Vehicle System Dynamics
Vol. 49, No. 7, July 2011, 1073–1111

Vehicle dynamics applications of optimal control theory

R.S. Sharp[a]* and Huei Peng[b]

[a]*Department of Mechanical, Medical and Aerospace Engineering, Faculty of Engineering and Physical sciences, University of Surrey, Guildford GU2 7XH, UK;* [b]*Department of Mechanical Engineering, University of Michigan, G036 Lay Automotive Laboratory, Ann Arbor, MI 48109-2121, USA*

(*Received 18 March 2011; final version received 4 May 2011*)

The aim of the paper is to survey the various forms of optimal-control theory which have been applied to automotive problems and to present illustrative examples of applications studies, with assessments of the state of the art and of the contributions made through the use of optimal-control ideas. After a short introduction to the topic mentioning several questions to which optimal-control theory has been addressed, brief reviews of automotive-applicable optimal-control theory are given. There are outlines of the Linear Quadratic Regulator, without and with state reconstruction and then with the addition of disturbance preview, the nonlinear regulator or state-dependent-Riccati equation method, general numerical optimal-control theory including indirect and direct methods, model predictive control and robust control. Applications of the theory to active and semi-active suspension design and performance, worst-case manoeuvring, minimum-time manoeuvring and high-quality driving are then discussed in detail. Application sections describe the problem, the theory that has been used, what has been discovered and what remains to be found. The record of optimal-control theory in aiding the understanding of the various issues, in helping with system designs and knowledge of what is possible, and in guiding future research is assessed. Some ideas about future work are included.

Keywords: optimal control; suspension control; rollover; optimisation; driver behaviour

1. Introduction

Whenever a system contains actuation and control, its operating circumstances and its objective are precisely defined, optimal-control theory can, in principle, be used to find the control that best achieves the objective. Of course, the problem must be properly structured to ensure that a solution is possible. Most often, the optimal control itself will be interesting mainly insofar as it enables the discovery of the best possible system performance. Occasionally, the optimal control will provide a basis for the design and operation of practical systems.

Optimal-control theory in various forms has been applied to (a) active and semi-active suspension design and performance evaluation; (b) worst-case manoeuvring; (c) minimum-time manoeuvring; (d) driver and rider control; (e) state estimation; (f) obstacle and collision avoidance; (g) integrated motion control; (h) spark-ignition engine air/fuel ratio control and (i) transmission clutch torque control. This list is probably not exhaustive.

*Corresponding author. Email: robinsharp@waitrose.com

ISSN 0042-3114 print/ISSN 1744-5159 online

DOI: 10.1080/00423114.2011.586707
http://www.informaworld.com

Forms of optimal-control theory that have been used in automotive problems include (i) the Linear Quadratic Regulator (LQR) or the Linear Quadratic Gaussian (LQG) theory; (ii) continuous or discrete LQR/LQG with road preview (LQRP); (iii) Nonlinear Quadratic Regulator or State-dependent-Riccati equation (SDRE) theory; (iv) indirect optimisation by the application of Pontryagin's minimum principle; (v) direct optimisation by nonlinear programming and the application of the Karush–Kuhn–Tucker (KKT) conditions; (vi) model predictive control (MPC) and (vii) the robust control theory. Again, the list is not exhaustive.

The breadth of applications and the number and complexity of the solution techniques used imply that it is necessary to be selective and to limit the scope of the paper. It is intended to discuss the problem areas (a)–(d), relating the problems to the techniques that have been most popular in connection with them. The focus will be on the advancements in the understanding of the problems that have come about as a result of the application of optimal-control theory, with as much guidance towards efficient future activities as possible. The account will not be complete in a historical sense.

Section 2 contains brief review material on the various forms of optimal-control theory referred to above, with the accent on the user and reliance on the extensive literature for mathematical details. Section 3 considers application to each of four problem areas, with background information to put the optimisation story into context. Section 4 is devoted to concluding statements, with emphasis on guidance for future research.

