US Federal limits, their conversion into European equivalent ones, and the European limits is given in Table 4.2. European

limits specific to the ECE-15 cycle have an amalgamated limit for NO_x and HC. In order to use this figure some decision would have to be made over the relative level of HC and NO_x which should be attributed to each. Without available expert knowledge this was seen to be another route through which errors could enter. Also there is no figure specified for particulate output. For those which are identified the American figures are comparable to the European ones. No figures were available for fuel and smoke for either legislative rulings. The optimisation figures shown in Table 4.1 do not take into account the nominal output during the resting phases, whereas the limit values do.

All single-objective cycle minima, to which limits are attached, achieve the legislative values. Particulate and HC are well under these levels, NO_x comes closer to its limit. However, whilst achieving the limit for the specific optimised function the other emissions can achieve values very close to their limits, *e.g.* particulate, HC and smoke minimisations produce NO_x values very close to its limit. Schedules, therefore, on which no optimisation had been carried out, would undoubtedly cross their limit values.

It was noted earlier that the minima of fuel, particulate, HC and smoke lie in the same region of the engine map. One may be tempted to conclude that a scheduler that passed roughly through this area would be acceptable. However, there are subtleties which must be taken into account. For example, minimisation of fuel produces a 30% increase in particulates over the particulate optimised solution whilst forcing a 34% decrease on the NO_x level. This slight movement has produced significant changes in the function values. The exact placement of the schedule therefore would be a very difficult decision making process and would entail a large degree of trial and error.

The manual analysis and attempted comparison of all five optimisation results has not been extensive here, but has given a flavour of the immense task facing the designer of a scheduler. Not only can complex interactions be hidden, found only upon inspection of the schedules on a second by second basis but those that are found can lead to extensive experimentation which is time-consuming and tedious.

The worth of single-objective optimisation lies not in the extraction of specific schedules but in the ability to give a first pass solution. One can highlight the interactions which are immediate, *e.g* that seen between NO_x and the other functions, and gain experience with the optimisation process before moving to more advanced methods. The process develops insight into the best starting points for the optimiser, the levels of tolerance for the optimisation process - some functions may need a greater degree of sensitivity, and most importantly, a feel of whether the optimisation is worth while. If an optimised function cannot achieve anywhere near its legislative limit through the process of scheduling, then the technique is inappropriate. If the technique is worthwhile then it is necessary to proceed to look at the compromises in more detail.

What has become evident is that optimisation of the engine torque and gear ratio variables is a worthwhile process. However, single-objective optimisation is not an adequate way of attacking the problem. It has given some experience in the technique, a feel for the individual functions and insight into some of the more obvious relationships between them. For realistic use however, more advanced optimisation methods are needed - those that automate the compromise process to a large extent, removing both the manual analysis and the time-consumption associated with it.

4.6 Driveability

Driveability is a term used to describe how well the vehicle handles. It is a term which has no precise definition and as such is very difficult to quantify. However, considerations which can be included under a driveability heading are engine roughness, engine transients, 'time-to-shift', and shift values.

There are many driveability constraints which are not included in the optimisation formulation. For instance gear ratios are usually restricted to those which do not force the engine speed down below about 150 *rad/s*, since below this region the engine operation becomes rough. To accomplish this on a CVT vehicle the transmission is declutched from a high regime to a low regime for low vehicle speeds [43]. This is not modelled here but could be quite easily included.

Also, the transient nature of the engine speed movement needs to be considered if the schedule is to be implemented. It is normal to judge engine driveability in terms of speed of response, flywheel acceleration, and the resultant driveline oscillations.

Also restricted is the transient nature of the gear shifting. Hunting, or continual shifting amongst gears, can appear for functions such as NO_x and smoke when the optimiser tends to jump between two different minima. In this case, some smoothing would have to take place either in the optimal schedule or in the controller. The 'time-to-shift' is also important. It is assumed here that the movement in gear ratio, over one second time intervals, is achievable. It is generally accepted that an experienced driver can change gear in under one second.

The general feeling is that the schedules produced in this chapter are driveable [79]. A greater test of this, is to simulate the schedules under a dynamic model. This is left to later chapters when better scheduling techniques have been introduced.

4.7 Possible Sources of Error

Before leaving this section on single-objective optimisation, it is worth investigating some of the possible sources of error in the problem formulation. The optimisation process is intrinsically based on empirical data and since one is solving nonlinear equations which can only be solved iteratively, inaccuracy will ultimately creep into the formulation. A more detailed look at the source and variety of these errors is shown in Figure 4.33. There are three basic sources as mentioned: the data collection, the feasibility of the empirical and phenomenological models, and the optimisation process.

No control over the accuracy of the data collection was possible. It is generally thought that emission data collected from an engine test bed is inaccurate to some extent. Results can differ with variations in test cycle driver, test bed operators and the condition of the vehicle. Tests are normally done in triplicate, which, if carried out in a short space of time, are generally comparable to a high degree. Uncertainty will continue to appear however, until better methods of collection and standardisation are developed.

Test bed data, in this work, was collected in the form of continuous traces on a chart recorder. However, it is becoming more usual to use computer datalogging techniques. Further error can thus be introduced when the continuous traces are digitised into specific engine operating points for regression purposes. The errors at this point are thus taken into the regression process itself.

The error in the empirical models of fuel and emission flow which result are controlled through statistical analysis. As was seen in Section 2.3.2, problems can still result when the functions are very complex.

The phenomenological models are reliant on the model parameters supplied, *e.g* vehicle mass and inertia values, variable bounds, *etc.*. These values were unavailable for the particular engine-vehicle test-bed combination and information was used from previous studies. The level of approximation in the model will also significantly affect the resultant solutions. The extremely dynamic nature of the vehicle is approximated by interpolation of steady state operating points. This will affect driveability of the schedules and their applicability to normal driving conditions.

The optimisation process itself, since the models are nonlinear, cannot be exact. Iterative routines are needed for solution, this in itself being dependent on the sensitivity of the optimiser and termination parameters chosen. These errors can be minimised however through experience and testing.

The possible source of error is enormous. The dominant source however, is the empirical data used, particularly the engine test-bed fuel and exhaust emission maps, whilst very little error will be contained in the optimisation process. It has been ensured that where control of errors can be instigated, this has been done effectively, and while some of the conclusions drawn from the numerical results will be subject to the data error, the applicability of the optimisation methods is unchanged.

4.8 Summary

This chapter has taken the steady state models developed in Chapter 2 and the nonlinear techniques of Chapter 3 and transferred the scheduling environment into an optimisation process. The scheduling process has been explained in depth, the models put into the context of fuel and emission minimisations, and the variables related to the physical process of selecting appropriate operating conditions. The

result has been the separate optimisation of fuel, particulates, NO_x, HC and smoke flow rates, the outcome of which has been discussed.

The conclusion to arise from this is that single-objective optimisation is necessary as a first pass solution. It allows the user to judge the worth of the method and to gain experience with it. However, it is impractical as a method of developing schedules which must contain beneficial elements for all fuel and emission functions. Techniques are needed which allow the optimisation of all functions simultaneously. The independent complexities of each function will then be inherent in the optimisation process, and can be used to produce a trade-off solution.

(g)	Fuel[a]	Partic	NO_x	HC	Smoke	limit
Fuel[b]	27.59	33.33	66.45	40.03	38.78	
Partic	0.097	0.075	0.237	0.094	0.096	0.34
NO_x	0.842	1.28	0.452	1.023	1.228	1.3
HC	0.160	0.144	0.386	0.127	0.203	1.0
Smoke	0.040	0.023	0.075	0.034	0.022	

[a] vertical columns denote functions optimisations
[b] horizontal columns denote cycle predictions

Table 4.1: Single-Objective Comparisons

US Federal			European Equivalent			European	
1975 FTP Test (4000-5999lb)						EEC 88/76/EEC (upto 2500kg)	
1988 models as of 1987						2.0l and above, as of Oct '92	
g/mile			g/test			g/test	
HC	NO_x	P	HC	NO_x	P	HC + NO_x	P
0.8	1.7	0.45	4.0	5.1	1.35	8.0	

Table 4.2: European, US Federal and Equivalent European Limits

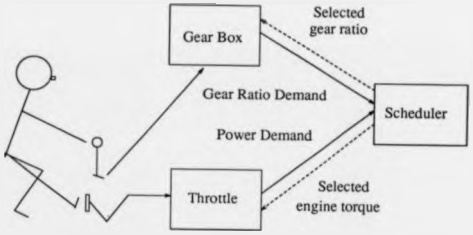


Figure 4.1: The Role of the Scheduler

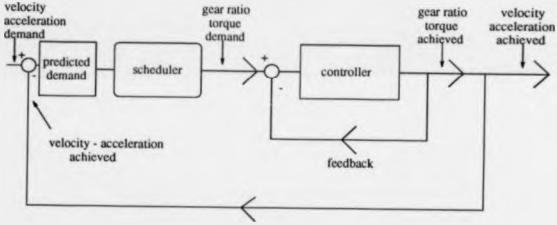


Figure 4.2: The Interaction of the Scheduler and Controller

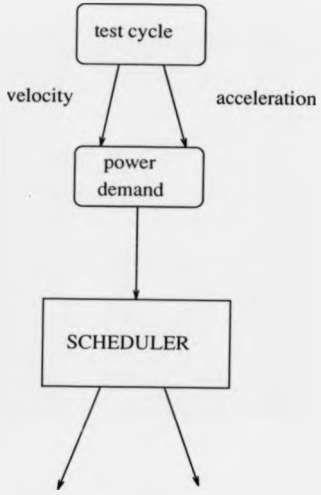


Figure 4.3: The Scheduling Process

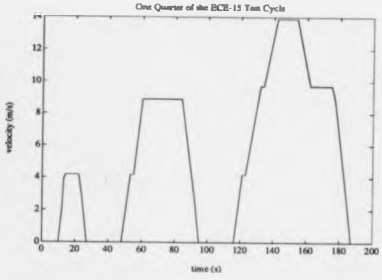


Figure 4.4: The European ECE-15 Test Cycle

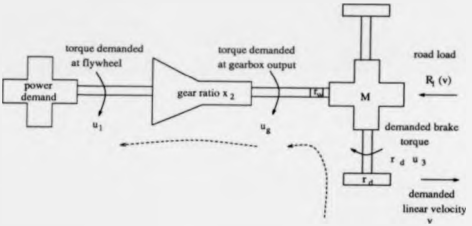


Figure 4.5: The Inverse Model Approach

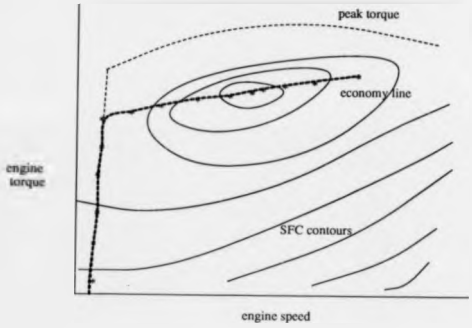


Figure 4.6: Schematic of Economy Line Formulation of Minimum SFC

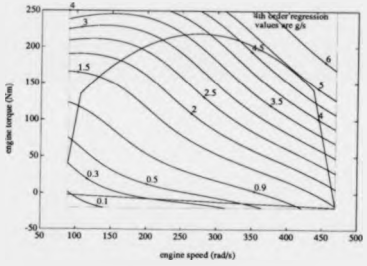


Figure 4.7: Fuel Contour Map

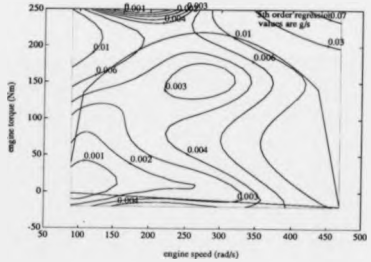


Figure 4.8: Particulate Contour Map

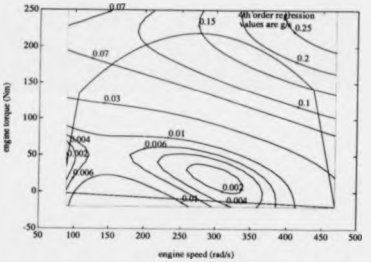


Figure 4.9: NO_x Contour Map

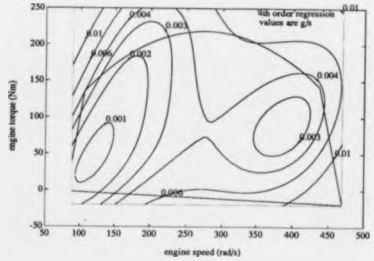


Figure 4.10: H.C Contour map

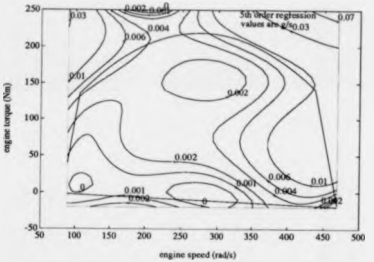


Figure 4.11: Smoke Contour Map

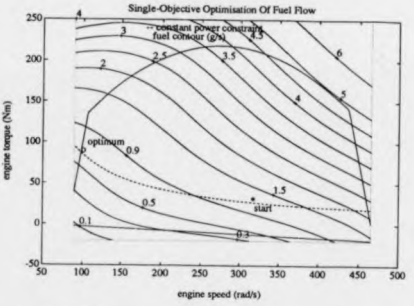


Figure 4.12: A Representation of Fuel Flow Optimisation In 2-D

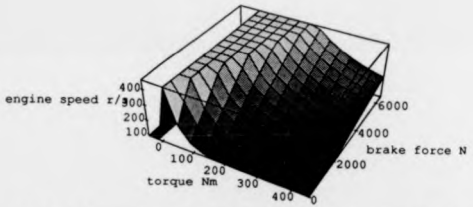


Figure 4.13: A Representation of the Optimisation Constraint Surface In 3-D

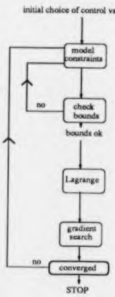


Figure 4.14: Logical Implementation Path for Optimisation

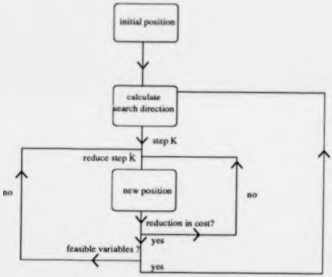


Figure 4.15: Search Direction Implementation

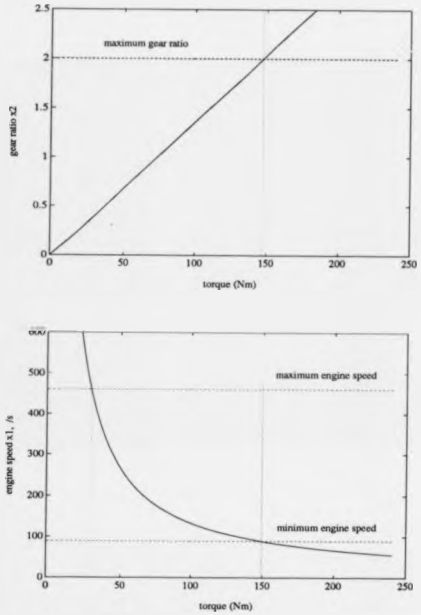


Figure 4.16: Initial Range of Control Variable For 2-D Problem

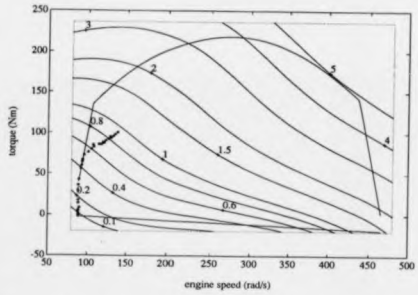


Figure 4.17: Single-Objective Solution Points for Fuel Minimisation

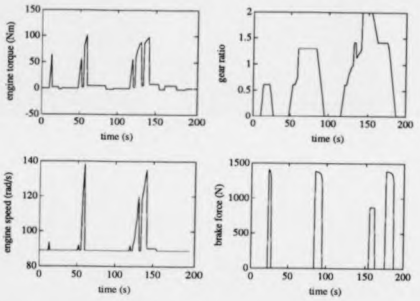


Figure 4.18: Fuel Optimal Schedules

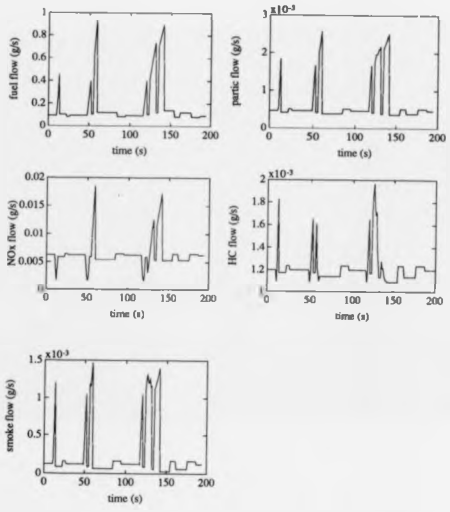


Figure 4.19: Emission and Fuel Flows for Fuel Optimal Schedules

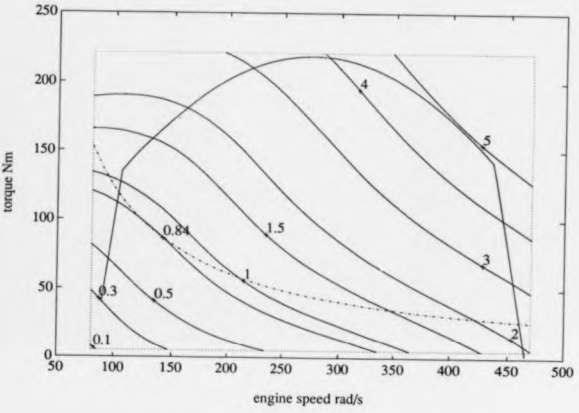


Figure 4.20: Shallowness in the Optimisation Contours

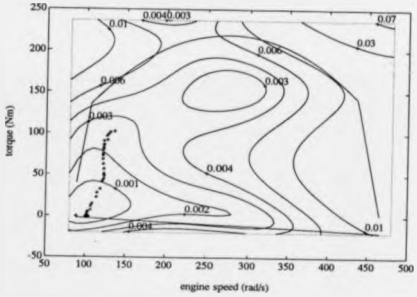


Figure 4.21: Single-Objective Solution Points for Particulate Minimisation

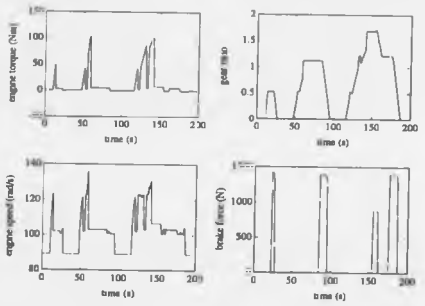


Figure 4.22: Particulate Optimal Schedules

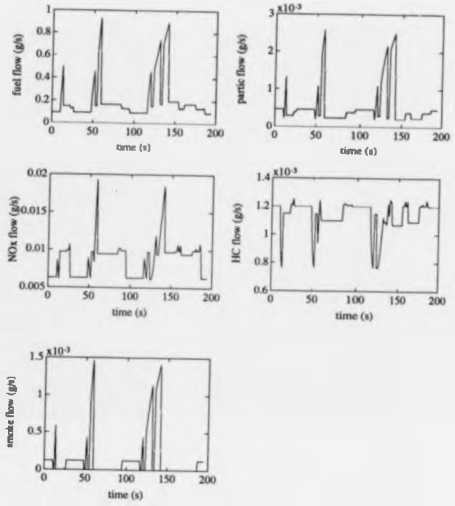


Figure 4.23: Emission and Fuel Flows for Particulate Optimal Schedules

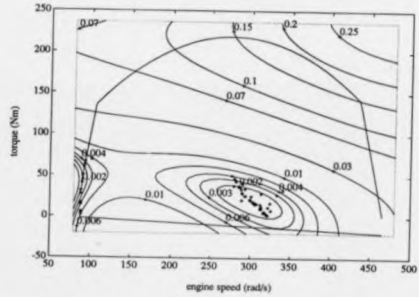


Figure 4.24: Single-Objective Solution Points for NO_x Minimisation

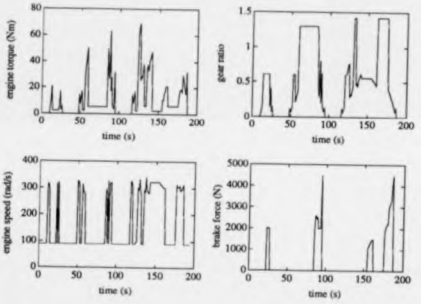


Figure 4.25: NO_x Optimal Schedules

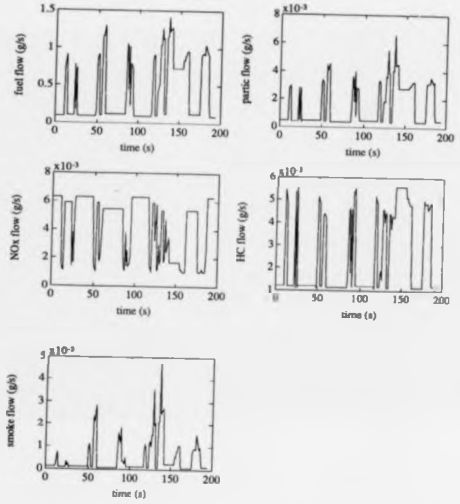


Figure 4.26: Emission and Fuel Flows for NO_x Optimal Schedules

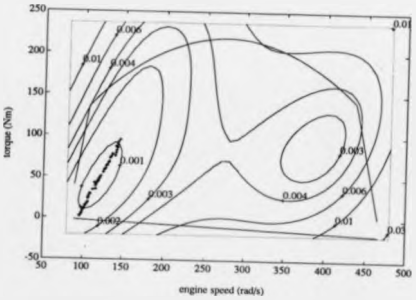


Figure 4.27: Single-Objective Solution Points for HC Minimisation

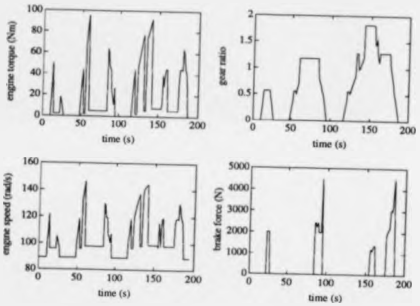


Figure 4.28: HC Optimal Schedules

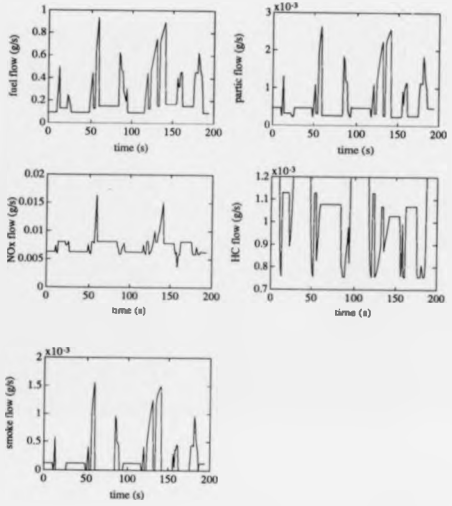


Figure 4.29: Emission and Fuel Flows for HC Optimal Schedules

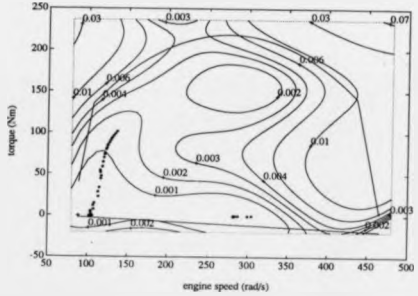


Figure 4.30: Single-Objective Solution Points for Smoke Minimisation

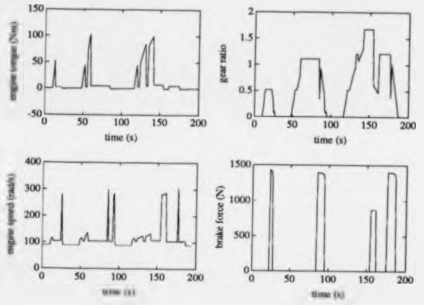


Figure 4.31: Smoke Optimal Schedules

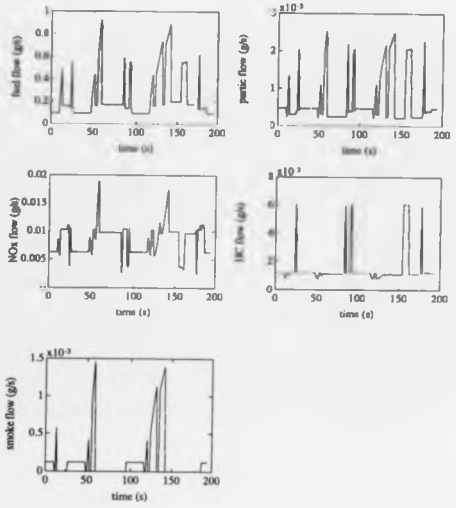


Figure 4.32: Emission and Fuel Flows for Smoke Optimal Schedules

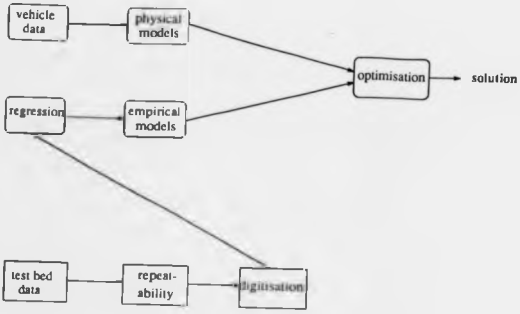


Figure 4.33: Possible Sources of Error in the Formulation

Chapter 5

Powertrain Scheduling as a Multi-Objective Optimisation Problem

5.1 Introduction

In the previous chapter, it was seen that single objective optimisation, as a minimisation of one of the inherent objective functions, is an inadequate solution to the powertrain scheduling problem. The real world is seldom as obliging in producing problems where some, if not all, of the important factors are non-interacting. It was shown that the relationships were often quite subtle and that the manual analysis needed to bring them under one umbrella was enormous. In fact, it was often the case that two or more of the objective functions were conflicting, a degradation in one producing improvement in another. It is this trait which makes the scheduling problem so suitable to multi-objective optimisation techniques.