2. Optimal-control theory reviews

2.1. *The LQR*

The LQR theory is standard and is covered in many textbooks [1–5]. Good accounts can also be found on the Internet. The control to be found is for a linear plant with a quadratic cost function to be minimised by the control. The cost function contains terms that represent outputs from the plant and control power, with weights to specify the relative importance of the factors. The cost function may also contain cross-terms involving both outputs and controls but that is not important in the present context. The state of the plant must be available for feedback, the optimal control being in full-state-feedback form, its existence depending on the system structure yielding stabilisability and detectability. Cost function terms may be frequency-weighted [4,6,7]. Automotive applications usually presume time-invariance and white-noise excitation with the optimisation extending over infinite time and concerning expected values. LQR solutions automatically produce controllers that are stable and somewhat robust, and they can be utilised straightforwardly. The LQR theory exists in both discrete-time and continuous-time forms and preview of the disturbance inputs can be included in the theory in either case.

2.1.1. *Restricted measurements*

If the full system state cannot be measured, it can be optimally reconstructed from the measurements made, based on the assumption that the measurements include uncorrelated Gaussian random noise of known covariance, by Kalman filtering. The Kalman filter is also a standard textbook problem [8,9] with good tutorial support available on the Internet. In discrete time, updating the state estimate at a particular time step involves a predictor step, in which the current measurements are not used, followed by a corrector step in which advantage of the current measurements is taken. Alternatively, an output-feedback control can be assumed and it can be iteratively improved towards the minimisation of a quadratic cost by gradient descent

[10–15]. In some problems, control based on a restricted set of measurements is almost as good as that based on full-state information [10,16,17].

2.1.2. *With preview of the disturbance*

If the disturbance input to the plant can be previewed, the controlled system may be much improved [18–20]. Optimal controls for the regulator with preview in continuous time were established in [15,21,22], in a form that requires on-line evaluation of the preview control. In discrete time, preview information can be incorporated into the LQR theory by augmentation of the plant with a shift register or delay line model of the disturbance, such that the cost function involves differences between future plant states and the disturbance. The regulation problem is transformed into a tracking problem. Finding the optimal control in the infinite-horizon case requires the solution of an algebraic Riccati equation which is probably inconveniently large, but the solution can be obtained from that of the non-preview problem, followed by recursive computations of successive preview gains. These gains diminish as the preview time corresponding to them increases [20], so that the calculations can be stopped when the gains are insignificantly small [23,24]. The optimal preview control can be determined offline as a sequence of gains for each control, making implementation especially convenient.

2.2. *The nonlinear regulator*

Nonlinear regulators can be constructed by creating a family of optimal linear regulators employing systematically changing weights and then fitting a nonlinear state-feedback control law, containing first- and third-order terms for example, to the linear regulator results [25]. A more direct route to a nonlinear control is through the SDRE control, surveyed recently by Çimen [26]. The method derives from LQR theory and involves extended linearisation. The SDRE control can be applied to systems that are full-state observable, nonlinear in the states and affine in the input, with infinite-horizon cost function to be minimised by the control. The cost function must be quadratic in the input but may be non-quadratic in the states and it may have state-dependent weights describing the priorities in the control design. The weighting matrices must be positive-definite or positive-semi-definite for all states. A stabilising nonlinear state-feedback control, which will approximately minimise the cost, is sought.

If such a control can be found, it will need to be computed at every time step of an application in order to follow the variations in state that cannot be known *a priori*. Each computation involves the solution of an appropriate algebraic Riccati equation, which limits the order of the system to which the method can be applied in practice, since the Riccati-equation solution becomes slow for higher-order systems. The first step in the solution is the factorisation of the nonlinear system equations by expressing the nonlinear function of states in the describing equations as a product of a matrix function of states and the states themselves. With two or more state variables, this is a non-unique operation, giving the control design degrees of freedom which are in addition to those available through the choice of weighting matrices. Different factorisations will yield different controls. Çimen describes how the factorisation can be parameterised by a vector, which may need to be a function of the states to minimise the cost, but finding the state-dependent vector which is best is typically more difficult than is warranted by the performance gain and the parameterisation vector is usually a constant. Determining a satisfactory factorisation is likely to demand trial-and-error calculations, although a rule of thumb provided is that the element (i, j) of the characteristic matrix $A(x)$ should be made non-zero whenever the ith state derivative depends on the jth state. Ideally, an appropriate form for the cost function weighting matrices will be definable from the nature of the problem of concern. Otherwise, the method demands the setting of a large number of parameters, contrary to the notion that one can easily obtain an optimal control.

At each state point, the state-dependent matrices are treated as constant and standard LQR computations are used to yield the control, which, if all goes well, will be stabilising. The closer the coefficient and weighting matrices are to being constant, the closer does the problem approach the standard LQR problem, for which minimisation of the cost and stabilisation of the system are guaranteed. Detailed stabilisability, detectability and optimality conditions are discussed by Çimen.