Multi-objective optimisation deals with minimisation problems for which there are many, possibly conflicting, criteria. For many years, it has been used successfully in financial and socio-economic fields [24]. Classical examples include manufacturing processes which require blending of ingredients : the manufacturer who wishes to optimise the ratio of ingredients and minimise deviations from a colour goal, whilst minimising waste and maximising profit; the investor who wants to maximise the return on his portfolio selection and minimise his risk, whilst maximising his dividends. These same conflicting requirements occur in engineering design problems but, as yet, the techniques being used widely elsewhere have not been widely applied to engineering problems.

The major applications of multi-objective techniques in engineering to date have been in the field of control system design. Grace [81], based on the work of Flemming [20], [82], has used a Control System Design Package (CSD), in conjunction with multi-objective methods, to minimise integral quadratic measures of control system performance. The work, mainly aimed at flight control systems, attempts to improve stability, speed of response, disturbance rejection etc., of the resultant control system. The few other applications have been concerned with the pitch-heave control of submarines [83], dam control problems [84], and gearbox design [85], [86]. Of the few dynamic formulations, Matsuda has also looked at satellite altitude control [87].

Available theory for multi-objective optimisation is much more abundant. Its close links with Pareto optimality [88] and mathematical gaming theory [89], have ensured a continued advancement. Pareto optimality tries to ensure that the designer is supplied with an acceptable compromise solution; gaming theory pits one competitor against another and tries to produce a compromise, over all their objec-

tives, which is acceptable to them both, *e.g* industrial arbitration. It is the former which is most closely related to multi-objective optimisation.

A major obstacle to the solution of meaningful optimisation problems is how to embody the various cost and constraint factors of a typical real world problem into a tractable mathematical model. The most obvious approach is to produce a description, or utility function, which is an exact mathematical representation of the problem. This is extremely difficult if not impossible for complex problems. The engineer must resort to techniques which in some way simulate the real objective of the optimisation.

Three fundamental techniques are available for solution of multi-objective optimisation problems and each, in its own way, overcomes the problem of producing a feasible cost function. The most basic of these is the method of Weighted Sums, or Linear Combinations [85]. The idea is to combine all individual objective functions into a single cost function, which is a linear combination of them all. In essence, the method reduces the problem to a single optimisation of the resultant cost function. This technique is discussed in Section 5.3. Alternatively, the problem can be turned into a constrained optimisation formulation, Section 5.4. One objective function is chosen for minimisation, whilst the others are written as added constraints on the optimisation process. Finally, a method which turns the problem into a constrained optimisation but does so using a mathematical representation of 'distance', is the Goal Attainment method, Section 5.5. Adapted from goal programming theory, the process tries to minimise the distance of each objective from a pre-defined goal value.

The concept and formulation of each method is explained in the appropriate sections and examples of its use given. The various merits and disadvantages of each will thus become apparent and their appropriateness discussed in terms of the

scheduling problem. Finally, justification is given for decisions made concerning the use of a specific method in the scheduling and control analyses of Chapter 6.

5.2 Multi-Objective Optimisation Methods

Before embarking on an explanation of the various methods of multi-objective optimisation, some preliminary concepts are needed. The first being the way in which the multi-object concept can be defined mathematically.

Multi-objective optimisation is the minimisation of a vector of objectives, $G(x)$, which may be subject to a number of bounds, *i.e*,

$$\min_{x \in \mathbb{R}^n} G_i(x) \quad i = 1, .., n. \tag{5.1}$$

subject to

$$\begin{aligned} g_i(x) &= 0, \quad i = 1, .., m \\ g_i(x) &\leq 0, \quad i = m+1, .., m \\ x_l \leq\ & x \leq x_u \end{aligned} \tag{5.2}$$

It is important to note that the function $G(x)$ is a vector of objectives, some of which may be competing. Therefore, there is no unique solution to the problem. Instead, the search is for one which is non-inferior. The term 'non-inferior' derives from Pareto Optimality [88], and is used to define those solutions which represent a compromise. A solution point is non-inferior if improvement in one objective, on movement away from the point, produces degradation in at least one of the others. In order to express this graphically consider a feasible region, Ω, in variable

space $x \in \Re^n$ which satisfies all the constraint equations 5.2. For a two-dimensional problem, Figure 5.1 shows how every point in variable space is mapped into a point in function or trade-off space Λ, such that,

$$\Lambda = \{f(x) | x \in \Omega\} \tag{5.3}$$

A point $\bar{x} \in \Omega$ can thus be defined as a non-inferior solution point if and only if for some neighbourhood of $\bar{x}$ there does not exist a Δx such that $(\bar{x} + \Delta x) \in \Omega$ and,

$$f_i(\bar{x} + \Delta x) \leq f_i(\bar{x}), \quad i = 1, .., m \tag{5.4}$$

For the two dimensional problem shown, the non-inferior solutions lie on the portion of the boundary AB, Figure 5.2.

Viable techniques of multi-objective optimisation are those therefore, which search amongst non-inferior solution points. Methods capable of this will now be discussed.

5.3 The Weighted Sum Method

The most basic approach to the problem of producing a cost function which adequately represents the minimisation of all the objective functions is the Weighted Sum method [85]. The basic idea is to add together each of the objective functions to form a new objective function, $\hat{G}(x)$, which is a linear combination of the individuals. Minimisation of the new objective function can be written,

$$\hat{G}(x) = \min_{x \in \Re} \sum_{i=1}^{m} \theta_i G_i(x)^2, \tag{5.5}$$

such that the set Θ of all weighting vectors can be defined by

$$\Theta = \{\theta \in \Re^n \mid \theta_i > 0, \sum_{i=1}^{n} \theta_i = 1\} \tag{5.6}$$

This is equivalent to defining a supporting hyperplane L, to the trade-off space Λ, as illustrated in Figure 5.3. The hyperplane is tangent to Λ at the terminal point of a non-inferior vector. Minimisation of the objective function $G(x)$ is equivalent to minimising a vector, extending from the origin, to the hyperplane L, in an L_2-norm sense, such that a non-inferior solution point is found. The weighting choices, in this case θ_1 and θ_2, define the slope of the tangent which, in turn, determines the specific non-inferior solution point reached. By choice of the various weighting parameters, θ_i, different combinations of objectives can be minimised. The solution should then reflect the relative importance attached to each of the objectives by the weightings.

A serious limitation of the Weighted Sum method is that existence of a supporting hyperplane for every non-inferior solution point is only guaranteed for convex trade-off spaces. Without convexity, a subset of the non-inferior vectors may terminate in the gap between the boundary of Λ and the hyperplane L, Figure 5.4. For problems with nonlinear objective functions it is unlikely a convex trade-off space could be guaranteed. Awareness of this problem does not become apparent to the user, since complex physical problems can seldom be interpreted graphically in this way. If a gap does exist compromise alternatives may be overlooked because they cannot be found using any combination of the weighting parameters.

A second, more implementational, problem is the choice of the weighting parameters themselves. Since one typically has little *a priori* knowledge of the weighting factors that should be chosen, the only way to proceed is to solve the problem for

a large number of different weighting factors. For large problems, of more than two or three objectives, this becomes unreasonable and only a limited number of trials can be attempted.

Various attempts have been made to overcome the difficulties of the method. Zadeh [90], and Geoffrin [91], have advocated more sophisticated approaches in which the cost functional is ordered in some manner.

Use of the Weighted Sum method for solution of the scheduling problem is discussed in the next section, where implementational problems will become more apparent.

5.3.1 Powertrain Scheduling as a Weighted Sum Problem

The minimisation of fuel and emission functions as a scheduling problem can be defined in terms of a weighted sum objective function as follows,

$$\bar{G} = \theta_f f(x_1, u_1)^2 + \theta_p p(x_1, u_1)^2 + \theta_n n(x_1, u_1)^2 + \theta_h h(x_1, u_1)^2 + \theta_s s(x_1, u_1)^2 \quad (5.7)$$

where functions f, p, h,n, and s are the polynomial representations of the fuel and emission functions. The optimisation problem,

$$\min \bar{G}(u) \quad (5.8)$$

subject to constraint equations, 4.10 to 4.14, is thus a constrained minimisation problem and can be implemented using standard constrained optimisation routines. In fact, the MATLAB routines of Appendix A can be simply adapted to incorporate the new cost function, although derivative calculation of all objective functions is

tedious. The choice of weights is dependent on the required emphasis. In this case, all results will be for the same weighting choices at every point of the test cycle. This is the simplistic approach. For greater control over the numerical procedure, a different weighting choice could be made at each test point but this is an unnecessary complication at this stage. Use of the routines, to produce results for various weighted optimisations, is discussed in the next section.

Optimisation Results

The Weighted Sum method involves combining all objective function definitions into one cost function. However, the individual objective functions are of differing range and in order to produce feasible results all functions must be scaled to some nominal magnitude. The simplest way of achieving this is to normalise all functions to the mean of the fuel flow function, the fuel function being chosen arbitrarily. The resultant scaling factors are shown in Table 5.2. The cost function then becomes,

$$G(u) = \theta_f f^2 + \theta_p(p_m p)^2 + \theta_n(n_m n)^2 + \theta_h(h_m h)^2 + \theta_s(s_m s)^2, \qquad (5.9)$$

where p_m, n_m, h_m, and s_m are the scaling factors as shown in Table 5.2.

The result of optimisations, run under the Weighted Sum formulation, is shown in Table 5.3. The table shows the overall test cycle fuel and emission output for three different sets of weighting choices θ_i. The first set is for an optimisation where all weights have been set equal. This is the most obvious choice to make for a first run of the problem since it shows how the optimiser reacts when all functions are on the same level. If the figures shown here are compared with the single-objective optimisation results of Table 4.1, some idea of the compromises

which have occurred can be gained. Overall, the variance from the single-objective optimum values varies from about 8% for HC to 66% for NO_x. However, this NO_x figure can seem deceiving, since the Weighted Sum output compared to the fuel, particulate, HC and smoke single-objective output values for NO_x is quite good. The engine map for this particular result in shown in Figure 5.5. The points are seen to lie in the bottom left hand corner of the engine map, where most of the function minima are found. On the whole, the result is quite a good compromise solution. However, a designer may feel that the particulate and NO_x output values are not quite as good as required, and that less importance can be attached to the fuel, HC and smoke output values. A second optimisation can thus be carried out.

The second level in Table 5.3 shows the results of an optimisation when more emphasis, *i.e* a greater weighting value, is placed on particulate and NO_x functions. In this case, particulate is valued equally with NO_x; fuel, HC and smoke take next equal rankings. The numerical values show that the extra weighting has been picked up on the NO_x output but not on the particulates, which have slightly increased. This suggests that either there are non-inferior solutions, giving lower particulate output, which cannot be reached, because of the non-convexity problems already discussed, or that the compromises are too great to reduce both functions at the same time. Increasing the weighting to favour particulates more does not, in fact, produce the improvement required.

The last set of values show the results of an optimisation when HC is deemed more important than the other functions, with NO_x ranked second. For the new weighting choices, the results show only a slight improvement in HC over the equally weighted solution. The compromises are such that NO_x output has increased despite a slightly increased weighting this time.

5.3.2 Conclusions

There are many flaws to the practical use of the Weighted Sum method - the most important of which is the nonawareness of the non-convexity problem. If one or more of the objectives are more important than the others, an adequate solution may not be found because of the nature of the trade-off boundary in the particular region. This has been illustrated in the results above. The scheduling problem has a total of five objectives and thus demands a search over a five-dimensional trade off space. Knowledge as to the nature of this space cannot be gained. For more subtle occurrences, the problem may go unnoticed. This means that certain solutions may be accepted as the best possible when other, 'better' ones, exist but cannot be found.

Implementationally, the method is difficult to use for complex problems. For such a large problem it is infeasible to try all combinations of weighting parameters when all the objective functions are included in the formulation. The weightings themselves give only a quantitative measure of the relative importance of each objective and knowledge of the interaction which individual choices define, must be gained from only a few runs of the optimisation. As has been seen, the result gained from a weighting choice is not always that expected.

Finally, for a problem such as scheduling, where legislative limits may have to be achieved, there is no inherent facility for including these into the formulation of the Weighted Sum method. Such added constraints have to be included as part of the variable constraints at the implementation stage.

As a useful method for the solution of the scheduling problem therefore, the Weighted Sum method is inappropriate. Despite the small implementational cost from single-objective routines, (the cost function changes only slightly although the increased cost of defining all derivatives is more troublesome), the method cannot

cope with either the size or the complexity of the problem. It has been useful however, as a first pass of the problem, aiding in its understanding and its requirements.

5.4 The ε-Constraint Method

A different approach to the multi-objective problem is to prioritise more explicitly the individual objective functions. Choosing a primary function for minimisation allows the other objectives to be represented as added constraints on the optimisation process. These sensitivity constraints restrict the path of the optimised function so as to comply with the bounds placed on the other functions. Essentially, the method converts the multi-objective problem into a single-objective optimisation of one of the independent functions subject to additional constraints.

The two-dimensional graphical interpretation of this is illustrated in Figure 5.7. Considering objectives G_1 and G_2, if G_1 is chosen as the primary objective the search along the boundary proceeds in the direction shown. A limiting value placed on the other objective G_2 forms the constraint boundary, beyond which the optimiser cannot move. The set of attainable non-inferior solution points therefore lies along the portion of the curve CD.

The formulation can be expressed mathematically as,

$$\min_{x \in \Omega} G_p(x) \tag{5.10}$$

subject to

$$G_i(x) \leq \varepsilon_i \quad i = 1, .., m. \quad i \neq p. \tag{5.11}$$

This hierarchical approach to the problem has been attempted by Waltz [92]

and others [93]. The method of Waltz, for example, requires that the objectives are ranked in their order of importance. The optimum value for the primary index alone is found first. The optimum of the secondary objective is then found subject to a constraint that the value of the primary index is held within a certain percentage of its optimum. The third index is treated similarly with the first two included as constraints, and so on until all objectives have been considered. If only a primary index can be identified, the method of Nelson [93] consists of finding the optimum value of the primary objective while all other indices are held within prescribed bounds.

As with the Weighted Sum method, the main problem encountered is the choice of the inherent parameters. The solutions are biased by the assumptions made on constraint values at each stage, when little or no justification can be made. In Waltz's method [92], the primary objective is obviously favoured. The same is true of Nelson [93], the added uncertainty being the choice of bounds placed on the remaining objectives. However, for problems in which there are inherent limits, such as legislative limits, this uncertainty can be reduced.

The main advantage of the ε-Constraint method , is the removal of the problem caused by non-convexity of the trade-off space. As seen in Figure 5.7, all non-inferior solutions, within the constraints of the problem, can be found, even when the surface is non-convex.

5.4.1 Powertrain Scheduling as an ε-Constraint Problem

The scheduling problem written as an ε-constraint problem is defined by,

$$\min_{u} f(x_1, u_1), \tag{5.12}$$

subject to,

$$\begin{aligned} p(x_1, u_1) &\leq \varepsilon_p, \\ n(x_1, u_1) &\leq \varepsilon_n, \\ h(x_1, u_1) &\leq \varepsilon_h, \\ s(x_1, u_1) &\leq \varepsilon_s. \end{aligned} \tag{5.13}$$

where, in this case, the fuel flow rate has been chosen as the primary function and the other functions have been constrained relative to predefined bounds. The choice of these bounds is discussed more fully later. The fuel function was chosen as the most obvious primary function since it is the major concern of both the motorist and the manufacturer. Defined in this way, the optimisation solves a minimum fuel flow problem, with added conditions linked to the adverse effect this may have on the emission outputs. An emission function could have been chosen but, since no clear legislative limit exists for fuel, choice of a bound would have added uncertainty. The same is true of the smoke function but, since it is not considered as important, a very loose, arbitrary, limit was chosen.

The overall formulation is of a constrained single-objective optimisation. The methods and routines of Chapter 4 and Appendix A apply but, because of the increased complexity due to the emission constraints, an available optimisation toolbox

was used to implement the problem [81]. The toolbox removes lengthy analytic calculations of derivative functions, for Lagrange Multipliers for example, and provides a flexible interface to a number of standard optimisation algorithms.

The toolbox is part of a MATLAB Control System Design environment, and although only a β-test version at the time of use, is now available commercially. The routines implemented within it are mainly second-order, quasi-Newton methods, as discussed in Chapter 3, with use of Lagrange Multiplier theory and active constraint sets to deal with constrained optimisation. The toolbox also has the facility to deal with single and multi-objective optimisation problems. It was tested, by comparing the results produced with the toolbox, to those produced by the routines of Appendix A. Within the confines of different convergence and termination criteria the results were the same.

The ε-Constraint problem is defined within a MATLAB header file, which then calls the appropriate optimisation routines. The header file also contains all other choices available within the toolbox, *e.g.* optimisation parameters, and any appropriate scaling.

Optimisation Results

The fuel minimisation problem as defined by equations 5.12 and 5.13 was implemented for specific values of ε_i, the adequate formulation of which is dependent on knowledge of appropriate limits. Legislative limits, as shown in Table 4.2 are available on a gram/test basis for particulate, NO_x, and HC emission outputs. However, since finite-dimensional optimisations are being carried out at each data point on the test cycle, the ε limits needed are defined on a gram/second basis. This is the fundamental stumbling block of the ε-constraint method for use in the scheduling

problem.

The most obvious choice for limits, given the information available, is to take the mean of the limit over the whole test cycle. This provides a crude approximation to the actual output of the engine, which is dependent on the velocity and acceleration state of the vehicle. Ideally, the limits should reflect this. The greatest disadvantage of this approximation however, is that the limits chosen must be achievable. In terms of the single-objective optimisations of Chapter 4, the constraint line has to pass within the working region of the engine, Figure 4.12. Now, the additional constraints impose the added requirement that the set of constraint lines move through a specific region of the engine map. If this is not the case, the problem is unsolvable. The region specified by the mean limits for particulates, NO_x, and HC is shown in Figure 5.8, the actual limits being shown in Table 5.1. These are the mean of the legislative limits shown in Table 4.2. The limits determine boundaries within which the solution must be found. Figure 5.8 shows that the region defined by the HC limit covers most of the engine's working range, whilst the particulate limit only envelopes the low engine speed/torque region. The NO_x bound determines two separate feasible regions in the low torque portion of the engine map. As can be seen, the intersection of bounded regions, leaves only a small feasible subset in the bottom left hand corner of the engine map, the maximum torque allowable in this region being only about 60Nm.

On running the optimisation, it is found that the mean limits do produce an infeasible problem at some points, *i.e* the constraint line does not pass through the reduced feasible region. Table 5.4 shows the output values when the mean limits are relaxed slightly; Figure 5.10 shows the new larger feasible region. The ε limits shown produce the smallest feasible region for which all constraints lines are

attainable. Although this has produced a solution, the new limits do not reflect the real problem and do not constitute a solution. Thus, a different limit formulation is needed. The ideal would be knowledge of the maximum allowable limits for each speed-acceleration data point on the cycle. This however, was unavailable and since there is no 'correct' vector of limits for the whole two hundred seconds of the test cycle portion, the task of intuitively producing such was considered to be inappropriate at this stage.

5.4.2 Conclusions

The ε-Constraint method in either the formulation of Waltz [92] or Nelson [93], is inappropriate to the demands of the fuel and emission scheduling problem. The recursive formulation of Nelson [93], was not implemented here because of the size of the problem. This method would need five independent single-objective optimisations, the latter of which would have four added constraints. The simpler formulation of one single-objective optimisation, subject to predefined limits on the other functions, has been attempted. That too, however, has been found to have problems.

The major advantage of the ε-Constraint method over that of Weighted Sums is the removal of the non-convexity problem. However, like Weighted Sums it converts the multi-objective problem into a single-objective equivalent. The choice of the constraint parameters is difficult, both on a general level and more specifically for the scheduling problem, the general disadvantage of the method is that the limits are being imposed on the solution at an early stage. How the choice of these limits affects the subsequent solution is very difficult to gauge. This is important if individual limits at specific test cycle points were to be attempted, since the choice could

seriously affect the results obtained.

The method does implicitly consider the legislative values which define limits on the emission output, but the formulation is such that they are inflexible. The limits must be achievable. As has been seen in the results presented here, the use of mean limits, which is the simplest choice, is inappropriate, and much more work would have to be done to find realistic limits for each data point. The method is also inflexible in that one function alone must be minimised. It is often the case that no one function stands out against any others since what is required is a true compromise solution. The ϵ-Constraint method does not accommodate this in any real sense.

5.5 The Goal Attainment Method

The Goal Attainment method of Gembicki [94], has its roots in goal programming, most applications of which have been in the fields of economics and management. The advantage of goal programming, which transfers across to the Goal Attainment method, is that the parameters have convenient intuitive interpretations. The bonus of the Goal Attainment method however, is that the non-convexity problems, associated with goal programming, are not incorporated.

The Goal Attainment formulation can be described mathematically by,

$$\min_{x \in \Omega, \gamma} \gamma, \tag{5.14}$$

subject to constraints,

$$G_i(x) - \omega_i \gamma \leq G_i^* \quad i = 1, ..m \tag{5.15}$$

where $\omega_i > 0$, G_i^* are scalar values, and γ is an unrestricted scalar variable. The values assigned to each G_i^* can be interpreted as the desired level of the corresponding objective function, $G_i(x)$, *i.e* the goal values. The quantity $(\omega_i\gamma)$ introduces an element of slackness and is related to the degree of under or over-achievement of the goal, with ω_i being the weighting coefficient of each objective. Note that the scalar parameter γ also becomes an optimisation variable. This is explained more clearly by a two-dimensional representation, Figure 5.9. The goal values define a vector position G^* in the function space. The weighting parameters then define a direction of search through that space to the goal value. The constraint equations impose boundaries shown, which change the size of the trade-off space as γ is minimised. The figure also shows that non-convexity of the trade-off space does not cause some non-inferior solutions to be overlooked, since no tangent hyperplane is needed.

Essentially, the method converts the multi-objective problem into the constrained optimisation of a scalar variable. In optimising γ, the method minimises how closely each objective achieves its goal. Since all objectives are simultaneously considered through the constraint equations it is more akin to a true multi-objective optimisation.

The method assumes that the engineer or decision-maker has a sufficient intuitive understanding of the problem at hand to know what values should be chosen for the goals. These are often well-defined by the problem and can be legislative limits, as in Section 5.4.1, single-objective minima, *i.e* the' best' we can do, or values borne out of experience of the problem, *i.e* intuitive values. The weighting parameters determine how near each objective will get to its goal value in function space. For any vector two possible conditions are apparent: the goal value is contained in the trade-off space, *i.e* it is attainable; the goal value lies outside the trade-off space

and thus is unattainable. In either case, the relative under or over-attainment that results will be proportional to the values of the corresponding weighting parameters and the compromises taking place.

Suppose that, for example, the goal vector, $G_k(x)^*$ for a particular k is attainable. If ω_k is large in relation to the other weighting parameters, the overattainment, $G_k(x) - G_k(x)^*$, will also be large. That is, the relative magnitude of ω_k determines the closeness of the solution point to its goal. A trade-off is thus made by sacrificing the attainment of one goal against another.

The formulation is much more flexible than that imposed by either the Weighted Sum or ε-Constraint methods. How the scheduling problem can be formulated as a Goal Attainment problem is discussed in the next section.

5.5.1 Powertrain Scheduling as a Goal Attainment Problem

The minimisation of fuel and emission functions can be written as a Goal Attainment problem as,

$$\min_{x,\gamma} \gamma \tag{5.16}$$

subject to ,

$$\begin{aligned} f(x_1, u_1) - \omega_f \gamma &\leq f^*, \\ p(x_1, u_1) - \omega_p \gamma &\leq p^*, \\ n(x_1, u_1) - \omega_n \gamma &\leq n^*, \\ h(x_1, u_1) - \omega_h \gamma &\leq h^*, \\ s(x_1, u_1) - \omega_s \gamma &\leq s^*, \end{aligned} \tag{5.17}$$

and the variable constraints of equations 4.10 to 4.14. Each fuel and emission function has a respective goal and weighting parameter as shown. As with the Weighted Sum method, it is usual to normalise the weightings such that,

$$\sum_{i=1}^{n} \omega_i = 1, \quad \omega_i \geq 0. \tag{5.18}$$

The problem is implemented through the use of the MATLAB Optimisation Toolbox [81]. The toolbox allows the use of the Goal Attainment method through modified use of the constrained optimisation routines. A special program, within the toolbox, converts the standard constrained description of the problem into the Goal Attainment formulation above, i.e all constraint equations are converted into the form of 5.17 each having a goal and weighting value. The new objective function, equation 5.16, is also defined. Variable bounds are handled by writing them as inequality constraints. By setting the weightings to zero they are defined as goals which must be strictly achieved. The equality constraint can be defined as such within the toolbox, a zero weighting again defining that it must be strictly achieved.

The following section shows the results of various choices of the weighting parameters and different definitions of the goal values.

Optimisation Results

Six separate optimisation results are presented within this section in order to show the consequence of different weighting and goal choices. As discussed, the goal choice can take one of three forms : intuitive values, legislative limits, and single-objective minima. Only the latter two will be discussed here, for various weighting choices. Each has its own advantages and these will become apparent.

Goals which are mean legislative limits, as they suggest, are constant over the

whole test cycle. It is more feasible to use them here since, unlike the ϵ-Constraint method, they do not have to be strictly achievable. This removes the need to define a specific limit at each point. Goals which are single-objective minima obviously have a different goal value at each data point on the test cycle. There are also various options over the weighting choice. Here, the weighting choice for a specific optimisation is constant throughout. This is the simplest approach. If greater control over the solution is needed in specific areas of the test cycle, as may be the case in a scheduling problem, this can be easily implemented.

As with the Weighted Sum method, the first sensible optimisation is that where all fuel and emission weights are equal. This gives an idea of how close a function gets to its goal value when all functions are optimised equally, i.e does it under or over achieve?. The result is a compromise which can be inspected for problem areas.

The optimisation results are shown in Table 5.5. The first set of data defines the results for an equally weighted optimisation which has mean legislative limits as its goal values. The value of the goals both on a second by second basis and for the quarter test cycle are also shown in the table. For completeness, goal values are also given to the fuel and smoke functions. Without legislative definitions, these have been chosen as relatively loose limits which are likely to be achieved. The resultant schedules and function flow rates over the cycle are shown in Figures 5.11 and 5.12.

The equally weighted optimisation has given quite a good compromise solution. Smoke and HC especially, have done well, both overachieving the mean limits at all points of the cycle. Fuel, particulate and NO_x have achieved their limits for all but the higher velocity-acceleration points. This initial optimisation has thus given better knowledge of the nature of the functions relative to their goal values. The exact nature of the resultant schedules, although shown, will not be discussed in

detail in this section. It is the nature of the method and its implementation which is important here.

Given knowledge of the functions' achievements relative to their goals, other optimisations can be conducted for various scenarios. The following two data sets, in Table 5.5, show the results of particulate and NO_x emphasized optimisations. The first is the NO_x optimised solution, Figures 5.13, 5.14. Since it is known that HC and smoke achieve their values quite comfortably, the weightings on these can be relaxed, to specify only a small amount of over-achievement. Fuel and particulates are given smaller valued weightings so that both over and under-achievement portions are optimised to some degree. This allows a larger weighting to be attached to the NO_x function. This will improve performance where the function can achieve its goal, but not necessarily improve those that it cannot. The result is seen in the table and in the schedule and flow diagrams of Figures 5.13 and 5.14. The NO_x output has been reduced considerably, and the compromises which have taken place to achieve this are quite obvious.

The third set of data shows what happens when the importance of the particulate function is increased. The result is an 8% drop in particulate output over the equally weighted solution, the change coming mostly from the deceleration phases, Figures 5.15 and 5.16. The compromise is however that NO_x has increased by 21% over its equally weighted solution, and 65% over its own emphasized solution, the changes again being due to a worse performance over the constant velocity phases. The engine speed over the cycle is generally low, especially during deceleration, when lube oil particulates, being sucked out from the pistons, can increase the particulate output if the engine speed is high. Similarly, the engine torque is low. The gear ratio goes well into overdrive and the braking is moderate. The effect of the weighting

change can be seen more clearly on the engine map comparisons of Figure 5.17. The top and bottom maps show the solution points for particulates and NO_x respectively. The greater weighting emphasis on the NO_x function has moved the majority of solution points into the high engine speed, low torque region of the engine map. This illustrates nicely the ability to manipulate the output solution through the relative weightings choices.

These first three sets of data have thus illustrated the use of the mean legislative data as goal values. However, certain disadvantages have become apparent. The relative size of an individual weight determines the over and under-achievement at every optimisation point. A large weight will cause both large under and over-achievement within the same cycle. It seems impractical therefore to use a blanket weighting choice over the whole cycle. This is where the equally weighted solution becomes important, in order to illustrate to the designer what achievement is possible, and where. Different portions of the cycle can then be given different weighting choices. If this is inappropriate, a different definition of the goals is needed and this is where the single objective minima become useful.

The single-objective (SO) minima are, as they suggest, the least possible values of each function within the constraints of the problem. Thus a multi-objective compromise solution can only hope to under-achieve these results. This removes the ambiguity found above. Table 5.5 shows the results of three Goal Attainment runs for equal weights, particulate emphasized and NO_x emphasized solutions. The latter do not have the same weighting choices as the mean goal solutions since the emphasis is on discovering the merits of the goal choices, and in gaining knowledge of the weights relative to those goals. Also shown in the table for comparison, are the individual SO minimum values over the test cycle.

The equally weighted solution shown, Figures 5.18, 5.19, can be seen to be very good for the fuel function and relatively good for the other functions. This is shown clearly in the flow functions of Figure 5.19. The fuel function only fails to achieve its value over the deceleration phases of the cycle. Particulates and smoke achieve their goals over the acceleration phases, with NO_x and HC doing less well over a greater proportion of the test cycle.

For better performance of the NO_x and particulate functions, a weight which defines a smaller degree of under-achievement is needed. This is the case for the second and third sets of SO goal data. The first is when particulate and NO_x are given the same weights, less emphasis being given to fuel, HC and smoke, Figures 5.20, 5.21. The results show a marked improvement in the NO_x value of 15% over its equally weighted value. Since NO_x and particulate are widely opposing functions though, the same improvement has not been seen in the particulate function. In order to improve this, the NO_x weighting must be relaxed to give more emphasis to a better particulate result. The result is seen in the table and in the diagrams of Figures 5.22, 5.23. The improvement is most noticeable over the high speed-acceleration phases. As might be expected, the higher emphasis on particulate, together with slightly reduced weights on the other functions, has forced a small improvement in the fuel, HC and smoke functions too. The change which has taken place over the previous two solutions is most clearly seen if the solution points are plotted on an engine map, Figure 5.24. The top graph show the particulate emphasised solution for the single-objective goals. The sharp movement of many of the points across the engine plane in the lower graph is the result of emphasizing NO_x. Slight changes in weighting parameters can thus produce dramatic movement of the solution points.

5.5.2 Conclusions

Overall, both goal definitions used here have shown some merit. The solutions which have been produced have been true compromise solutions, and the ease of manipulating them has been shown through the various weighting choices. The respective weights will be dependent on the specifications of the engineer. The goal choices will be determined by the problem, *i.e.* if mean limits are chosen, is it easy to define different weights at different portions of the cycle. This is removed when single-objective minima are used but the solution tends to be less flexible, although again blanket weightings need not be used.

The use of the MATLAB Optimisation Toolbox and its support of the Goal Attainment method has been especially useful. It provides an efficient user-interface which allows the complex workings of the optimisation algorithms to become transparent to the user. The supplementary information it provides is also of value. This includes a list of those constraints which are active at the solution, especially important in scheduling since the engineer needs to know which bounds are restricting further improvement.

What has become apparent in this section is the worth of the Goal Attainment method as a multi-objective optimisation tool, its ease of implementation, and especially, its suitability to the scheduling problem.

5.6 Summary

The results of this chapter clearly illustrate the worth of the Goal Attainment method over other methods discussed. Their major stumbling blocks are overcome to produce an appropriate method of solution for the scheduling problem.

The method is able to handle the complexity and the size of the problem : non-

convexity does not restrict the non-inferior solutions which can be found, and the method can quite easily handle the optimisation of five objectives simultaneously. It has also overcome the inflexibility of the ε-Constraint method, which imposed strict bounds on the optimisation process. Here, the bounds can be under or over-achieved dependent on the specifications of the engineer. The choice of bounds therefore does not have such far reaching consequences on the feasibility of the method.

One advantage of the ε-Constraint method was that it implicitly incorporated the emission constraints into the optimisation process. This is the case for Goal Attainment too, but with the added ability to weight the relative importance of each. In being able to meter how close an objective gets to its goal, the weighting choices are much more quantitative. The weighting choices of the Weighted Sum method simply emphasized amongst the objectives functions.

In effect, the Goal Attainment method is retaining a greater degree of the multi-objective nature of the problem. It does not simply reduce the problem back to a single-objective optimisation, be it a combination of all the functions, or only one of them. By its nature, it combines the advantages of the other methods without incorporating their downfalls, and in doing so, produces a true compromise solution.

With respect to the scheduling problem, the Goal Attainment method has been able to handle its complexity, and provide the flexibility needed. The results have not been discussed in detail here, only the method's general use. It has been inappropriate to discuss the details until a suitable technique had been found. That technique is now known to be the Goal Attainment method. However, whatever the worth of the method, the resultant schedules, must be implementable in control terms and Chapter 6 will now discuss its use for a particular goal specification, the respective results and how these can be reduced to implementable control laws.

Limit Values (g/s)			
Partic	NO_x	HC	Smoke
0.0017	0.0065	0.0051	0.002

Table 5.1: Mean Legislative Limits Over The ECE-15 Test Cycle

Partic	NO_x	HC	Smoke
273.1	55.9	469.5	353.0

Table 5.2: Weighted Sum Scaling Factors

Weights					Cycle Predictions				
θ_f	θ_p	θ_n	θ_h	θ_s	f	p	n	h	s
0.2	0.2	0.2	0.2	0.2	35.45	0.103	0.754	0.137	0.039
0.1	0.35	0.35	0.1	0.1	38.63	0.133	0.718	0.167	0.070
0.05	0.05	0.25	0.6	0.05	37.87	0.103	0.826	0.133	0.0395

Table 5.3: Weighted Sum Optimisation Results

Cycle Predictions				
f	p	n	h	s
27.77	0.097	0.841	0.159	0.039

Table 5.4: ε-Constraint Optimisation Results

Goal	Weights					Output (g) [a]				
	ω_f	ω_p	ω_n	ω_h	ω_s	f	p	n	h	s
mean	0.2	0.2	0.2	0.2	0.2	30.68	0.085	0.889	0.146	0.029
	0.18	0.18	0.3	0.18	0.16	58.28	0.208	0.652	0.374	0.068
	0.01	0.6	0.18	0.2	0.01	32.93	0.078	1.077	0.144	0.025
mean limits (g/s) [b]						0.5	0.0017	0.0065	0.0051	0.002
goal for 1/4 cycle (g) [c]						100	0.34	1.35	1.0	0.4
S.O.	0.2	0.2	0.2	0.2	0.2	27.82	0.096	0.816	0.155	0.038
	0.33	0.01	0.01	0.33	0.32	31.95	0.105	0.681	0.167	0.050
	0.29	0.01	0.1	0.3	0.3	29.08	0.091	0.860	0.149	0.034
sing-obj. minima (g)						27.59	0.075	0.452	0.127	0.022

[a] output for the test cycle
[b] goal value at each optimisation point
[c] overall goal for test cycle optimisation

Table 5.5: Goal Attainment Optimisation Results

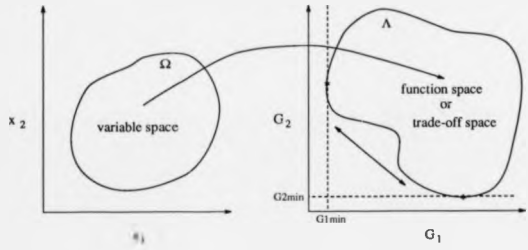


Figure 5.1: A Graphical Interpretation of Multi-Objective Optimisation

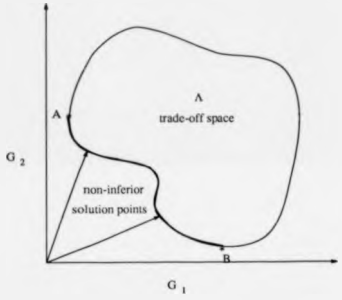


Figure 5.2: The Set of Noninferior Solution Points

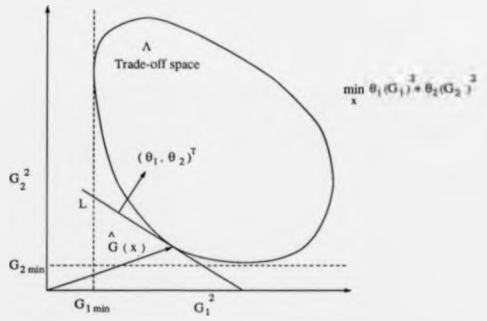


Figure 5.3: A Graphical Interpretation of the Weighted Sum Method

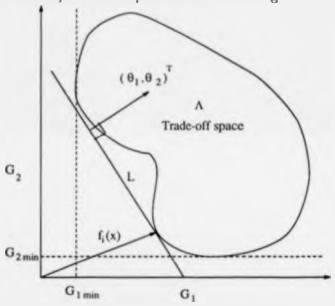


Figure 5.4: Non-Convexity of the Trade-Off Space

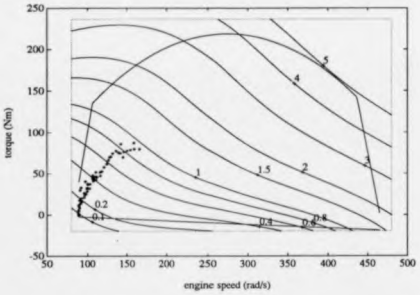


Figure 5.5: Solution Points for Weighted Sum (Equally Weighted)

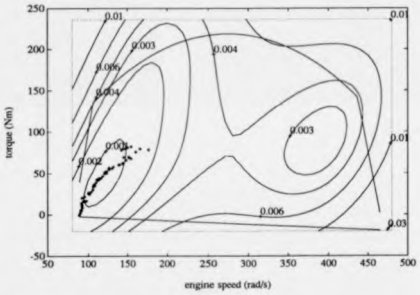


Figure 5.6: Solution Points for Weighted Sum (HC Emphasised)

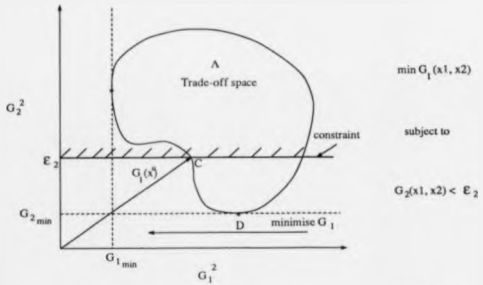


Figure 5.7: A Graphical Interpretation of the ε-Constraint Method

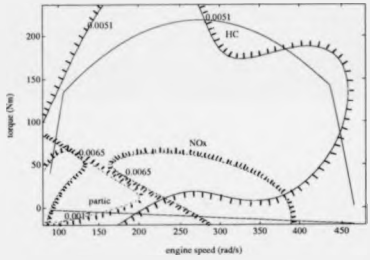


Figure 5.8: The Feasible Region Imposed By Emission Constraints (Exterior of Boundary)

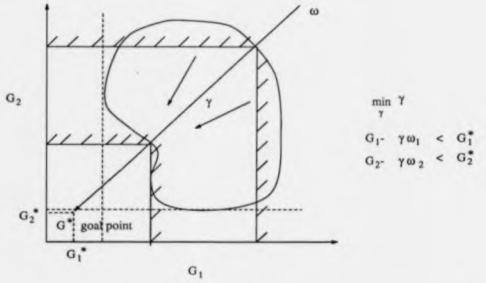


Figure 5.9: A Graphical Representation of the Goal Attainment Method

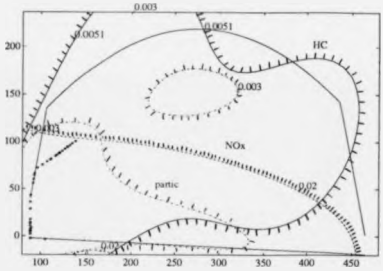


Figure 5.10: ε-Constraint Solution for Loose Limits ($\mathfrak{A}$ Exterior of Boundary)

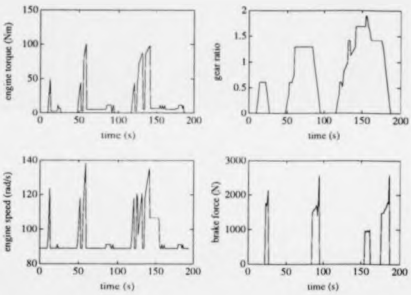


Figure 5.11: Schedules For Goal Attainment : Equal Weights, Mean Goals

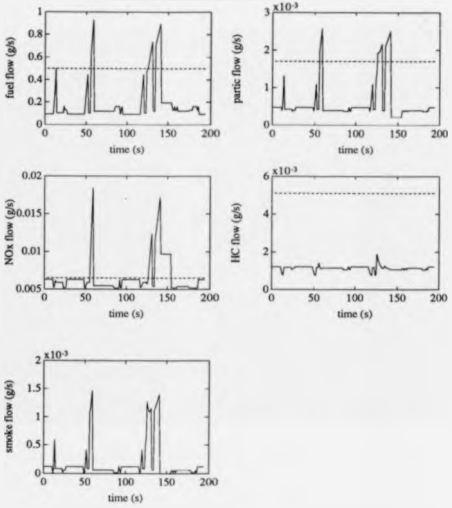


Figure 5.12: Fuel and Emissions For Goal Attainment : Equal Weights, Mean Goals
(- optimal flow g/s, - - mean limit level g/s)

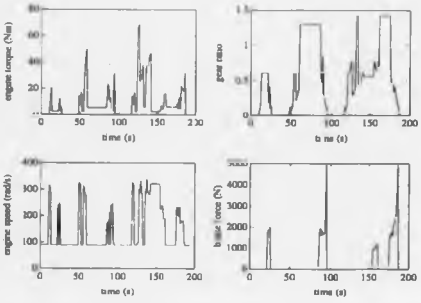


Figure 5.13: Schedules For Goal Attainment : NO_x Weighted, Mean Goals

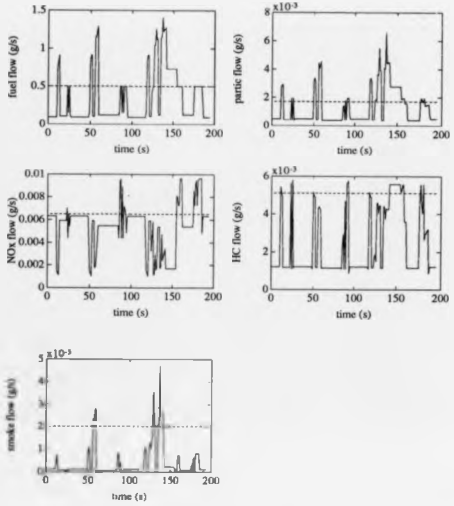


Figure 5.14: Fuel and Emissions For Goal Attainment : NO_x Weighted, Mean Goals (- optimal flow g/s, - - mean limit level g/s)