2.3. *General nonlinear optimal control*

Optimisation in general is supported by a large number of textbooks, for examples [27–29]. A general optimal control problem involves finding a control which is continuous in time over an interval, which drives a nonlinear system through a trajectory that minimises a cost function specified for the interval and the final state. The control will be constrained to have upper and lower limits and other constraints may apply, but constraint violations can be represented as costs, without changing their significance much, so that both constrained and unconstrained systems are of interest. A mathematical model of the plant to be controlled, a statement of any system constraints and the cost to be minimised provide a starting point. One of two approaches to the solution of an optimal-control problem is normally taken, indirect and direct approaches.

2.3.1. *Indirect optimisation*

Detailed treatments of indirect optimisation can be found in [3,5,30]. The necessary theory comes from the calculus of variations. Consider a system with state vector $\boldsymbol{x}$, control vector $\boldsymbol{u}$ and independent variable $t(0 \leq t \leq t_f)$ to have equations of motion $\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}, \mathbf{u}, t)$ or in the alternative implicit form $\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{u}, t) = \mathbf{0}$, to have specified initial conditions $\mathbf{x_0}$, and to have its performance given by a cost functional $J = h(\mathbf{x_f}) + \int_0^{t_f} \mathbf{g}(\mathbf{x}, \mathbf{u}, t)\,\mathrm{d}t$. This general cost includes a term relating to the final state. Let us adjoin the state equations to the cost functional through Lagrange multipliers $\boldsymbol{\lambda}$ to give $\bar{J} = h(\mathbf{x_f}) + \int_0^{t_f} \{\mathbf{g}(\mathbf{x}, \mathbf{u}, t) + \boldsymbol{\lambda}^{\mathrm{T}}[\mathbf{a}(\mathbf{x}, \mathbf{u}, t) - \dot{\mathbf{x}}]\}\,\mathrm{d}t$. Integrating the last term $-\boldsymbol{\lambda}^{\mathrm{T}}\dot{\mathbf{x}}$ by parts to eliminate $\dot{\mathbf{x}}$ and introducing the Hamiltonian function $H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, t) = \mathbf{g}(\mathbf{x}, \mathbf{u}, t) + \boldsymbol{\lambda}^{\mathrm{T}}.\mathbf{a}(\mathbf{x}, \mathbf{u}, t)$ allow the establishment of the conditions for the cost functional to have a stationary value, the Euler–Lagrange equations, $\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}, \mathbf{u}, t)$ with given $\mathbf{x_0}$, $\dot{\boldsymbol{\lambda}} = -[\partial\mathbf{a}(\mathbf{x}, \mathbf{u}, t)/\partial\mathbf{x}]^{\mathrm{T}}.\boldsymbol{\lambda} - [\partial\mathbf{g}(\mathbf{x}, \mathbf{u}, t)/\partial\mathbf{x}]^{\mathrm{T}}$ with final values of the Lagrange multipliers given by $\boldsymbol{\lambda}_{\mathbf{f}} = [(\partial h/\partial\mathbf{x})(\mathbf{x_f})]^{\mathrm{T}}$, and with $[\partial\mathbf{a}(\mathbf{x}, \mathbf{u}, t)/\partial\mathbf{u}]^{\mathrm{T}}.\lambda + [\partial\mathbf{g}(\mathbf{x}, \mathbf{u}, t)/\partial\mathbf{u}]^{\mathrm{T}} = 0$ determining the control values. These equations express necessary, but not sufficient, conditions for a solution. For an nth-order system, $2n$ differential equations need to be satisfied. Half of these are the state equations with specified initial conditions, while the other half are the co-state equations, needing to reach particular final conditions. This is a two-point, boundary-value problem. In principle, it can be solved by integrating the state equations forwards in time with an assumed control history from the known initial state to yield a final state and integrating the co-state equations backwards in time from the final state obtained, yielding the history of the sensitivity of the cost functional to variations in the controls. Iterative improvement of the control history by some form of gradient descent, usually but not necessarily steepest descent, to satisfaction of the necessary conditions follows. The derivatives of the Hamiltonian with respect to the controls are zero when the cost functional has a stationary value. The sufficiency of the control can normally be judged from the problem context.