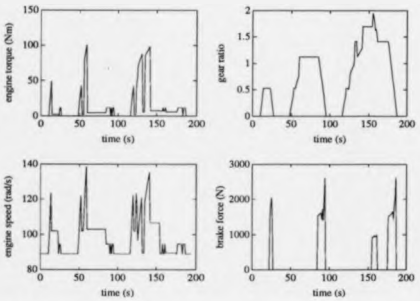


Figure 5.15: Schedules For Goal Attainment : Partic. Weighted, Mean Goals

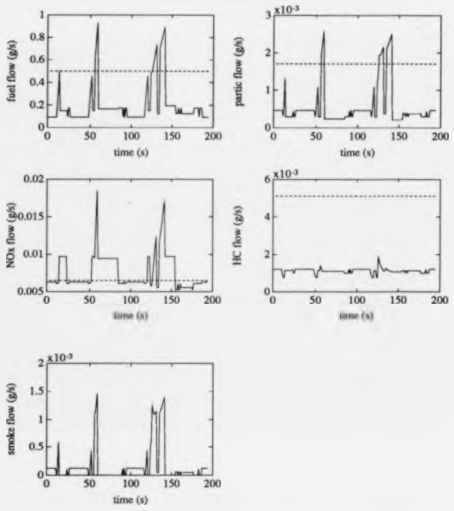


Figure 5.16: Fuel and Emissions For Goal Attainment : Partic. Weighted, Mean Goals, (- optimal flow g/s, - - mean limit level g/s)

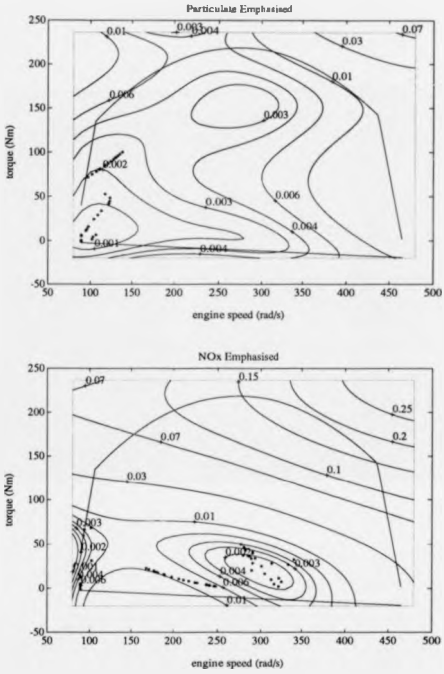


Figure 5.17: A Comparison of Particulate and NO_x Solutions For Mean Goals

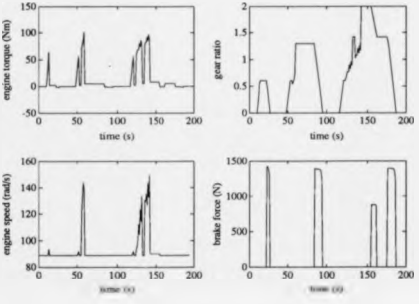


Figure 5.18: Schedules For Goal Attainment : Equal Weights, SO Goals

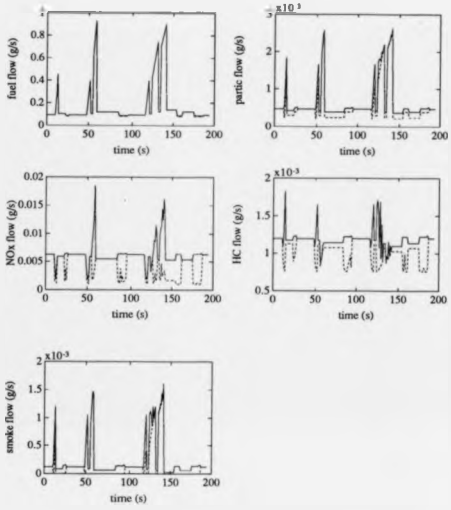


Figure 5.19: Fuel and Emissions For Goal Attainment : Equal Weights, SO Goals, (- optimal flow g/s, - - mean limit level g/s)

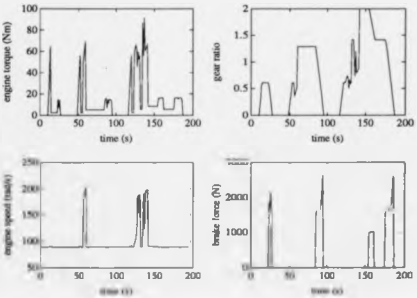


Figure 5.20: Schedules For Goal Attainment Method : NO_x Weighted, SO Goals

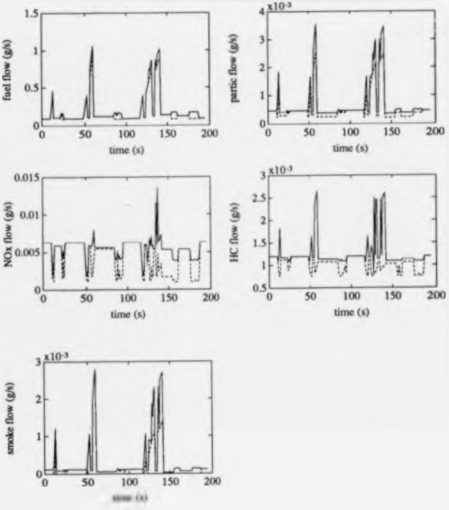


Figure 5.21: Fuel and Emissions For Goal Attainment : NO_x Weighted, SO Goals, (- optimal flow g/s, - - mean limit level g/s)

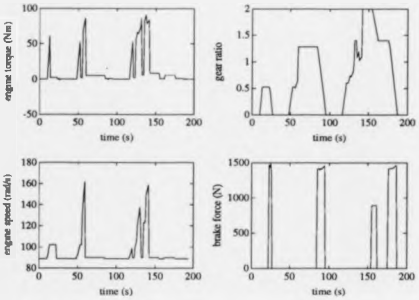


Figure 5.22: Schedules For Goal Attainment : Partic. Weighted, SO Goals

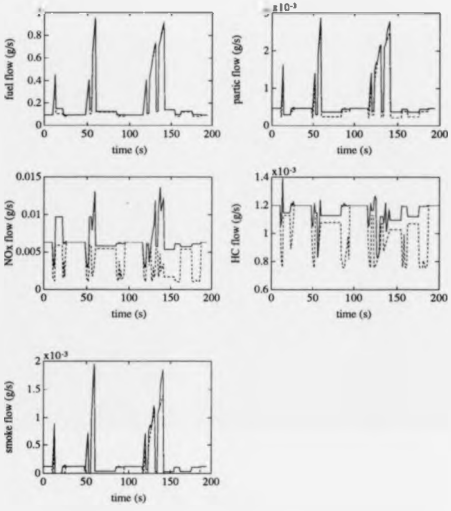


Figure 5.23: Fuel and Emissions For Goal Attainment : Partic. Weighted, SO Goals, (· optimal flow g/s, - - mean limit level g/s)

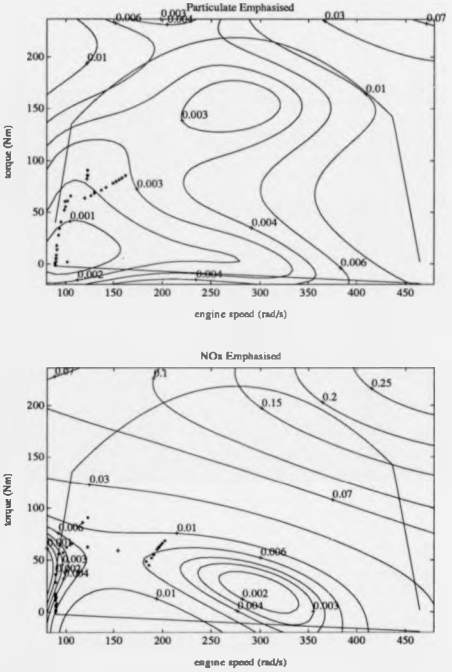


Figure 5.24: A Comparison of Particulate and NO_x Solutions For SO Goals

Chapter 6

Optimal Powertrain Schedule Implementation Using Rule-Based Control

6.1 Introduction

The traditional approach to control system design has become increasingly inadequate for the proper formulation and solution of the type of problems arising in modern technology. This is particularly so in the field of automotive control, where the dynamics are fast and models nonlinear, [95]. The aim of this chapter is to illustrate the effective use of optimal control schedules for the development and evaluation of practical control system designs.

This modern control theory approach can be described in three steps: formulation of a mathematical model, including the essential dynamics, constraints and objectives of the physical problem; optimisation of the control strategy with respect

to relevant performance criteria; and synthesis of the optimal control design. The model was introduced in Chapter 2, its relation to the optimisation process discussed in Chapter 4, and feasible methods of optimisation discussed in Chapter 5.

More specifically, Chapter 5 introduced the Goal Attainment method and illustrated its use in producing effective solutions to the scheduling problem. Here, the flexibility of the method is extended, to show how a greater degree of specification can be given to the optimisation objectives. Section 6.2 discusses the specific design requirements. Allowing different weighting values to be chosen at different points along the test cycle emulates a more realistic scheduling design problem. How this optimal schedule is used, is covered far less widely in the literature than the optimisation process itself. Few authors consider its implementation in control terms. This is not the case here, and it is the aim of this chapter to discuss the mapping of the optimal schedules into a control environment. Section 6.3 shows that splitting of the schedules into separate, well defined, regions results in a feasible rule-based control environment.

In order to test the feasibility of the resultant control laws, a dynamic model is used to simulate driving conditions over the relevant test cycle. The simulation allows incompatibility between the finite-dimensional optimal demands, and the dynamic state of the vehicle, to become apparent. This together with modelling improvements is discussed in Section 6.4. The simulation also measures the performance of the control laws against that of the optimal schedules, and allows an assessment of the driveability of the schedules, Section 6.4.4.

The generality of the control scheme is also a very important consideration. The U.K. automotive industry, in an international market, must take into consideration worldwide legislation. As already seen, different countries have different test pro-

cedures and legislative limits. It is thus important to know whether an optimal schedule is applicable to many test cycles or whether a separate optimisation process must be undertaken for each. All results up to this point have been directed towards European procedures and limits. Section 6.5 however, tests the transferability of the control scheme to the American EPA test cycle. That is, does the control scheme track the new test cycle? and, more importantly, does it retain the primary objectives of the optimisation?.

Finally, some conclusions are drawn, in Section 6.6, on the effectiveness of the control design procedure, and its viable implementation.

6.2 An Optimisation Specification

This section develops the Goal Attainment ideas discussed in Chapter 5, and directs them more specifically at a scheduling design problem. Results were shown, in Chapter 5, for two possible goal specifications : mean limits and single-objective minima. It was said that, whilst use of the mean goals presented problems concerning weighting choice, this could be overcome if a more detailed weighting specification was used, and that the approach was essentially more flexible. Rather than a blanket weighting over the whole cycle, this section produces optimal schedules which are the result of assigning different weightings to different points on the test cycle. Doing so, brings together the many results of the previous chapter and is an approach more akin to the needs of the engineer. It is not unrealistic that a designer may feel that acceleration manoeuvres, for example, should be given more importance than periods of cruise.

The design problem considered here can be stated as follows:

- Minimisation of particulate output is of greatest important,
- If particulate output can achieve its limit it should do so,
- Hunting problems and the use of fuel should be minimised during deceleration manoeuvres,
- The compromise between all objectives should be retained.

Although quite a complex specification, the result will be a mixture of the type of weighting choices applied in the previous section. The only hard specification is that particulate should meet its limits where possible. This dictates a very strong weighting in these areas and requires some prerequisite analysis of the previous results. The remaining weightings are quite flexible and it is the job of the designer to transform these demands into weighting choices which will produce the desired result.

The weighting choices which reflect the specification above are shown in Table 6.1. The main consideration is that of particulate minimisation. Thus, over the constant velocity phases, the particulate emphasised weighting choice of the previous chapter, shown in Table 5.5, is used. The results of Table 5.5 showed that an equal weighting choice gave good output results for all objectives. Thus to retain some of the compromise amongst all objectives, *i.e.* no one function does really badly, equal weightings are placed on the functions over the acceleration phases. Analysis of the schedules however, and the fuel and emission outputs of Figure 5.12, show that the particulate output is above its mean limit during the high velocity/acceleration phases. Closer analysis of these phases shows that some points could achieve the particulate limit but are not doing so, *i.e.* the second acceleration phase of the third portion of the ECE-15. This is seen as a high peak above the goal value around

the 125s point. The same weighting is being applied over the whole manoeuvre. Since the weighting is specifying equal under or over-achievement of the functions, the particulate output is under achieving at the expense of some other function. A weighting is thus given in this region which stipulates strict achievement of the particulate goal, as shown by the second set of weights in Table 6.1. The zero weight indicates strict achievement. Although only a minor point of the cycle, doing this illustrates the potential of the method and gives a feel for the sensitivity which can be achieved. It also illustrates the close interaction between the designer and the optimisation process.

Thus far, the compromises of the equally weighted solution have been retained over the acceleration portions, and some emphasis has been given to the particulate output in specific regions. Not dealt with yet is the hunting, seen in both the equally weighted and particulate emphasized optimisations of Figures 5.11, 5.15. This is a deceleration phase problem and can be solved by allocation of specific weights to this phase. The hunting is brought about by the existence of two local minima and the the trade-offs being made as the solution jumps between them on the engine map. The specification asked for particular attention to fuel in this region. Some fuel systems actually cut off during overrun, a feature not modelled here. The trade-off is such that the optimal result is to use the braking force at the wheel, rather than engine braking, to produce the required velocity reduction. Since the optimiser therefore leaves torque in the positive region of the engine map it is feasible to request an emphasis on fuel in this area. The resultant weighting choice in Table 6.1 shows the greater weight given to the fuel function, whilst all other objectives have approximately equal weights.

6.2.1 The Optimal Schedules

The schedules which result from the weighting choice discussed above, and the functions with respect to their goal values, are shown in Figures 6.1, and 6.2. The goals are the mean legislative limits of Table 5.5 and the position of the solution on the engine map is shown in Figure 6.3. The total output values over the cycle are shown in Table 6.2 together with a repeat of applicable results from Chapter 5. The subtleties of the results are best seen when compared to the results of Chapter 5, and more specifically the schedules of Figures 5.11 through to 5.16.

With regard the output results, it can be seen that the emphasis on particulate has been retained, the value not changing from that of the particulate emphasized optimisation of Chapter 5. All the other functions have reduced slightly from these values, due mainly to the equal weighting over the cruise phases. The result is thus still a good compromise solution. The behaviour over specific portions of the cycle can be seen in Figure 6.1. The emphasis on particulate around the 125s point has brought the engine torque and speed down in this region. The effect on the function values is shown in Figure 6.2. The peak in particulates above the goal value has been brought down to and below the goal. As a consequence, the fuel smoke and NO_x outputs have reduced, whilst the HC output has increased. Other acceleration and cruise phases remain as in their respective diagrams in Chapter 5. The similarity ends in the deceleration phase. It can be seen that much of the hunting has been removed from the engine torque schedule, especially when compared to the particulate emphasized solution of Figure 5.15. As a consequence the fuel during the deceleration phases has reduced from $4.79g$ to $4.57g$, roughly a 5% reduction. A greater weighting on fuel would produce an even greater reduction. The general result is thus low torque and engine speed, high gear ratio and a moderate amount

of braking.

Compared to the particulate engine map of Figure 5.17, the result is very similar. Some spreading of the points has taken place around the 50Nm region due to the greater range of specifications now brought together into one design. The result is a retention of the particulate schedule characteristics, with added specifications at little cost.

In this section, it has been shown how the Goal Attainment method can be used in practice. Optimal schedules for the fuel and exhaust emission minimisation problem have been produced subject to a complex design specification, achieved through manipulation of the weighting parameters over different portions of the test cycle. The following sections will show how the optimal schedules produced can be transformed into a feasible control system design.

6.3 Control Law Design

Having obtained a schedule which complies with specific design requirements, the next step is to transfer these optimal schedules into control laws. Two general methods have previously been used in the automotive sector to accomplish this. The first is through the use of regression techniques. Prabhabkar [33], used multi-variable regression to provide laws for spark advance and A/F from his optimisation results. Tennant *et. al.* [12], [36], used the same procedure over a statistical data base description of the test cycle, by breaking the data base into segments and fitting curves to each. Radtke *et. al.*, [21], fitted first order lines to shift patterns as functions of driveshaft torque and speed, whilst Kuzak *et. al.* [19], approximated the optimal shift pattern as a function of throttle angle and driveshaft speed. The control

laws are thus dependent on two input variables for which sensors must be available. This approach is feasible as long as the laws are simple and the resultant optimal equations of low order. A second approach, only suitable to certain applications, is to formulate laws through visual analysis of the schedules, *i.e.* simple rule-based control laws. Anti-lock braking systems, whose structure can be simply represented as,

if *wheel is slipping* then
brakes off

are typical of this type of control. This bang-bang approach to control is not sophisticated enough for the powertrain scheduling problem. However, a combination of this and the regression type of approach can produce feasible control laws. The first step in such a process is recognition of the available sensed inputs.

In Chapter 4, the scheduling process was described in terms of the optimisation algorithm. The inputs were shown to be the demands of the test cycle, *i.e.* the instantaneous velocity and acceleration. It was also shown how these replaced the inputs of the driver under normal operating conditions. Both open and closed loop control schemes were described, the inputs coming from an instantaneous demand or predicted demand respectively, Figure 4.2. This chapter considers the open-loop control scheme only.

Although velocity has long been a recordable variable for control, the acceleration measurement is more difficult. An attempt has been made therefore, to extract a relation between the velocity demand and control variables only, whilst still encompassing the optimisation results. Figure 6.4 shows the optimal schedule points for engine torque, gear ratio and brake force plotted against demanded velocity. On

doing this, it becomes apparent that some relationship does exist. If the solution points are separated into acceleration, deceleration, and cruise phases, as shown on the diagram, the relationships become even clearer. For each mode of travel, linear and low dimensional polynomial relationships can be fitted to the data, using least-square techniques in the MATLAB environment. Visual analysis then determines the rule-base under which the regression equations act.

The approach here is thus to take the optimal control variable outputs of the finite-dimensional Goal Attainment optimisation of Section 6.2.1 and to formulate a rule-base which supplies optimal engine torque, gear ratio and brake force control variables as a function of vehicle speed only. The relationships are derived through regression analysis in the form of low order polynomials, together with a rule-base strategy. This strategy is centred on the different modes of vehicle acceleration, deceleration and cruise together with a prescription for switching between them.

The result can be seen in Figures 6.5 and 6.6, and is a piecewise collection of low order control rules, similar to many fuzzy logic control schemes. It can be seen that most points within the individual phases, can be described by first order straight lines; others are described by polynomials up to a maximum of third order, this maximum being set by the limits of microprocessor implementation. The control laws thus provide simple laws for which engine torque gear ratio and brake force variables could be modulated in a drive-by-wire system.

In some cases, decisions on when to switch and the nature of the regression will not be obvious at this stage. Spreading of the optimal points can mean that decisions have to be taken over their relative importance, that is, which should be reflected in the control law, and which should not. The optimal results are dependent on the design specification. The greater the specification, and the more

acceleration dependent it becomes, the more likely spreading is to occur. Like all physical processes there is some compromise to be taken into consideration. In this case, the control law chosen may reflect an improvement in vehicle tracking at the expense of regression error. The reason for the decisions taken will become clearer in Section 6.4 when simulation of the control laws is undertaken.

The engine torque control laws are shown in the top diagram of Figure 6.5. The deceleration control law is a simple switch between three constant values, dependent on vehicle speed. The switches take place at 1.0, and $6.6m/s$, the torque levels being 0, 6.2 and $11.84Nm$. From the original schedule, Figure 6.1, it can be seen that movement between the two more prominent torque levels during deceleration, produces little movement in either fuel or particulate output. The stepping down of the torque, as the velocity reduces, stops the sharp cut off of larger torque values at low velocities which could set up compliance oscillations in the driveline. The constant velocity curve is a simple switching between two first order lines, the switching taking place at $9.0m/s$. A single straight line was placed through the points as a first pass solution but it was found that the two separate control laws produced better tracking, as will be seen in Section 6.4. As the cycle stands, there are not enough constant velocity solution points to either support or contradict this decision.

The regression of the acceleration solution points is more difficult. The points are more spread out, the torque values between 5.0 and $9.0m/s$ being particularly troublesome. In this region, two quite separate first order laws seem to exist. The reason for this is the specification placed on the particulate output during the third phase of the test cycle. A similar velocity range exists on the second phase but at a much higher acceleration value, *i.e.* the upper straight line in Figure 6.5 Some compromise must thus be made in this region. The result is a three portioned

control law. The first portion is a second order fit to the points up to and including $6.0m/s$. This takes in some of the lower line. At this point however, a switch is made to a first order line which encompasses the higher torque points for a small interval. Above $7.0m/s$ control passes to a first order line which again encompasses some of the lower line and the remaining higher velocity points. This jump to higher values for a short period allows better tracking of higher accelerations and would be equivalent to a quick depression of the accelerator pedal to pick up speed rapidly. However, most emphasis is given to the lower torque values since these are linked to the specification.

The gear ratio control laws are shown in the bottom diagram of Figure 6.5. The constant velocity control law is made up of two first order lines as in the torque demand case, the switch occurring again at $9.0m/s$. The deceleration control law is a simple second order line which fits the data to a good degree. The acceleration law is simpler here than its torque counterpart and has been regressed to a second-order line.

The regression of the remaining control variable, brake force, is shown in Figure 6.6. There is quite a spread of points due to the shallowness of the constraint surface in this direction, Figure 4.13. Again some decisions have been made with prior knowledge of tracking ability. The control law is made up of four piecewise linear portions which reduce the brake force as the velocity increases. The first, third and fourth sections incorporate constant force levels while the second follows a first order line.

The schedules produced through the control laws just discussed and the original optimisation output are compared in Figures 6.7 to 6.9. That is, the upper figures in each are the optimal schedules as output from the Goal Attainment optimisation of

Section 6.2.1, whereas the lower figures are the schedules produced from a dynamic simulation of the rule-base control laws .

The torque control law, Figure 6.7, smooths out the jumps during deceleration and reaches a peak value during deceleration of the same order as the optimal schedule. The controlled gear ratio schedule, Figure 6.8, is again smoother than the optimal but the schedules are very similar. Finally, despite the more complex set of laws describing the brake force the controlled schedules resembles closely the optimal, Figure 6.9.

Thus, the engineer now has at hand simple rule-based control laws which mimic the optimal schedules and are implementable on an on-board microprocessor. The technique has incorporated both visual analysis and low order regression and is a viable method for control law design. The laws are based however, on finite-dimensional optimisation, and data, and how the control laws perform under a dynamic simulation will now be discussed.

6.4 Dynamic Simulation of the Powertrain Under Rule-Based Control

6.4.1 Modelling Inclusions

Chapter 2 introduced a compliant dynamic model for which simulation to a step input was shown. The important dynamics were the inclusion of the engine flywheel, the gearbox dynamics and the compliance effects. The nature of the compliance was illustrated in Chapter 2.

The gearbox dynamics are such that the transmission is able to store and move

energy across the gearbox. The delay in actuation was modelled as,

$$\dot{x}_2 = \frac{(-x_2 + u_2)}{T_c} \tag{6.1}$$

where x_2 is the actual gear ratio, and u_2 is the demand. Figures 6.10, 6.11 and 6.12, show the magnitude of the dynamic effect for $T_c = 0.1$, 0.5 and 1.0, over a typical simulation of the test cycle. During accelerations the actual gear ratio lags the demand and $\dot{x}_2$ is positive; during deceleration the reverse is true. Increasing the time constant increases the delay between demand and response. The magnitude of the derivative $\dot{x}_2$, generally decreases and becomes smoother. The relative effect of the gearbox dynamics is therefore greater for smaller time constants. The time constant of the C.V.T. modelled here is 0.5, allowing a complete pass through the gearbox in approximately two seconds.

Unlike the simple step inputs of Chapter 2, actual vehicle motion, and even motion of the test cycle, is more varied. The ECE-15 test cycle also includes periods of rest during which the engine speed must be brought down to idle and the vehicle speed brought to rest. In order to handle this adequately, two additional effects must be included in the dynamic model. The first is a change to the way the road load R_l is modeled, equation 2.10.

Road load, a function of velocity, comprises of aerodynamic drag and rolling friction, and is modelled to act at the tyre/road surface. The constant term A is representative of the energy loss in the tyres. When the vehicle achieves zero velocity the road load equation becomes,

$$R_l = A, \tag{6.2}$$

a positive constant. In practice, the energy loss term disappears when the tyre

stops rolling. The retention of the positive term at zero velocity is thus a modelling anomaly, and to overcome it the road load is now defined by,

$$\begin{aligned} &\text{if} \quad v_a < \epsilon \\ &\qquad R_l = 0 \\ &\text{else} \\ &\qquad R_l = A + Bv_a + Cv_a^2, \end{aligned} \tag{6.3}$$

where v_a is actual velocity achieved by the vehicle. The constant ϵ is set, through experimentation, to be sufficiently small. In this case it is $0.01\,m/s$. The effect of this can be seen in Figure 6.13. By the nature of the C.V.T., the vehicle velocity does not settle to zero immediately. As the driveline torque oscillates, the motion, when translated through to the road surface, causes the road load to switch in momentarily.

The second modelling addition is the inclusion of engine friction. The engine flywheel now has an inertia which will resist changes to motion. Engine friction, the main part of which comes from the action of the piston, acts to decelerate the flywheel when no load is applied. Friction data applicable to the specific engine of this project was not available. However, the friction coefficient used in the model is a scaling of that for a Ford Continental [49], and corresponds to a deceleration from maximum to idle engine speed in $1.35s$, and has a value of $0.99 Nms/rad$. The engine torque u_1, modelled as a peak torque and overrun curve from engine test bed data, will have friction effects included over the modelled region. It is only the area around zero torque therefore, for which the effect is lacking. Hence the friction is

only switched in as follows,

$$\begin{aligned} \text{if} \quad u_1 &> u_{fric} \\ u_{1_c} &= u_1 \\ \text{else} & \\ u_{1_c} &= u_{fric} \end{aligned} \tag{6.4}$$

where

$$u_{fric} = -\mu_{fric}(x_1 - x_{min}) \tag{6.5}$$

and u_{1_c} is the simulated torque demand. Thus, if the optimal demanded torque is less than the calculated frictional force, the latter is dominant. The effect can be seen in Figure 6.21. The frictional force switches in as the engine torque drops and decelerates the engine flywheel to its idle value.

The model now stands as a four state, three control input model, with compliance and frictional effects. A full dynamic simulation of such, with optimal control inputs, is discussed in the next section.

6.4.2 Dynamic Control Law Simulation

This section takes the optimal control laws and simulates them using the ACSL simulation package. The aim is to highlight both any mismatch between the finite-dimensional schedules and the result of a dynamic simulation of such, and to gauge the optimality of the resultant control laws. The time constant of the C.V.T. modelled here is $0.5s$, roughly equivalent to a complete pass through the gearbox in

two seconds. It is believed that an experienced driver can pass through a manual gearbox, ratio 0-1, in approximately four seconds [79].

Figure 6.14 shows the optimal demand inputs u_1, u_2, u_3. The engine torque, in this case, is the demand since friction affects are not switched in. The resultant state variables over the ECE-15 cycle are shown in Figure 6.15. The effect on the engine speed is quite different to the finite-dimensional engine speed response, Figure 6.17. The gear ratio response, driveline deflection, and driveline torque are all reasonable given the inputs. The demands are such, however, that the engine speed is decelerated far beyond the idle speed region during deceleration manoeuvres, so much so that the result is infeasible. The reason is the imbalance of the dynamic response to the steady state demands. The gear ratio reduces so quickly that the engine speed is decelerated rapidly; the compliance in the driveline is unable to restore the engine speed, which as a result plummets. The imbalance between the finite-diemensional and dynamic cases is due mainly to the ratio lagging behind the demand. The one second approximations to the test cycle in the finite-dimensional case will also introduce some error. The imbalance can be shown more clearly by a closer look at the equations. Engine speed is determined by the equation,

$$(I + x_2^2 I_g)\dot{x}_1 = u_{1_e} - x_2 u_g - I_g x_1 x_2 \frac{(-x_2 + u_2)}{T} \tag{6.6}$$

The modelled torque demand u_{1_e} must be large enough to balance the dynamic driveline torque u_g and the torque due to the momentum change across the gearbox. This is shown more clearly in Figure 6.18. The driveline torque dominates the optimal engine torque demand during deceleration and itself reflects the balance between gearbox output torque and external forces translated through to the driveline.

The difference between the steady state and the dynamic case can be illustrated as below. Assume that the engine torque demand is large enough to maintain constant engine speed. From equation 6.6,

$$u_{1_c} = x_2 u_g - x_3 x_1 I_g \frac{(-x_2 + u_2)}{T_c} \tag{6.7}$$

where u_g is the driveline torque and assume that friction effects are not switched in. For convenience define the torque needed to overcome the momentum change as u_m, i.e.

$$u_m = x_3 x_1 I_g \frac{(-x_2 + u_2)}{T_c} \tag{6.8}$$

Assume that the driveline speed, represented by

$$\mu^2 M \dot{x}_3 = u_g - \mu(A + B v_a + C v_a^2) - \mu u_3 \tag{6.9}$$

is constant. The driveline torque needed to balance the external forces is thus,

$$u_g = \mu((A + B v_a + C v_a^2) + u_3). \tag{6.10}$$

From equation 6.7 however, the driveline torque needed is

$$u_g = \frac{u_{1_c} + u_m}{x_2} \tag{6.11}$$

Substituting equation 6.11 into 6.10

$$u_{1_c} + u_m = \mu x_2 (A + B v_a + C v a_2) + \mu x_2 u_3 \tag{6.12}$$

Call the first right hand term in equation 6.12 the dynamic road load torque, R_{l_d} and the second the dynamic brake torque u_{b_d}. Now compare this with the finite-dimensional equations from which the optimal demands are calculated.

$$u_1 = \mu u_2(A + Bv + Cv^2) + \mu u_2 u_3 \tag{6.13}$$

where the static road load, R_{l_s} is based on the demanded vehicle speed v, and the static brake torque u_{b_s} is based on the demanded ratio u_2. Comparing right hand sides of equations 6.12, 6.13, Figure 6.19 shows that the relative magnitude of both road load and brake terms is greater in the dynamic case. During deceleration the actual gear ratio x_2 lags behind the demand u_2 and thus is greater in magnitude. Thus the dynamic brake torque is larger than in the static case. Despite the lag in the velocity, Figure 6.16, the difference in ratio demand is enough for the dynamic road load to be greater than the static road load. The result is that the torque needed to balance the driveline torques is greater in the dynamic case than in the static. In both cases however, the input torque u_1, u_{1_s} is of the same magnitude. In the dynamic case the driveline torque also has a component due to the dynamics of the C.V.T.. In fact, during deceleration the term u_m is negative, Figure 6.11, detracting again the input torque's ability to counter the dynamic effects. In this case no friction torque was included, but by its very role, this will again reduce the torque input during deceleration. The overall effect will be the large decrease in engine speed seen. The comparisons here have been made in steady state: during transient manoeuvres when energy is needed to move from point to point, the imbalance is even greater.

It would seem therefore that for the model in question, the steady state optimal

schedules are inappropriate to dynamic control. However, in practise C.V.T. vehicles contain a clutch or torque converter between the flywheel and gearbox to decouple the engine from the driveline. This removes the ability of the C.V.T. to decelerate the engine at the expense of winding up the driveline. For simulation purposes the degree of imbalance will depend on the optimal schedules chosen, the regression of the control laws, and the accuracy and level of approximation of the model itself. The effect is only seen here during deceleration manoeuvres and as shown in the next section improved control, and more realistic modelling can overcome the imbalances which are taking place.

6.4.3 Inclusion of a Governor

Diesel engines use either mechanical or electronic governors to control the speed of the engine. The minimum-maximum speed governor, as its name suggests, acts only at idle or when the engine reaches its maximum speed. Inclusion of this type of governor into the dynamic model, makes it more realistic, and overcomes the infeasible engine speed problems being incurred. Based on governor modelling elsewhere [49], the action only, and not the dynamics of the governor, are modelled. It is assumed that the governor is implemented electronically and that certain control variables are available. Its action provides a variable gain feedforward loop around the torque demand in order to control engine speed. Figure 6.20 shows a schematic of the control system.

If the engine speed drops below idle, $x_{1_{min}}$ or rises above the maximum, $x_{1_{max}}$, the governor feedforward control is switched in. It compares the engine torque demand against the driveline torque reflected through the gearbox and assumed available. The difference, passed through a gain k, is added to the optimal torque demand.

This effectively produces a new torque demand which is able to at least balance the dynamic effects. The control system shown is a variable gain system since a different gain is needed for idle and maximum speed operations. The gains are $k_i = 0.99$ for idle control, and $k_m = -0.98$ for maximum speed control. The observable variables, u_{1_i} and u_g were chosen so that the new torque demand is still based on the optimal demand. Although a logical approach, this is not achievable with the current state of torque sensors. With technology improving at an ever increasing pace, however, it was felt a justifiable approach to the problem. In practice the inclusion of a torque converter would remove the ability of the C.V.T. to snatch energy from the flywheel and remove some of the feedforward control needed.

The effect of the inclusion of the governor can now be shown. The friction and feedback adjusted engine torques are shown in Figure 6.21 and Figure 6.22. The engine torque demand is replaced by the frictional force during low torque demands. The governor control loop is then applied to the modelled torque demand to produce the new torque shown in Figure 6.22. The effect of the control is to increase the torque above the modelled demand in the deceleration regions. The complete set of control demands is shown in Figure 6.23. The resultant state variables are shown in Figure 6.24. The engine speed now resembles the optimal finite-dimensional torque trajectory of Figure 6.1. The engine speed has been brought up from below idle by the feedforward control. The frictional force then reduces it to idle as the vehicle speed drops to zero. The driveline deflection and torque shown are the result of transfering the engine torque and speed through to the gearbox output.

The test of the control laws is the ability to track the test cycle and to retain the optimality of the fuel and emission output. The velocity tracking is shown in Figure 6.25. The control laws follow the test cycle especially well during low

velocity movement and deceleration. The acceleration is seen to deteriorate slightly at higher velocities due to the nature of the regression and the compromises as discussed in Section 6.3. The result is that the tracking under-achieves the constant velocity demand on the second portion of the cycle by $0.75m/s$ and over-achieves the higher velocity portion by a maximum of $1.0m/s$. Standard tracking error during legislative testing is of the order of $\pm\ 0.9m/s$, or $\pm\ 1s$ [79]. The control laws therefore adequately track the test cycle.

The fuel and emission outputs over the cycle are shown in Figure 6.26. Compared to the optimal outputs of Figure 6.2, all functions are very similar. With regard to the specifications, the particulate output, which was deemed the most important, is very close to its optimal value. The reduction to the limit around the $125s$ region has not been met, in this case, because of decisions taken over vehicle tracking. Improvement on this small section would worsen the output and tracking over the second portion of the cycle. However, if this specification was felt to be of the utmost importance revision of the control laws could produce the desired effect. The higher engine speed in the dynamic case during deceleration increases the particulate output in these sections, as it does for the fuel output. However, the last deceleration phase, in which governor action does not take place, produces good fuel output as specified. The outputs of HC and smoke are very close to their optimal values and thus the compromises between the variables are being retained. The higher engine speed during deceleration improves the NO_x flow over these regions. The reason can be seen in the engine map of Figure 6.27. Compared to the optimal solution points of Figure 6.3 the engine map points are still in the region of the particulate minimum, but have shifted slightly toward the idle speed boundary, where NO_x is at an optimum during deceleration, as in Chapter 4.

The overall test cycle figures, and a comparison to the finite-dimensional optimal values, is shown in Table 6.3. The table also shows the offset as a percentage from the optimal value. The control laws have kept all functions within a reasonable degree of their optimal values, and all values are below their legislative limits.

6.4.4 Driveability

Driveability is a somewhat ambiguous term which describes the ride and handling of the vehicle as seen by the driver. Its definition can include many dynamic phenomenon dependent on the context of the problem. Since powertrain scheduling is the consideration here, only issues linked to this will be considered.

There are two main considerations linked to the driveability of a particular schedule on a vehicle. The first is the behaviour of the engine, the second is the demand placed on the gearbox. Related issues such as shunt and noise are important, but outside the scope of this project. The engine speed profile, driven by the optimal control laws, was shown in Figure 6.24. By the nature of the optimal schedule the engine speed needs to operate at low, near to idle values. When driven dynamically the effect is such that a governor is needed to maintain the engine speed above the idle speed point. As mentioned, it is usual to include a torque converter or clutch in the driveline of the C.V.T. . The first smoothes the transient flow of energy to and from the engine, the second decouples it completely. To some extent this could be included in the optimisation process by restricting the gear ratio choice to those that do not drive the engine speed down below a nominal level. Similarly, a simple clutch mechanism could be included. The result is that the schedule would be applicable to a wider range of engines since new engine mounts, noise control analysis *etc.* would not have to be undertaken for each. These are both topics of future work.

Despite the low nature of the engine speed, the trajectory is quite smooth, with few erratic jumps. The largest transient occurs during the high speed/acceleration motion, with the engine speed dropping by approximately $40 rad/s$ in one second, due mainly to the sharp kick-up in gear ratio at that point. The rest of the ratio schedule consists of smooth increases and decreases for accelerations and decelerations respectively. Transients occur mainly at the start of a deceleration because of the different nature of cruise and deceleration constraints. The kick-up in gear at the 120s mark can be explained along the same lines. A sudden constant velocity demand after an acceleration sends the gear ratio momentarily high. The shift demands are such that the first portion acceleration requires a change in ratio from 0 to 0.5 in six seconds, the second a change in ratio from 0 to 1.1 in fifteen seconds and the third, largest acceleration manoeuvre, requires a change in ratio from 0 to 1.9 in thirty seconds, affecting a velocity rise to $13.89 m/s$. These are quite moderate demands.

Downstream of the gearbox the transients set up compliance effects in the driveline, as seen in Figure 6.24. The oscillation of the driveline manifests itself in the smoothness of the vehicle motion. The deflection here is less than 1 radian at all times and the vehicle is brought to rest smoothly as in Figure 6.25.

Overall therefore, the schedules are quite moderate. The main concern is the level of the engine speed which may cause engine roughness. However, refined optimisation and simulation models could overcome this.

The result of the control law study has thus been the formulation of control laws which although simple, are able to reproduce the optimality of the finite-dimensional solutions whilst also producing reasonable tracking of the test cycle. Many of the design specifications have been upheld. Others have not quite been achieved either

because of choices made at the control law design stage, or because of dynamic effects inherent in the the powertrain. Like the optimisation process itself, the control design process is an iterative one, with interaction between the designer and package inherent at all times. Decisions made at the design stage will be reflected in the simulation output and again it is a process of compromise.

The control design and optimal schedules produced have been specific to the European ECE-15 test cycle. An important question is the robustness[1] of the optimal design when used under different driving conditions. To test this, the next section looks at the application of the control laws to an American E.P.A. test cycle.

6.5 Robustness of the Optimal Control Laws

The ECE-15 cycle is a relatively moderate, non-transient, test cycle. It consists of clear phases of acceleration, deceleration, and cruise, the maximum vehicle speed reached being $13.89 m/s$. A large proportion of the cycle is also taken up with idle. The Japanese cycles are not dissimilar, with smooth accelerations and decelerations, and maximum vehicle speeds of 16.7 and $11.0 m/s$ for the 11 and 10-mode cycles respectively. It is assumed that the optimal control design would thus be applicable to the Japanese test cycles. The American FTP-75 cycle however, is a much more transient cycle. The test consists of three different stages : a transient cold phase, a stabilized phase, and a repeat of the first phase run under hot conditions. The overall cycle covers $2477 s$. For the purposes here, only the first stage of the test cycle is used, the maximum speed of which is $25.2 m/s$. The cycle is shown in Figure 6.28 together with the acceleration levels involved.

[1] Robustness here relates to the appropriateness of a control system to test cycles other than that for which it was derived. It should not be confused with robustness in terms of H^{∞} control.

The aim is thus to use the control laws of Figure 6.5 over the extended range of the FTP-75 cycle. Simple extrapolation of the control laws over the new range is inadequate. As Figure 6.29 shows, the gear ratio extends beyond its bounded range. To correct this the ratio is held at its maximum value at speeds of over $20m/s$ and $15m/s$ for constant velocity and deceleration manoeuvres respectively. To improve tracking other modifications were made to the control laws above the ECE-15 maximum speed limit. It is assumed that these do not affect the optimality of the controller to any great extent. The extended torque and gear ratio laws are shown in Figure 6.29, and the extended brake law in Figure 6.30.

The result of a dynamic simulation of the extended control laws for the FTP cycle is shown in Figures 6.31 through to 6.37. The friction adjusted and feedback adjusted torque demand is shown in Figures 6.31 and 6.32. The friction again brings the engine torque down into overrun during deceleration. The action of the governor is again mainly centred around the idle engine speed boundary and increases the torque in this region. The maximum engine speed boundary is also crossed at the end of the first stage and the action of the governor is to reduce the amount of overrun torque demanded. The governor action is bounded such that it cannot take the torque outside approximate bounds of $200Nm$ and $-20Nm$ for the peak and overrun torque curves respectively.

The control inputs are shown in Figure 6.33. The ratio demand reaches its maximum value over a significant part of the cycle. The transient nature of both the gear ratio and brake force demands reflect the nature of the cycle and the constant switch between acceleration and deceleration. The effect on the state variables is shown in Figure 6.34. The engine speed is reasonably well behaved staying at low values except at a number of points. However, the feedforward control is such that

the governor gain is not quite large enough to bring the engine speed above its idle value as effectively for the more transient behaviour encountered. Also the large decelerations of the FTP cycle to rest, drive the engine speed up in these regions. This is especially so at the first deceleration to rest around the 150s region, at which point governor action is required. The large transients taking place in the engine speed are reflected in the driveline deflection and torque trajectories.

The resultant tracking of the extended control laws is quite good however, and is shown in Figure 6.35. The driveline speed shown is reflected through to the road surface to produce the actual velocity in the lower graph. The tracking between the demanded speed (dashed line) and the actual speed (bold line) is generally good. The vehicle over-achieves the maximum demand on the first phase (0-150s), which includes demands equivalent to the third portion of the ECE-15 cycle, and generally under-achieves on the last three phases which contain speed values above that of the ECE cycle.

Perhaps of more importance than the tracking however, is the optimality of the output functions. Figure 6.36 shows the output together with the optimisation goals. Fuel output is above its goal value at all but low velocities. However, the goal was an arbitrary limit based on output values over the ECE-15 cycle. Particulates are above their goal values at speeds higher than the ECE-15 as expected - the top speeds of the ECE cycle could not achieve the limit. The output of NO_x is generally above its goal value except during deceleration, as in the ECE-15 case. HC and smoke have outputs below their goal values at all but a few places. In fact, it is a feature of all functions that where the governor has taken idle speed action and increased the torque demand the outputs of all functions are particularly bad. This is reflected in the total function outputs for the cycle. Table 6.4 shows the simulation output

figures compared to estimated legislative output and goals summed over the cycle. Despite the large rises in functions at governed stages, HC and smoke are well within their goal values. Fuel and particulates are within 7% of their goals. This is despite the engine being taken into lower than idle speed regions, very bad for particulates and fuel. NO_x output is 50% above its goal value for the whole cycle. However, when the cycle output is converted to a gram/mile figure, as in the original EPA legislation, the comparisons are even better. Particulate and HC are well under their legislative values, while NO_x is very close. The gram/mile figures are based on an estimation of the test cycle length.

Figure 6.37 shows the nature of the engine map points. The crudeness of the engine torque bounds and the idle gain problem mentioned has meant that a fair number of the solution points lie outside the operating region. This accounts for the higher levels of function output. The control system is trying to place the engine in the low engine speed/high torque region, as in the optimal case, but the effects mentioned mean that scattering is taking place. Retuning of the controller specifically for the FTP cycle could overcome this problem, as would the inclusion of a torque converter or clutch into the model.

6.6 Summary and Conclusions

This chapter has shown how optimal schedules for powertrain control can be converted into practical control system designs. The approach has been to take the somewhat simpler optimisations in finite-dimensions and to transfer them into control laws for dynamic vehicle operation. A more realistic model at the optimisation stage of the process would have resulted in a extremely complex optimisation and

synthesis. Using finite-dimensional results means that control laws are of low order and implementable in a rule-based controller.

It has been shown that the Goal Attainment method is a flexible process for specifying scheduling design problems, and that the use of mean legislative limits as goal values produces results for which the overall result is below the true legislative value. Carried through to controlled dynamic simulations the results are upheld to a good degree.

The use of the finite-dimensional control laws resulted in some imbalance when dynamic effects were incorporated. However, these have been overcome to a large extent, the results proving the worth of the control laws. The inclusion of a governor to provide control of the engine speed response is a feasible method of engine control to supplement the fuel and emission control system.

The robustness quality of the schedules and the control laws is complicated by the actions of the governor, and the crudeness of the bounds. The gain of the feedforward loop was not as suited to the FTP cycle as it was to the ECE-15 for which it was tuned. Since the gain was tuned to overcome the dynamic affects of the ECE-15 cycle, it is credible that a more transient cycle, such as the FTP-75, would need a different tuning. The overall result is not unacceptable however. The majority of functions were below or near to their goal values over most of the cycle. All functions will have suffered to some extent by movement outside of the working area, NO_x especially so. Those solution points contained within the bounds are in the position of the original optimal engine map, i.e. low engine speed/high torque. The result is due to modelling anomalies and in practice solution points would be clipped to the variable bounds. Also, modifications to the control laws outside the optimised range were done primarily for tracking purposes. The effect on the

optimality of the results, however, has not been too great. The result as it stands is a good compromise solution with promise of even further improvements with better modelling.

The three stage process of optimal fuel and emission design is thus complete. Schedules can be produced which comply with specifications through use of the Goal Attainment method. The resultant schedules can then be successfully implemented into simple control laws which then form part of a rule-based controller. Moreover, within the bounds of the modelling environment the schedules are transferable to other cycles. It was noted that the control law modification was based to some extent on tracking without knowledge of optimisation. Optimisation over the new cycle would not only solve this but lead to even greater advantages. Optimisation over a greater range of which the ECE-15 is a subset would remove extrapolation uncertainties and produce controllers tuned for the whole range.

Region	Weighting Choice				
	F	P	N	H	S
constant velocity	0.01	0.6	0.18	0.2	0.01
acceleration	0.25	0.25	0.25	0.25	0.25
	0.25	0.0	0.25	0.25	0.25
deceleration	0.3	0.18	0.16	0.18	0.18

Table 6.1: Weighting Choices For Design Specification

Optimisation	F	P	N	H	S
equal weights	30.68	0.085	0.889	0.146	0.029
partic emphasized	32.93	0.078	1.077	0.144	0.025
design spec.	32.75	0.078	1.059	0.143	0.024

Table 6.2: Output of Fuel and Emissions Over Design Schedule

(g)	F	P	N	H	S
simulation output	38.0	0.094	1.099	0.143	0.0298
offset from optimal (%)	16	20	4	0	20

Table 6.3: Results of Controlled Simulation

(g)	F	P	N	H	S
simulation output	268.52	0.924	6.20	0.903	0.711
goal value over EPA	254	0.864	3.30	2.59	10.1
output in g/mile	74.80	0.257	1.72	0.251	0.198
EPA legislation g/mile		0.45	1.7	0.8	

Table 6.4: Results of EPA Simulation

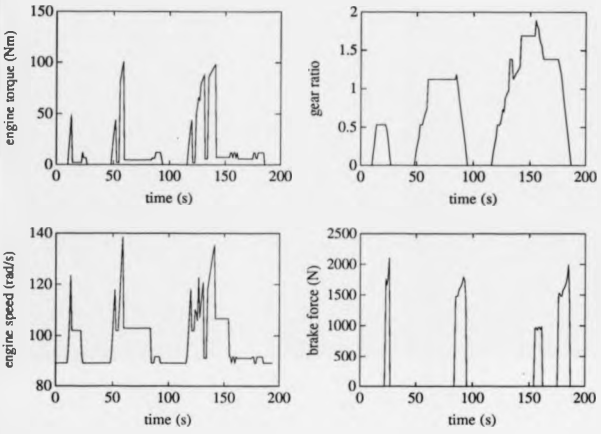


Figure 6.1: Schedule for Control Law Design

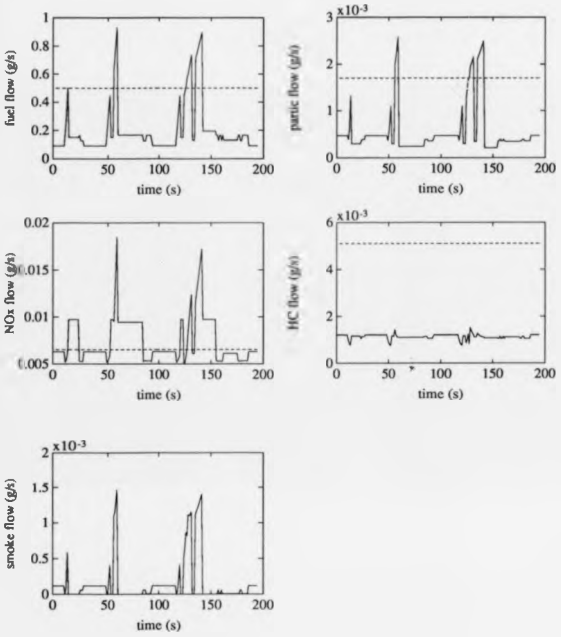


Figure 6.2: Flow Functions for Control Law Design, (- optimal flow g/s, - - mean limit level g/s)

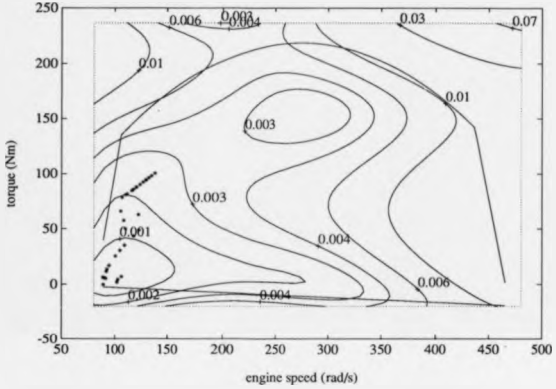


Figure 6.3: Engine Map for Control Law Design Showing Particulate Contours

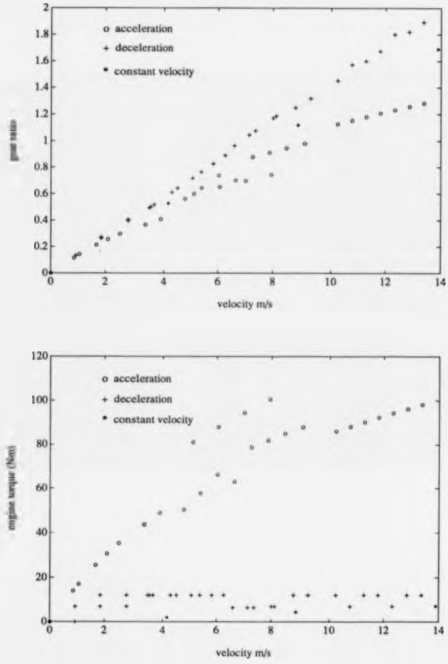


Figure 6.4: Regression Points From Optimal Schedule

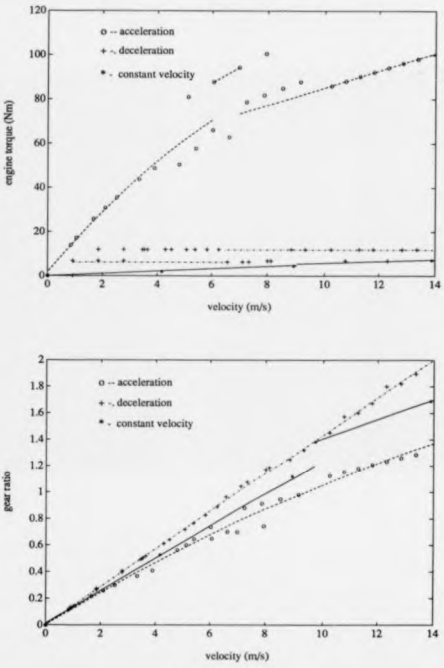


Figure 6.5: Torque and Ratio Control Laws For ECE-15 Cycle

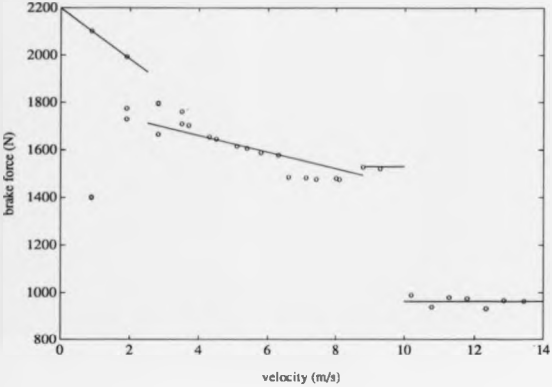


Figure 6.6: Brake Control Law For ECE-15 Cycle

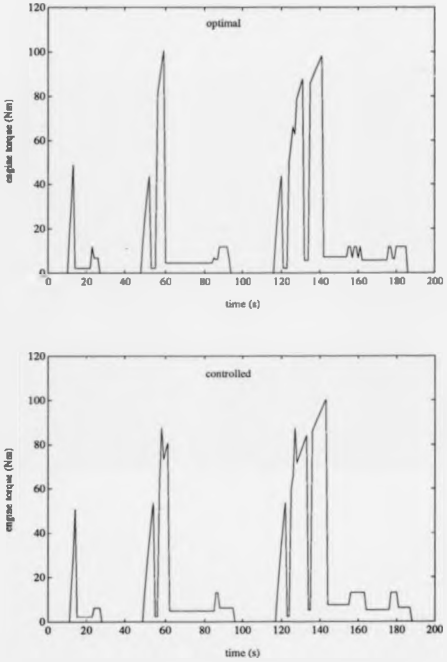


Figure 6.7: Comparison of Optimal and Controlled Torques

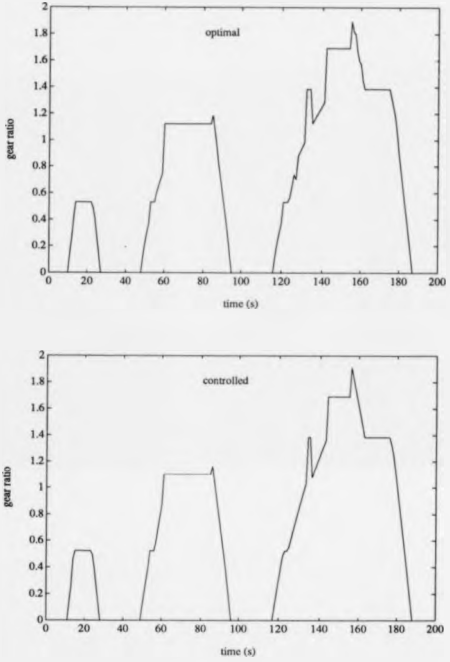


Figure 6.8: Comparison of Optimal and Controlled Ratios

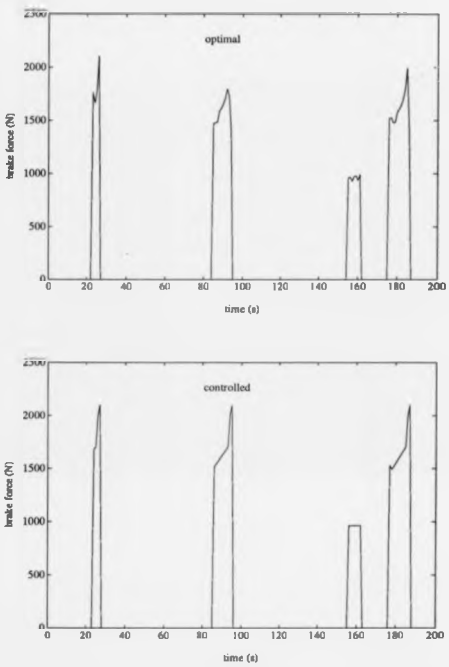


Figure 6.9: Comparison of Optimal and Controlled Brake Forces

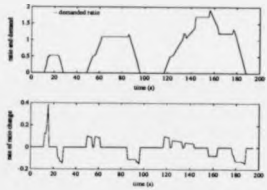


Figure 6.10: Gear Ratio and Demand For T = 0.1

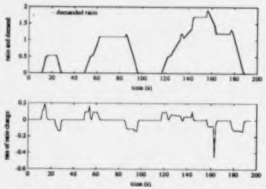


Figure 6.11: Gear Ratio and Demand For T = 0.5

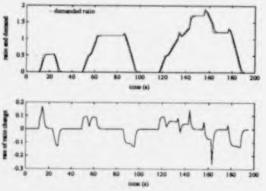


Figure 6.12: Gear Ratio and Demand For T = 1

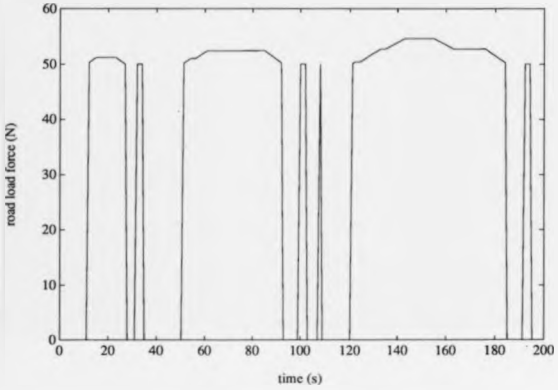


Figure 6.13: Road Load Under Dynamic Model

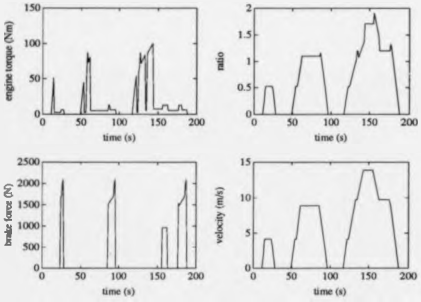


Figure 6.14: Dynamic Simulation of Control Laws For T=0.5 : Demand

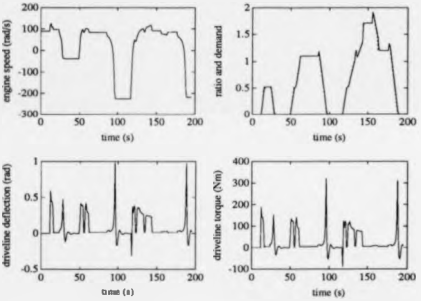


Figure 6.15: Dynamic Simulation of Control Laws For T=0.5 : States

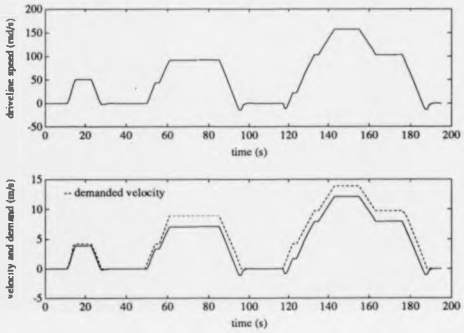


Figure 6.16: Dynamic Simulation of Control Laws For T=0.5 : Tracking

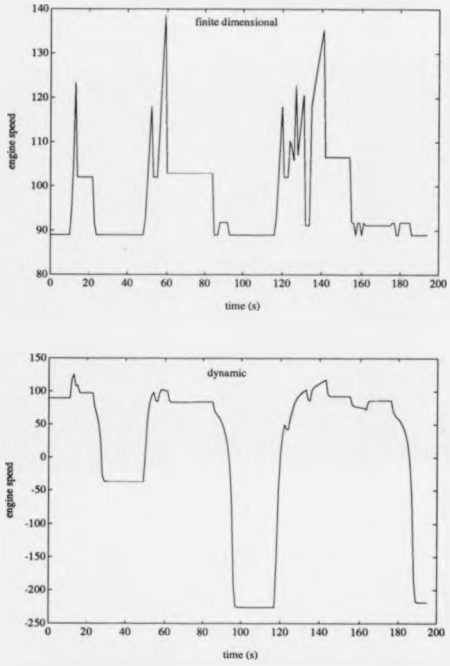


Figure 6.17: Comparison of Steady State and Dynamic Engine Speeds

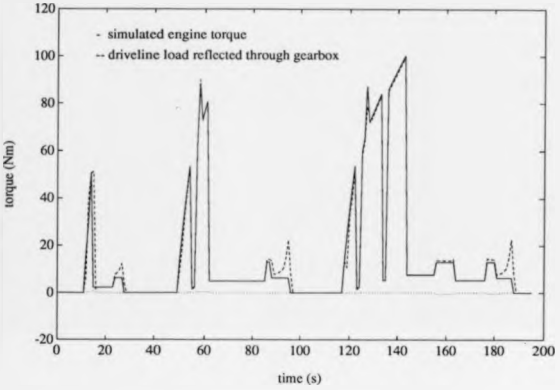


Figure 6.18: Powertrain Torque Balance

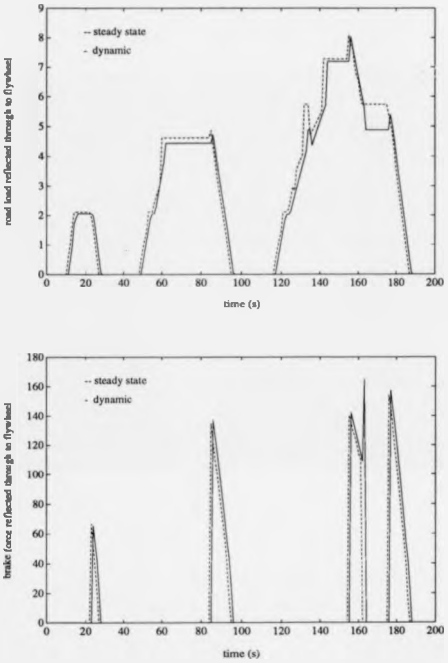


Figure 6.19: Imbalance Between Static and Dynamic Models

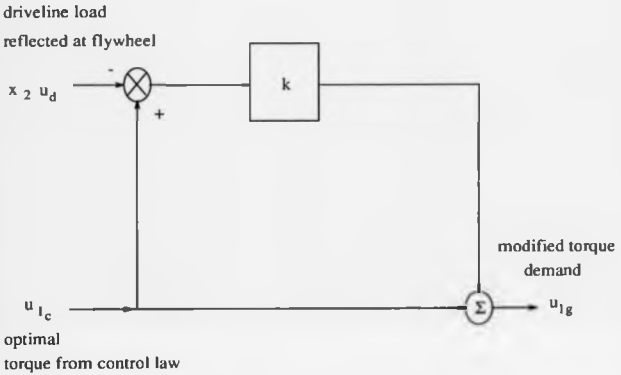


Figure 6.20: Govenor Control Scheme

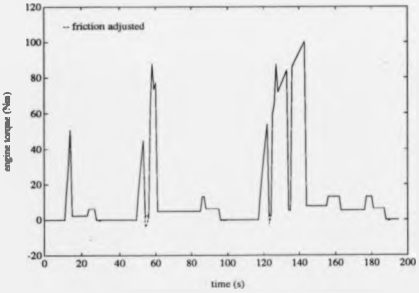


Figure 6.21: Friction Adjusted Engine Torque

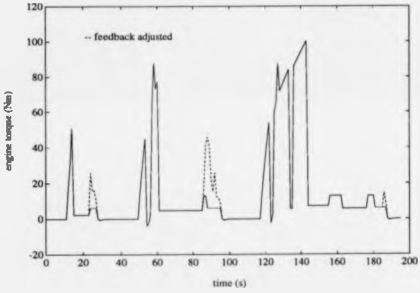


Figure 6.22: Govenor Adjusted Engine Torque

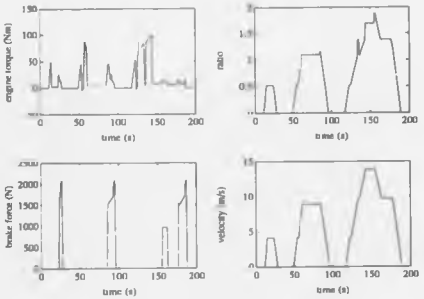


Figure 6.23: Control Input Demands

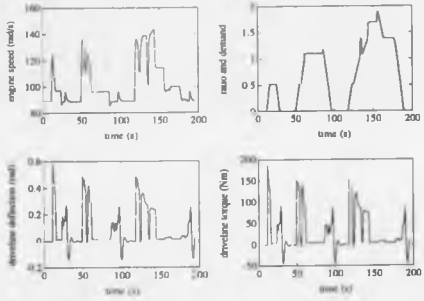


Figure 6.24: State Variables Under Dynamic Simulation

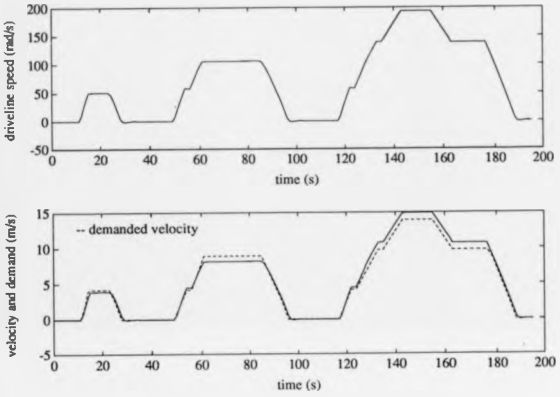


Figure 6.25: Tracking of Test Cycle

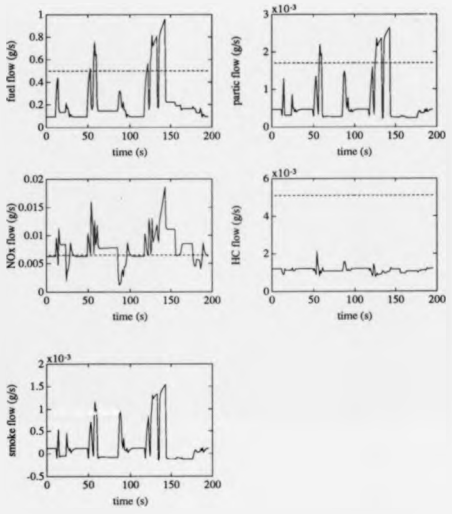


Figure 6.26: Function Output Under Control Laws, (- optimal flow g/s, - - mean limit level g/s)

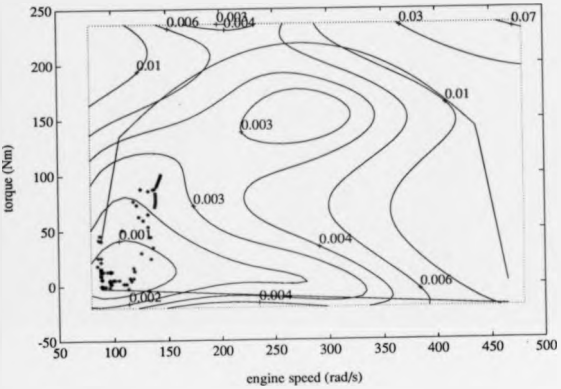


Figure 6.27: Engine Map For Control Law Simulation

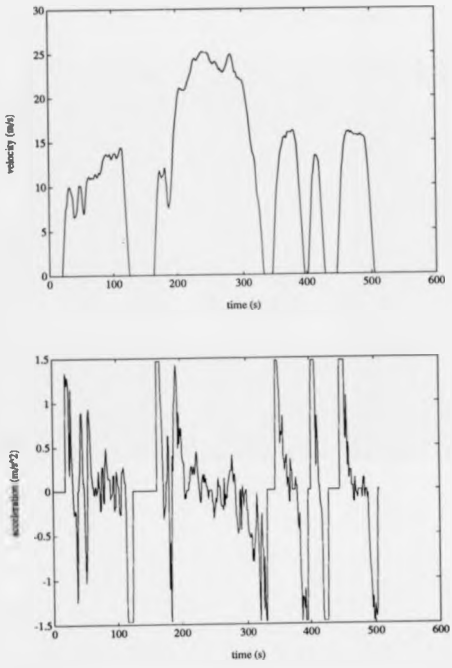


Figure 6.28: Velocity and Acceleration Profiles For FTP-75 Cycle

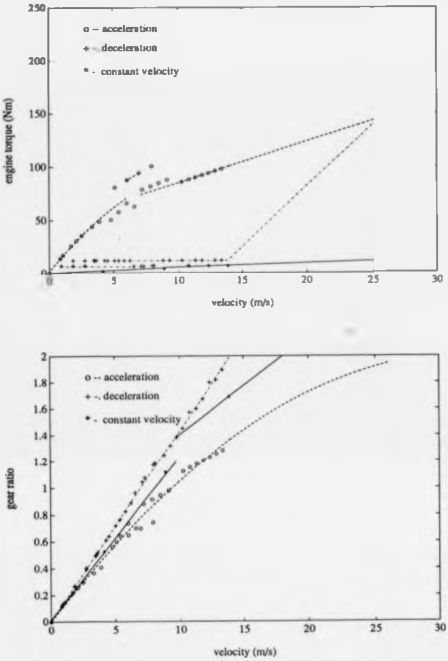


Figure 6.29: Modified Control Law For EPA Cycle

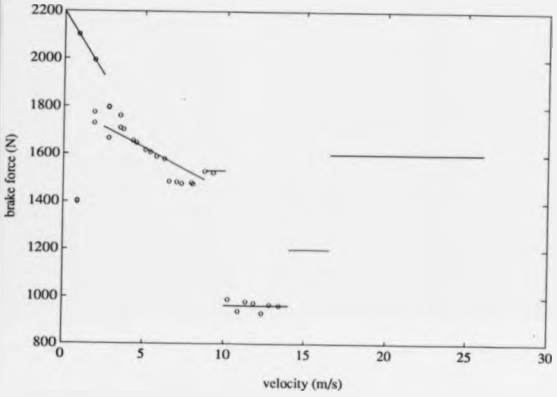


Figure 6.30: Modified Brake Control Law For EPA Cycle

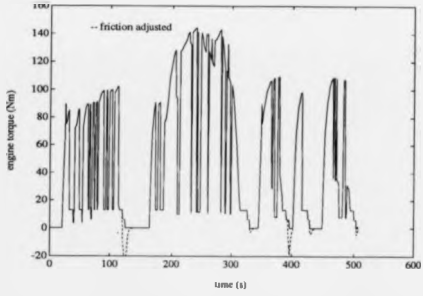


Figure 6.31: Friction Adsjusted Torque Demand Over EPA Cycle

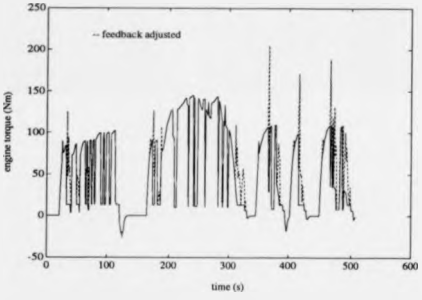


Figure 6.32: Govenor Adjusted Torque Demand Over EPA Cycle

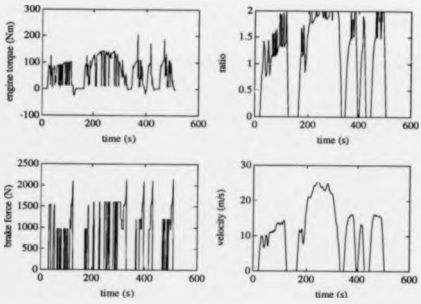


Figure 6.33: Dynamic Simulation of Control Law Inputs Over EPA Cycle

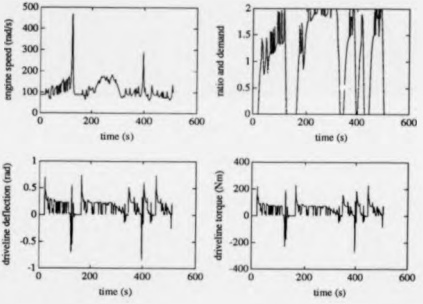


Figure 6.34: Dynamic Simulation of State Variables Over EPA Cycle

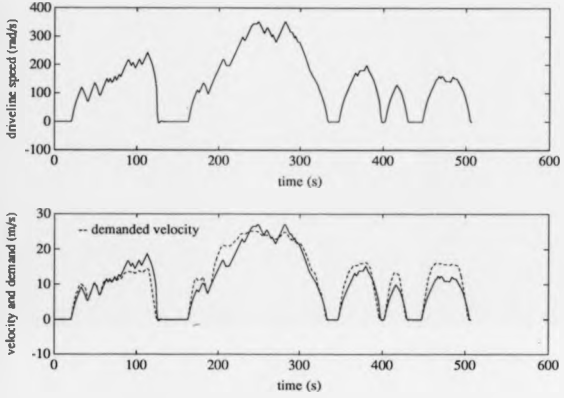


Figure 6.35: Dynamic Simulation of Tracking Over EPA Cycle

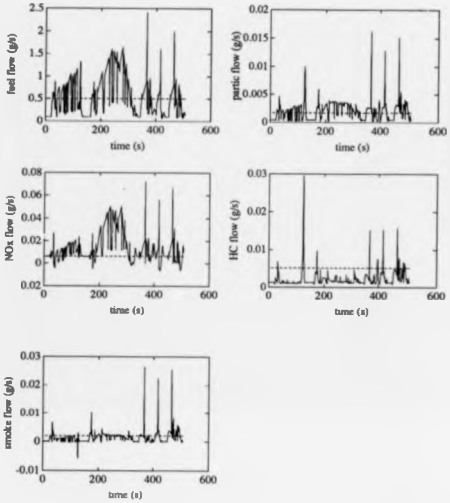


Figure 6.36: Dynamic Simulation of Function Output Over EPA Cycle, (· optimal flow g/s, - - mean limit level g/s)

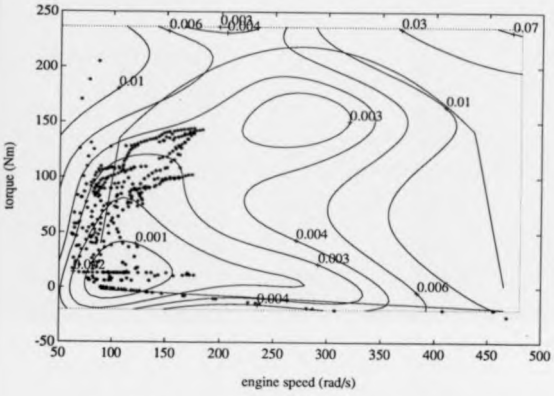


Figure 6.37: Engine Map For Control Laws Over EPA Cycle

Chapter 7

Conclusion

The power of optimisation methods, in engineering, lie in their ability to determine the optimal design at a minimum of cost. The advantage of multi-objective optimisation is its ability to provide a compromise solution in the face of conflicting design requirements. The aim of this thesis has been to present a novel approach to the fuel and emission minimisation problem in an automotive vehicle through the effective use of nonlinear optimisation, and more specifically through the novel use of multi-objective methods applied to powertrain scheduling.

The control of fuel and emission output through powertrain scheduling is a complex and nonlinear design problem, for which the traditional methods of control are inappropriate. The approach here has been a three stage process : formulation of a mathematical model, its essential dynamics, constraints and the objectives of the physical problem; optimisation of the control strategy with respect to the relevant performance criteria; and synthesis of the optimal control design. The interrelationship between these three stages is shown in Figure 7.1.

The foundation of the process is the formulation of a mathematical model. The

optimisation result will only be as good as the accuracy of the model and the correctness of the optimisation formulation. Chapter 2 introduced the powertrain model, Chapter 4 placed it in the context of the optimisation problem. Single-objective optimisation of each function, subject to the constraints imposed, was undertaken, the basic methods of which were discussed in Chapter 3, along with the nature of nonlinear problems. The worth of the single-objective optimisation was seen to be in its role as a first pass solution to the problem, highlighting the conflicting nature of the fuel and emission requirements. However, despite the improved knowledge of the underlying system characteristics which it provided, as an aid to the production of trade-off schedules, it was impractical. The need was clear for a method which simulated the real design environment to a greater extent.

Multi-objective problems arise in all fields of science and finance and considerable attention has been given to the development of methods for their solution. The scope of their application in the field of engineering however, has been limited. Chapter 5 discussed the three basic methods of multi-objective optimisation and applied them to the scheduling problem. Whilst all improved the situation to some extent both the Weighted Sum and ε-Constraint methods were found to be unsuitable. The Goal Attainment method however, essentially combined the advantages of the other methods without also containing their drawbacks. Not only did it represent the simultaneous optimisation of all objectives realistically, but was flexible enough to accommodate the specifications of the scheduling problem. The key to the method is the high degree of interaction between the engineer and the toolbox, as shown in Figure 7.1. The optimisation process is an iterative one, the engineer changing the weights of the functions in order to achieve the desired result. Chapter 6 took this interaction further and showed how the Goal Attainment method could be used to

accommodate a more realistic design specification, different weightings being applied to different portions of the cycle. Although complex specifications were imposed, the result highlighted both the sensitivity of the method and its scope as a tool for the scheduling engineer.

The optimal schedules thus reflect the design specifications laid down by the engineer, and in achieving them, improve his knowledge of the trade-offs which are taking place. The manual comparison needed to combine the single-objective schedules is replaced by a flexible design environment, through which, interaction of the objectives can be investigated. The open loop control laws are such that they require no *a priori* assumptions, as in the case of feedback control, and represent a truly multi-objective design solution. This is a vast improvement over previous studies whose efforts have concentrated on the single-objective optimisation of one function subject to added constraints imposed on the others, [19], [21].

The effectiveness of the optimal schedules is furthered by their successful transference into implementable control laws. Chapter 6 introduced a design process by which rule-based control laws were produced, of low enough order, for microprocessor implementation to be considered. Under dynamic simulation, the control laws upheld both the tracking of the vehicle test cycle and the optimality of the design specification. This was despite some imbalance between the finite-dimensional schedules and the dynamic implementation. The degree of imbalance, as explained, is dependent on many things, not least of which is the degree of modelling. The results of the control law simulation thus have implications on the optimisation model, as shown in the figure, and especially on the constraint formulation. It transpired however, that the problem could still be overcome in this case, without this reformulation, by the inclusion of a governor.

The robustness of the control laws, with regards their application to other test cycles, was also tested, both tracking and optimality surviving well under a considerably more transient test cycle. The conclusion to emerge however, was that an optimisation over the more transient test cycle, of which the other cycles are a subset, would improve the global use of the optimisation results. The compromise may be however, that the control laws become more complicated and thus beyond the scope of controller implementation. The decisive benefit of the control laws, as they stand, is the simplicity of the rule-base and the inherent equations. Only further work can assess whether the second approach would be as viable.

The multi-objective design process presented, thus has the ability to encompass the needs of the engineer into a modular, flexible, interactive design process, links between each stage of the process being apparent at all times. The optimisation process is able to provide information on the optimal control variables; the rule-base is able to implement them in an straightforward manner. Thus, one has the essence of a drive-by-wire control system, capable of making decisions for the driver which will ensure the reduction of both fuel and exhaust emission output. Drive-by-wire throttle and gearbox control systems are already beginning to appear in the market place and the automotive industry's goal of 'total vehicle control' : all vehicle control under one centralised microprocessor, is coming ever nearer.

The limitation of the multi-objective design process is dependent only on the limitations of the modelling process. The data used here has been specific to a diesel engine, for reasons already discussed. The approach however, is equally applicable to petrol and other types of engines for which regression surfaces of fuel and emissions can be obtained. Similarly, other types of transmissions can benefit from the results. Here, the powertrain has included a C.V.T., capable of achieving the true optimal

schedule. Manual, discrete ratio, gearboxs can only approximate this. However, the results of Chapters 5 and 6 show that discrete gear ratio levels could be extracted from the continuous results, to provide scheduling information for manual gearboxes. Similarly, the methods need not be constrained to the scheduling problem. Other automotive control problems, for which there are conflicting demands and suitable models, are equally applicable to the multi-objective design environment, *e.g.* engine management and the manipulation of A/F and spark advance; clutch control and the smoothness of shifts in minimal time; brake control and the minimisation of wheel slip at the optimal brake force.

For the powertrain scheduling problem alone, there is much scope for future work. The model and the constraint specifications can be improved. With ever improving test procedures, the reliability of the data can be enhanced. The design process can be made faster through the more efficient use of package communication. The use of estimation techniques and formulation of a more advanced control scheme would allow prediction of the driver's demand, removing dependence on instantaneous information, and enabling the controller to act accordingly.

Finally, in the same way that a predictive control scheme would incorporate movement of the vehicle from second to second, dynamic optimisation of the scheduling problem would incorporate the dynamic effects within the resultant optimisation schedules. The optimisation study here has been in the finite-dimensional regime. A whole wealth of infinite-dimensional optimisation theory exists however, which could be developed to handle the multi-objective scheduling problem. Initial work has already been undertaken for the dynamic minimisation of fuel use, [23]. Transference into a multi-objective domain could essentially remove the modelling imbalance between the optimisation and application processes and explicitly incorporate drive-

ability constraints at the optimisation level. However, the price again may be the loss of simplicity of the design process and the resultant control laws.

The aim of this thesis has thus been achieved through the development of a multi-objective design and control environment. The engineer is provided with a tool which helps not only in his understanding of the system but provides him with control laws which are implementable with today's technology. The powertrain management approach, treated very little in the literature before, has been followed, and applied for the first time to a diesel-engined vehicle. Moreover, the use of multi-objective techniques, to minimise both vehicle fuel and emissions, has not been seen previously. The formulation here, using finite-dimensional methods, goes a long way to providing optimal control for dynamic vehicle motion. The promise of even better results with improved modelling and optimisation can only help the automotive industry's quest to stay ahead of current legislation, and in doing so play their role in the protection of the environment.

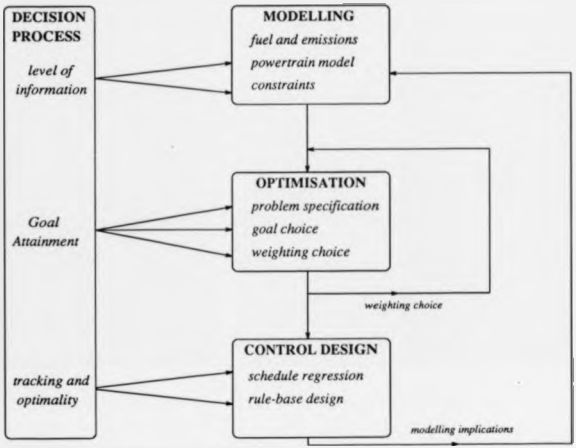


Figure 7.1: CAD Environment

Bibliography

[1] Flemming W.H, Chairman, *Future Directions in Control Theory : A Mathematical Perspective*, S.I.A.M. Report on Issues In The Mathematical Sciences, 1989.

[2] Costa A., Jones R.P., *Motion Management for Automotive Vehicles*, I.E.E. International Control Conference 1991, Edinburgh, March 1991.

[3] Tokuda T., *Cars In The '90s As A Humanware*, S.A.E. Paper 885049, (1988).

[4] Rivard J.G., *Automotive Electronics in the Year 2000*, S.A.E. Paper 861027, (1986).

[5] Braess H.H., Thomson B., *The Motor Vehicle - A Good Example of Wide Range of Application of Modern Control Engineering*, I.F.A.C. World Congress, 1987

[6] Wachter W.F., *Analysis of Transient Emission Data of a Modelyear 1991 Heavy Duty Diesel Engine*, S.A.E. Paper 900443, (1990).

[7] Cartellieri W.P. Wachter W.F., *Status Report on a Preliminary Survey of Strategies to Meet U.S-1991 H.D Diesel Emission Standards Without Exhaust Gas Aftertreatment*, S.A.E. Paper 870342, (1987).

[8] Robert Bosch Gmbh, *S.A.E. Automotive Handbook*, 2nd Ed., Delta Press, 1989.

[9] Barber P., Jones R.P, Scotson P.G.; *Pseudo-Random Binary Sequence (PRBS) Identification Techniques Applied to Multipoint Fuelling Systems in a Modern Automotive Vehicle*, Control and Instrument Systems Centre Report No. 10, University of Warwick, 1990.

[10] Dobner D.J., *Dynamic Engine Models for Control Development Part 1: Non-linear Model Formulation*, Int. J. Vehicle Design, Series SP4, *pp 54-74*, (1983).

[11] Blumberg P.N, Auiler J.E, and Wu H., *A Methodology for Evaluation and Optimisation of Three-Way Catalyst Based, Low NO_x Emission Control Systems*, S.A.E. Paper 810273, (1981).

[12] Tennant J.A., Cohen A.I., Rao H.S., and Powell J.D., *Computer-Aided Procedures for Optimisation of Engine Controls*, Int. J. of Vehicle Design, Vol. 4, No. 3, *pp 258-269*, (1983).

[13] Blumberg P.N., *Powertrain Simulation: a Tool for the Design and Evaluation of Engine Control Strategies in Vehicles*, S.A.E. Paper 760158, (1976).

[14] Powell J.D., *A Review of I.C Engine Models For Control System Design*, I.F.A.C. World Congress, 1987.

[15] Kiencke U, *A View of Automotive Control Systems*, I.E.E.E. Control Systems Magazine, August, *pp 11-18*, 1988.

[16] Thring R.H., *Engine Transmission Matching*, S.A.E. Paper 810446, (1981).

[17] Wong L.T., *Powertrain Matching For Better Fuel Economy*, S.A.E. Paper 790045, (1979).

[18] Christenson B.C., Frank A.A., Beachley H., *Fuel Saving Potential of Cars With a C.V.T and an Optimal Control Algorithm*, A.S.M.E. Proceedings, Paper No. 75-W A/A U T 20, (1975).

[19] Kuzak D.M., Shields B.D, Freedman R.J., Lee J.J, and Rittmueller D.L., *Powertrain Control Strategy Determination for Computer-Controlled Transmissions*, Int. J. Vehicle Design, Vol 8, No.1, *pp 13-36*, 1987.

[20] Flemming P.J.; *Computer Aided Design of Regulators Using Multi Objective Optimisation*, Control Applications of Nonlinear Programming and Optimisation, Pillo G.D.I, (Ed.), Proc. 5th I.F.A.C. Workshop, Italy, June 1985, Pergamon Press.

[21] Radtke R.R, Unnewehr L.E, and Freedman R.J., *Optimisation of a Continuously Variable Trasmission with Emission Constraints*, S.A.E. Paper 810107, (1981).

[22] Jones R.P, *Fuel Optimal Control Of An Automotive C.V.T Powertrain*, in Owens D.H and Nichols N. (Eds), I.M.A Conference on Control Theory, Academic Press, 1991.

[23] Garbett K.S, Jones R.P., *Optimisation of Fuel Economy Through Dynamic Scheduling of Engine Torque and Gear Ratio*, I.E.E. Colloquim on Powertrain Control, May 1990.

[24] Steuer R.E, *Multiple-Criteria Optimisation, Theory, Computation and Application*, Wiley Series in Probability and Mathematical Statistics, and Applied

Probability , Wiley 1986.

[25] Kuriger I.F., *Some Aspects of Modelling and Control of Automotive Power Systems'*, Ph.D Thesis, Engineering Department, University of Warwick, 1985.

[26] Hrovat D., Powers W.F., *Power Train Computer Control Systems'*, International Federation of Automatic Control, World Congress, 1987

[27] Dobner D.J., *A Mathematical Engine Model for Development of Dynamic Engine Control'*, S.A.E. Paper 800054, (1980).

[28] Grimm R.A., Bremer R.J., Stonestreet S.P., *GM Micro-computer Engine Control System*, Implementation of Engine Control Strategies With Electronics, S.A.E. Publication, SP-477, Paper 800053, (1980).

[29] Williams T.J., Whitehouse N.D., *Investigation Into Some Aspects of the Computation of Diesel Engine Combustion*, Proc. IMechE, V. 190 42/76, 1976.

[30] Bowns D.E., *The Dynamic Transfer Characteristics of Reciprocating Engines*, Proc. IMechE, V 185 16/71, 1971.

[31] Winterbone D.E., Thiruarooran C., Wellstead P.E., *A Wholly Dynamic Model of a Turbocharged Diesel Engine for Transfer Function Evaluation*, S.A.E. Paper 770124, (1977).

[32] Fehr W.L., Medanic J.V., *Using an Angle-Based, Sampled-Data Approach to Modeling a Diesel Engine*, International Federation of Automatic Control, World Congress, 1987.

[33] Prabhabkar R., *Optimisation of Automotive Engine Fuel Economy and Emissions*, A.S.M.E. Paper No. 75-WA-AUT-19, 1975.

[34] Rishavy E.A., Hamilton S.C., Ayers J.A., Keane M.A., *Engine Control Optimisation for Better Fuel Economy with Emission Constraints*, S.A.E. Paper No. 770075, (1977).

[35] Cassidy J.F., *A Computerized On-Line Approach to Calculating Optimum Engine Calibrations*, S.A.E.Paper No. 0078, (1977).

[36] Rao H.S., Cohen A.I., Tennant J.A., VanVoorhies K.L., *Engine Control Optimisation Via Nonlinear Programming*, S.A.E. Paper No. 790177, (1979).

[37] Auiler J.E., Zbrozek J.D., Blumberg P.N., *Optimisation of Automotive Engine Calibration for Better Fuel Economy - Methods and Applications*, S.A.E.Paper No. 770076, (1977).

[38] Dohner A.R., *Transient System Optimisation of an Experimental Engine Control System Over the Federal Emissions Driving Schedule*, S.A.E. Paper No. 780286, (1978).

[39] Dohner A.R., *Optimal Control Solution of the Automotive Emission-Constrained Minimum Fuel Problem*, Automatica, Vol. 17, No 3, *pp 441-448*, 1981.

[40] Matsomuto K., Inoue T., Nakanishi K., Matsushita S., Koganemaru S., Ooshika H., *Engine Control Optimisation for Smaller Passenger Cars*, S.A.E. Paper No. 780590, (1978).

[41] Hales J.M, May M.P., *Transient Cycle Emissions Reduction at Ricardo - 1988 and Beyond*, S.A.E. Paper No. 860456, (1986).

[42] Wade W.R., Hunter C.E., Trinker S.P., Hansen S.P., *Future Diesel Engine Combustion Systems for Low Emissions and High Fuel Economy*, F.S.I.T.A.-Congress, Belgrade, Paper No. 865012, 21, 1986.

[43] Ironside J.M., Stubbs P.W.R., *Microcomputer Control of an Automotive Perbury Trasmission*, Proceedings of 3rd International Conference on Automtive Electronics, Mechanical Engineering Publications, 1981.

[44] Tsangarides M.C., Tobler W.E., *Dynamic Behaviour of a Torque Converter with Centrifugal Bypass Clutch*, S.A.E. Paper No. 850461, (1985).

[45] Kotwicki A.I., *Dynamic Models for Torque Converter Equipped Vehicles*, S.A.E. Paper No. 820393, (1982).

[46] Moskwa J.J., Hedrik J.K. *Sliding Mode Control of Automotive Engines*, Proceedings of ACC (American Control Conference), Pittsburg, U.S.A, 1989.

[47] Cho D., Hedrick J.K., *Sliding Mode Control of Automotive Powertrains With Uncertain Actuator Models*, Proceedings of ACC (American Control Conference), Pittsburg, U.S.A, 1989.

[48] Holman C.D., *Pollution from Heavy Goods Vehicles*, The Impact of Emissions Legislation on the Fleet Operator, IMechE Conference, London, 1990.

[49] Crossley P.R., Jones R.P., Howarth S.I., *Modelling and Simulation of a HGV Powertrain for Tansmission Control Studies*, Int. Symposium on Automotive Technology and Automation, Graz, 1985, I.S.A.T.A. Paper No. 85047.

[50] Dittrich O., *The Continuously Variable Chain Transmission in Motor-Cars*, S.A.E. Paper No. 885064, (1988).

[51] Mason P., Advanced Reasearch Centre, Lucas Automotive Ltd., *Private Communication*, Pro-Matlab m-files.

[52] Mitchell E.E., Gauthier J.S., *ACSL : Advanced Continuous Simulation Language - User Guide/Reference Manual*, Mitchell and Gauthier Assoc., 1981.

[53] Gear C.W., *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, 1971.

[54] Cunningham W.J., *Introduction to Nonlinear Analysis*, McGraw-Hill, 1958.

[55] Lamda S.S., Kavanagh R.J., *The Phenomenon of Isolated Jump Resonance and Its Applications*. Proc. I.E.E., V118, *pp* 1047-1050, 1971.

[56] Utkin V.I., *Sliding Modes and Their Application in Variable Structure Systems*, MIR, Moscow, 1978.

[57] Minorsky N., *Theory of Nonlinear Control Systems*, McGraw-Hill, New York, 1969.

[58] Popov E.P., Pal'tov I.P., *Methods of Approximation in the Theory of Nonlinear Control Systems*, Fizmatgiz, Moscow, 1960.

[59] Graham D., McRuer D., *Analysis of Nonlinear Control Systems*, J. Wiley & Sons, 1961.

[60] Thaler G.J., Pastel M.P., *Analysis and Design of Nonlinear Feedback Control Systems*, McGraw-Hill, 1962.

[61] Atherton D.P., *Nonlinear Control Engineering : Desribing Function Analysis and Design*, Van Nostrum and Reinhold, London, 1975.

[62] Mees A., *Desribing Function, Circle Criteria and Multi-Loop Feedback Systems*, Proc. I.E.E. Volume 120, *pp 126-130*, 1973.

[63] Gelb A., Van der Velde W.E., *Multiple-input Desribing Functions and Nonlinear System Design*, McGraw-Hill, New York, 1968.

[64] Mayne D.Q., Sahba M., *Design of Feedback Controllers for Nonlinear Systems*, Int. Conf. on Control and its Applications, Computing ans Control Division, I.E.E., University of Warwick, March 1981.

[65] Janáč K., Vojtasek S., *Solution of Nonlinear Systems*, Luffe, 1969.

[66] Adby P.R., Dempster M.A.H., *Introduction to Optimisation Methods*, Chapman and Hall, London, 1974.

[67] Nelder J.A., Mead R., *A Simplex Method for Function Minimisation*, Computer Journal, V. 7, *pp 308-313*,, 1965.

[68] Rosenbrock H.H., *Automatic Method for Finding the Greatest or Least Value of a Function*, Computer Journal, V. 6, *pp 175-184*, 1960.

[69] Fletcher R., Powell M.J.D., *A Rapidly Converging Descent Method for Minimisation*, Computer Journal, V. 6, *pp163-168*, 1963.

[70] Gill P.E., Murray W., Wright M.H., *Practical Optimisation*, Academic Press, 1981.

[71] Luenberger D.G., *Optimisation by Vector Space Methods*, Wiley, New York, 1969.

[72] Polak E., *Computational Methods in Optimisation: A Unified Approach*, Academic Press, 1971.

[73] Noton M., *Modern Control Engineering*, Pergamon Press Inc., 1972.

[74] Quintana V.H., Davison E.J., *Clipping-Off Gradient Algorithms to Compute Optimal Controls With Constrained Magnitude*, Int. J. Control, V20,*pp* 243-255, 1974.

[75] Leiber H., Czinczel A., *Antiskid System for Passenger Cars with a Digital Electronic Control Unit*, SAE Paper 790458, (1979).

[76] Thomson B., *The Increase of Cornering Force And Traction By a Wheel-Slip-Control-System*, 9th Symposium I.A.V.S.D., Switzerland, 1985.

[77] Hayes J.G. (Ed), *Numerical Approximations to Functions and Data*, Athlone Press, 1967, *pp* 87-89.

[78] *AutoExpress*, July 9-16th, 1991.

[79] *Private Communication*, with P. Scotson, P. Mason., Advanced Research Centre, Lucas Automotive Ltd.

[80] *Federal Register*, V.42, No. 124.

[81] Grace A.C.W., *Computer-Aided Control Sysytem Design Using Optimisation Methods*, Ph.D Thesis, University College of North Wales, Bangor, 1989.

[82] Flemming P.J., *Application of Multi-Objective Optimisation to Compensator Design for SISO Control Systems*, Electronic Letters, Vol. 22, No. 5, *pp* 258-259, 1986.

[83] Gall D.A., *A Practical Multifactor Optimisation Criterion*, Recent Advances in Optimisation Techniques, Lavi .A., Vogl T.P., (Ed.), Proc. of Symposium on Recent Advances in Optimisation Techniques, J. Wiley & Sons, 1965.

[84] Matsuda T., Fujii K., *Mulit-Objective Optimisation Method and it's Application to Dam Control Problems*, Trans. Inst. Electon. Commun. Eng., Section E, Vol. E62, No. 11, Nov. 1979, *pp* 741-748.

[85] Osyczka A, *Multi-Criteria Optimisation Engineering With Fortran Programs*, Ellis Horwood, 1984.

[86] Stadler W., *Multi-Criterion Optimisation in Mechanics : A Survey*, Applied Mechanics Review,Vol. 37, No. 3, March 1984, *pp* 277-286.

[87] Matsuda T., Fujii K., *An Approach to Dynamic Optimal Control Problems With Multi-Criteria*, Technol. Rep. Osaka Uni. Vol. 29, No. 1492-1516, Oct. 1979, *pp* 383-393.

[88] Censor Y., *Pareto-Optimality in Multi-Objective Problems*, Applied Mathematics and Optimisation, Vol. 4, *pp* 41-59, 1977.

[89] Owen G., *Gaming Theory*, 2nd Ed.,Academic Press, 1982.

[90] Zadeh L.A., *Optimality and Non-Scalar-Valued Performance Criteria*, IEEE Trans., Aut. Ctrl., AC-8, 1963.

[91] Geoffrin A.M., *Solving Bi-Criterion Mathematical Programs*, Operational Research, Vol. 15, *pp* 39-54, Jan./Feb. 1967.

[92] Waltz F.M., *An Engineering Approach : Hierachical Optimisation Criteria*, IEEE Trans., Vol AC-12, April, 1967, *pp* 179-180.

[93] Nelson W.L., *On the Use of Optimisation Theory for Practical Control System Design*, IEEE Trans., Vol. AC-9, Oct, 1964, *pp* 469-478.

[94] Gembicki F.W, *Vector Optimisation for Control with Performance and Parameter Sensitivity Indices*. Ph.D Thesis, Case Western Reserve Univ., Cleveland, Ohio, USA, 1974.

[95] *Real Control Problems: No Solution Yet?*, I.E.E. Colloquium, London, May 1991.

Appendix A

Matlab Files

The optimisation and simulation of the fuel and emission functions is implemented through user-written MATLAB *m-files*. This appendix displays some of the important files for the single-objective optimisation study of Chapter 4. The files are commented throughout, but some additional explanation of their use will be given. The overall program has been written in a modular top-down fashion to aid in debugging and flexibility. It has also been written such that files needed for all optimisations need not be duplicated. Only those specific to the particular function, *e.g* cost function, gradient, and Lagrange Multiplier calculation files differ and are switched in where necessary.

The controlling file is **control.m**, page 278. This calls the optimisation header file for each data point on the test cycle. This file is independent of the function definition. Specific function files are called where necessary for implementation of the separate optimisations. That is, files which are specific to the function being minimised, *e.g.* the cost function evaluation file, not shown here; the **adjoint.m** file which calculates the Lagrange Multipliers from a differentiation of the cost function;

and the **gradsa.m** file which calculates the search direction, again from a differentiation of the cost function. The latter two are shown on pages 282, 283. The specific calculations which they solve are explained in Chapter 3, Section 3.4.6.

The controlling file calls the optimisation header file **steda.m**, pages 279,280, again independent of the function. This file reads in initial variables from **initiala.m** and checks their feasibility. If the state independent variables are violated a clipping routine is called (not shown) to clip them to their boundary values. The initial multiplier values and gradient direction is then calculated through calls to **adjointa.m** and **gradsa.m**. The files shown on page 282, 283 are specific to the fuel minimisation problem. Files for minimisation of emissions differ only in the analytic equations used within. Note that before the iterative gradient search routine is entered the direction of search is set to that of steepest descent, see Chapter 3, Section 3.4.6. The iterations proceed while the convergence criteria is satisfied and the maximum number of iterations has not been reached. The Fletcher-Reeves Algorithm is embedded in this iterative procedure, a successful pass of which will lead to the iteration counter being moved on and the new variables set to be the current ones. The remaining section of **steda.m** records the reason for exiting the convergence test and calculates the function values. State and control variables are rescaled for output and passed back to **control.m** for inclusion in the output matrix.

The initialisation file **initiala.m** and the state equation file 'calcxa.m' are shown on pages 281, 284. The first defines the system parameters, calculates the initial conditions, scales the problem, and sets up the global parameter *gv_param*. **Calcxa.m** performs state equation calculations, once the optimisation process has started.

Other files, not shown, check the feasibility of the state variables, **feasiblel.m**, and calculate the state dependent bounds, **u1maxa.m**, **u1mina.m**.

control.m Page 1

```
%*****************************************************************
% CONTROL.M                                                       *
% FILE TO AUTOMATE STEDA PROGRAMS OVER WHOLE TEST CYCLE           *
% inputs -- inputfile.m contains test cycle description and initial*
%           engine torque, brake force values.                     *
% output -- data vector output built up over cycle and saved in    *
%           The****.mat.                                           *
%        -- total fuel and emission figures are output to screen   *
%*****************************************************************

diary fuel

%LOAD IN CYCLE AND INITIAL VALUES
[y0,ydot0,u0]=inputfilef;
u01 = u0(:,1);
u02 = u0(:,2);

if length(y0) ~= 132
   disp('error in size of input vectors')
end

%INITIALISE AND SET UP GLOBALS
output=[];
count = 0;
global gv_param

%ITERATE THE OPTIMISATION PROCEDURE
for f=1:length(y0)
     count = count + 1
     [datalist] = steda(y0(f),ydot0(f),u01(f),u02(f));
     output = [output;datalist];
     if datalist == []
        disp('datlist  = empty')
        pause
     end
end;

%ASSIGN VARIABLES
y0=output(:,1);
F=output(:,2);
pmu=output(:,3);
hmu=output(:,4);
smu=output(:,6);
nmu=output(:,5);
a1=output(:,7);
a2=output(:,8);
u1=output(:,9);
u2=output(:,10);
iter=output(:,11);
Jn=output(:,12);
a01=output(:,13);
flag =output(:,14);
ydot = output(:,15);
ui = output(:,16);

%CALCULATE FUEL AND EMISSION VALUES FOR CYCLE
tt=[0:131]';
totf=simpson(tt,F)
disp('fuel usage in g/s =') ,totf
totn = simpson(tt,output(:,5))
disp('nox output in g/s='),totn
totp = simpson(tt,output(:,3))
disp('part output in g/s = '),totp
toth = simpson(tt,output(:,4))
disp('hc output in g/s'),toth
tots = simpson(tt,output(:,6))
```

control.m Page 2

```
disp('smok output in g/s'),tots

save Thefuel output
diary off
```

```
%*************************************************************************
% STEDA.M  -automated optimization header file                           *
% Steady state optimization of the desired test cycle.                   *
% Specific function minimization is controlled by the input files        *
% cost.m, adjoint.m and grad.m.                                          *
% A Fletcher Reeves algorithm is used with constraints incorporated      *
% as Lagrange multipliers. State and control                             *
% bounds are handled by a simple approach which attempts to prevent     *
% violation of the constraints at each iteration.                        *
% The initial solution is assumed to satisfy the constraints.            *
%                                                                        *
%*************************************************************************
%       xc(1) = engine speed ,rads/sec                                   *
%       xc(2) = vel. ratio                                               *
%       uc    = [engine torque ,Nm/s , brake force, N]                   *
%       ydd   = desired vehicle speed, m/s                               *
%       ydot  = desired acceleration rate, m/s/s                         *
%       K     = iteration step size                                      *
%       maxit = maximum iteration count                                  *
%       tol   = convergence criterion                                    *
%       Jc    = cost function                                            *
%       gc    = gradient value                                           *
%                                                                        *
% ASSOCIATED M-FILES :                                                   *
%       initiala.m  = set up nominal control, scale problem, etc         *
%       feasible.m  = check bounds on variables satisfied                *
%       costa.m     = calculate cost function                            *
%       adjointa.m  = calculate lagrange multipliers                     *
%       gradsa.m    = calculate gradient                                 *
%                                                                        *
% global variable gc param - model//cost parameters                      *
%*************************************************************************

function [details]= steda(y0,ydot0,u10,u20);
[y0,uc,w0,K,tol,maxit] = initiala(y0,ydot0,u10,u20);
ac = w0
ui = uc;

%CHECK FOR INFEASIBILITY AND TAKE ACTION
uc = clip(uc,ydot0);
subiter = 1;
if ~feasible(ac,uc)
    disp('nominal control violates state indep constraints')
    pause
end

%CALCULATE COST AND INITIALISE GRADIENT AND MULTIPLIERS
Jc =costa(ac,uc)

iter =0
progress = [iter Jc subiter ];

subiter = 0;
pc = adjointa(ac,uc,ydd);
gc = gradsa(ac,uc,pc);

%GRADIENT SEARCH
ac = - gc;
while K*norm(gc) > tol*norm(uc)          %convergence test
      uc = uc + K*ac                     %update u
      uc = clip(uc,ydot0);               %clip at bound
      ac = calcsa(ac,ydd);               %new state variables
      subiter = subiter +1
      Jc = costa(ac,uc)
      if Jc > Jc
            disp('Jc > Jc')              %if cost increases 1/2 step
```



```
            K = 0.5*K;
      elseif ~feasible(ac,uc)            %if bounds not satisfied
            K = 0.5*K;                   %1/2 step
      else
            iter = iter + 1
            if iter > maxit
              disp('maximum iterations exceeded')
              break
            end
            progress = [progress;iter Jc subiter];
            subiter = 0;                 %reset
            pc = adjointa(ac,uc,ydd);    %multipliers for new control
            gp = gradsa(ac,uc,pc);       %gradient for new control
            if rem(iter,3) ~= 0          %reset conjugate gradients
            be = norm(gc)^2/norm(gc)^2;
            else be=0;
            end
            ac = -gc + be*ac;            %new gradient direction
            ac = ac;                     %reset to current variables
            gc = gc;                     %for next iteration
            Jc = Jc;
            uc = uc;
            ac = ac;
            K = K*2;                     %iteration successful
      end                                %double step
end

%LOG REASONS FOR STOP
%MAKE SURE IT WASNT IN THE MIDDLE OF AN INFEASIBLE SUBITER
flag = 0;
disp('exiting convergence test')
if iter == 0
   disp('zero iterations: optimizer failed to move')
   ac = ac;
   uc = uc;
   Jc = Jc;
   flag = 3;
   progress
elseif ~feasible(ac,uc)
   ac=ac;
   uc=uc;
   Jc=Jc;
   flag = 1;
elseif Jc > Jc
   Jc=Jc;
   ac=ac;
   uc=uc;
   flag = 2;
end
if iter <  maxit
          disp('convergence achieved')
end

clear ac uc pc Jc
ale

%CALCULATE FUNCTION VALUES
load xycoeff
xx = [ac(1)*100^2 - [af(1)+af(2)]/[af(1)-af(2)];
ux = [ac(2)*100^2 - [af(3)+af(4)]/[af(3)-af(4)];
fac = genxyfun(xx,ux,coeff);
gac = genxyfun(xx,ux,coeffp);
hac = genxyfun(xx,ux,coeffh);
nuc = genxyfun(xx,ux,coeffn);
euc = genxyfun(xx,ux,coeffe);   % continued ...
```

steda.m continued


```
CHECK COST AND FUNCTION VALUE TALLY
   if abs(fnu - fn) < 0.01
      disp('WARNING : were up the creek!')
      pause
      flag = 4;
   end

   %OVERCOME NEGATIVE SMOKE PROBLEMS
   if xnu < 0
      xnu = 0;
      flag = 4;
   end

   %RESCALE VARIABLES FOR OUTPUT
   X1 =100*xn(1);
   X2 = xn(2);
   U1 =100*un(1);
   U2 = 1000*un(2);
   yd = yd*10;
   xf(1) = xf(1)*100;
   pdot = gv_param(24)*1000;
   xi(1)=xi(1)*100;
   xi(2) = xi(2)*1000;

   datalist=[yd fnu pnu hnu cnu snu X1 X2 U1 U2 iter Jn xf(1) flag...
             pdot xi ];
```

initiala.m Page 1

```
%*************************************************************
%INITIALA -  INITIAL VALUES AND SCALING                      *
%  inputs -- lamda in cycle and initial information from inputfile*
%  output -- to steda.m                                      *
%*************************************************************

function [y,u,x0,R,tol,maxit] = initiala(y,ydot,u10,x20)
format short e

%DEFINE SYSTEM PARAMETERS
M = 1800                % vehicle mass (kg)
A = 50;                 % road load coeffs
B = 0.34;
C = 0.0255;
RD = 0.275;             % final drive
RW = 0.28;              % wheel radius (m)

%EVALUATE MODEL PARAMETERS
alpha = 1;
beta = M*(RD*RW)^2;
gamma = A*RD*RW;
delta = B*(RD*RW)^2;
eps = C*(RD*RW)^3;
mu = RD*RW;

%BOUNDS ON STATES
x1max = 640.77;
x1min = 89.2;
x2max = 2.0;
x2min = 0;

%USE DAMAGED OUTPUT
%CALCULATE INITAIL STATES
x0(2) = u10/mu*((A + B*y + C*y^2) + M*ydot + u20);
x0(1) = y/mu*x0(2);

%DEFINE NOMINAL STEP SIZE
R = 0.01;

%DEFINE NOMINAL TOLERANCE
tol = 1.0e-8;

%DEFINE MAX ITERATIONS
maxit = 100;

%DEFINE CONTROL VARIABLE CONSTRAINTS
load pkvar
clear fmax1 fmax2 fmax3 fmin xc c el data1 data2 data3 data4
u1max = 6.5e+03;
u1min = 0.0;

%SCALING
su1 = 1.0e2;
su2 = 1.0e2;
su3 = 1.0e3; %

u(1) = u10/su1;
u(2) = u20/su2;
x0(1) = x0(1)/su1;

gamma = gamma/su1;
delta = delta*su1/su2;
eps = eps*su1^2/su2;
mu = mu*su3/su1;

x(1) = x(1)*su1^2/su1;
```

initiala.m Page 2

```
x(2) = x(2)*su1/su1;
x(3) = x(3)/su1;
b(1) = b(1)*su1/su1;
b(2) = b(2)/su1;
d(1) = d(1)*su1/su1;
d(2) = d(2)/su1;
e(1) = e(1)*su1/su1;
e(2) = e(2)/su1;

x1min = x1min/su1;
x1max = x1max/su1;
vmax = vmax/su1;
vmin = vmin/su1;
u1max = 1max/su3;
u1min = u1min/su3;

y = y*su3/(su1*su1);
ydot = ydot/su3;

A = A/su3
B = B*(su3/(su1*su1))^2;
C = C*su3/(su1*su1);

%WRITE MODEL PARAMETERS TO GLOBAL VARIABLE
gv_param = [ gamma, delta, eps, mu, u ...
             x1max, x1min, x2max, x2min, x, b ,A ,B,C ...
             vmax,vmin,M,ydot,ps,ks,e,u,u1max,u1min,d,e ];
%
```

adjointaf.m Page 1

```
%*************************************************************************
%ADJOINTAF.M -  DETERMINATION OF LAGRANGE MULTIPLIERS FOR FUEL           *
%calculate lagrange multipliers through differentiation of Hamiltonian  *
%*************************************************************************

function p =adjointaf(x,u,y)

mu = gv_param(8);
gamma = gv_param(1);
delta = gv_param(2);
eps = gv_param(3);
M = gv_param(28);
ydot=gv_param(29);

%CALCULATION OF FUEL FUNCTION
load mycoeff
ax = (x(1)*100^2 - (xf(1)+xf(2)))/(xf(1)-xf(2));
ux = (u(1)*100^2 - (xf(3)+xf(4)))/(xf(3)-xf(4));

fuu1=coeff(1) + coeff(2)*ax + coeff(3)*ux +coeff(4)*(ax^2)+ ...
     coeff(5)*ax*ux + coeff(6)*(ux^2) + coeff(7)*(ax^3) + ...
     coeff(8)*(ax^2)*ux+coeff(9)*ax*(ux^2)+coeff(10)*(ux^3);
fuu2=coeff(11)*ax^4 + coeff(12)*(ax^3)*ux + ...
     coeff(13)*(ax^2)*(ux^2) + coeff(14)*ax*(ux^3) + ...
     coeff(15)*(ux^4);
fuu = fuu1 + fuu2;

%DF/DX
fdx1 = coeff(2) + 2*coeff(4)*ax +coeff(5)*ux +3*coeff(7)*ax^2 ...
     + 2*coeff(8)*ax*ux + coeff(9)*(ux^2) + 4*coeff(11)*ax^3 ...
     + 3*coeff(12)*(ax^2)*ux + 2*coeff(13)*ax*(ux^2) + coeff(14)*(ux^3);

fdx2 = 0;

%CONSTRAINTS AND THEIR DERIVATIVES
con1 = x(1)*x(2)*mu - y;
con2 = -x(1) - x(2)*(gamma + delta*x(1)*x(2) + eps*(x(1)*x(2))^2) ...
       - mu*x(2)*M*ydot - mu*x(2)*x(2);
%DC/DX
con1dx1 = x(2)*mu;
con1dx2 = x(1)*mu;
con2dx1 = -1 delta*x(2)^2 + 2*eps*x(1)*x(2)^3 ;
con2dx2 = -(gamma + delta*2*x(1)*x(2) + 3*eps*x(2)^2*x(1)^2) ...
          -mu*M*ydot - mu*x(2);
cond = [con1dx1 con1dx2;con2dx1 con2dx2] ;
l = inv(cond);
lambda= -[fdx1 fdx2]*l;
p=lambda';
```

gradsaf.m Page 1

```
%*********************************************************************
*
%GRADSAF - DETERMINATION OF DE/DU  (E = EARCLIDINIAN)
 *
%calculation of the search direction for the optimization algorithm
*
%*********************************************************************
*
function g = gradsaf(e,u,p)

mu = gv_param(4);
load xycoeff

%SCALED VARIABLES FOR FUEL FUNCTION EVALUATION
xx = (e(1)*100^2 - (af(1)+af(2))/(af(1)-af(2));
ux = (u(1)*100^2 - (af(3)+af(4))/(af(3)-af(4));

hdu(1) = coeff(3) + coeff(5)*xx + 2*coeff(6)*ux + coeff(8)*(xx^2) + ...
          2*coeff(9)*xx*ux + 3*coeff(10)*ux^2 + coeff(12)*xx^3 + ...
          3*coeff(13)*(xx^2)*ux + 3*coeff(14)*xx*(ux^2) + 4*coeff(15)*ux^3 + p(2)
;
hdu(2) = -e(2)*mu^p(2);
g = hdu;
```

calcxa.m Page 1

```
%*************************************************************
%CALCXA - CALCULATION OF STATES                              *
%calculate the state variables from the constraint equations*
%*************************************************************

function x = calcx(x,y)

mu = gv_param(4);
gamma = gv_param(1);
delta = gv_param(2);
eps = gv_param(3);
M = gv_param(28);
ydot = gv_param(29);
A = gv_param(23);
B = gv_param(24);
C = gv_param(25);

x(2) = x(1)/(A + B*y + C*y^2)*mu + M*ydot*mu + x(2)*mu;
x(1) = y/(x(2)*mu);
```

Multi-Objective Scheduling and Control of a Nonlinear Automotive Powertrain

AUTHOR ..... K. S. Garbett .....

DEGREE .....

AWARDING BODY University of Warwick

DATE ..... September 1991

THESIS
NUMBER .....