Control constraints imply that only a restricted set of controls will be allowable. Typically, steering, throttle and braking controls will have maximum and minimum values and their rates of change may be limited. Due to the constraints, the unconstrained solution may not be admissible. Track-boundary constraints, expressing the idea that a vehicle reference point must remain within a specified distance of a track centre line, for example, will constitute inequalities on selected states. It remains the case that the admissible controls that minimise the Hamiltonian are optimal, which is Pontryagin's Minimum Principle. A relaxation (or collocation) method is likely to provide the best approach to solving the optimisation problem with state and control constraints retained, that is, not moved into the cost functional [31]. To represent state constraints in the cost functional, a suitable form of penalty function is needed. A useful continuous and twice differentiable logarithmic form, $-\log(b_i - a_i^T \cdot x)$, can be found in [32]. In the collocation method, differential equations are replaced by finite difference ones and starting values for states and co-states are assumed for each mesh point. Iterative improvement to bring the values into agreement with the finite difference equations and the boundary conditions then follows [33].

If the problem of concern involves minimum-time manoeuvring, the final time is not given. Rather, it is the main constituent of the cost functional. An appropriate modification of the above form can be achieved by deriving the relationship between distance travelled resolved along the track centre line, s, and elapsed time, t, and changing the independent variable from t to s. The differential $(\mathrm{d}s/\mathrm{d}t)$ will account for the combined influences of the path length and the vehicle speed in the revised form and the cost functional will be evaluated over the range zero to s_f, instead of zero to t_f. In connection with minimising the time taken for a flying lap of a complete racing circuit, $\mathbf{x_0}$ will not be known, but the constraint $\mathbf{x_0} = \mathbf{x_f}$ can be imposed.

For a problem of some complexity, it may be difficult to perform the necessary symbolic differentiations to determine the co-state equations and to identify starting values for the co-states sufficient to yield convergence of the iterative scheme to a proper conclusion [34]. Therefore, it is often preferred to solve large optimal-control problems by a direct method. However, major improvements in computational symbolic algebra in recent times might tilt the scales in favour of indirect methods, even for large problems [35].

2.3.2. *Direct optimisation by nonlinear programming*

In the direct approach, the control history is discretised such that control decisions can only be made on an assumed grid. The grid may be time-based or distance-based. In the latter case, the description of a track as a succession of way points is useful. Continuous control histories can be constructed by interpolation over the grid. Determining the optimal controls becomes a finite problem instead of an infinite one. Now over the problem interval, a sequence of control values, $\mathbf{u}_n$, which minimises a multivariable cost or objective function, $J(\mathbf{u}_n)$, subject to equality and inequality constraints is sought.

Initially assuming only m equality constraints, $c_i(\mathbf{u}_n) = 0$ for i=1…m, a Lagrangian is formed by combining the objective function with the constraint expressions through a vector of Lagrange multipliers, $L(\mathbf{u}_n, \boldsymbol{\lambda}) = J(\mathbf{u}_n) - \sum_i \lambda_i.c_i(\mathbf{u}_n) = J(\mathbf{u}_n) - \boldsymbol{\lambda}^{\mathrm{T}}.\mathbf{c}(\mathbf{u}_n)$. Whenever the constraints are satisfied, the constraint expressions are zero, so that the Lagrangian is then equal to the objective function. A necessary condition for a minimum cost is that the derivatives of the Lagrangian with respect to the controls and to the Lagrange multipliers are zero; $\partial L(\mathbf{u}_n, \boldsymbol{\lambda})/\partial \mathbf{u}_n = \partial J(\mathbf{u}_n)/\partial \mathbf{u}_n - \boldsymbol{\lambda}^{\mathrm{T}} \cdot (\partial \mathbf{c}(\mathbf{u}_n)/\partial \mathbf{u}_n) = 0$ and $\partial L(\mathbf{u}_n, \lambda)/\partial \boldsymbol{\lambda} = -\mathbf{c}(\mathbf{u}_n) = 0$. The latter condition is simply a re-statement that the equality constraints are satisfied, that is that the controls are admissible, while the former implies that the gradient of the objective function at the stationary condition must be a linear combination of the gradients of the constraint functions and that the Lagrange multipliers are the coefficients in that linear combination. Allowing inequality constraints in addition, such constraints can be divided into those which are active and those which are inactive at the stationary condition. Each Lagrange multiplier belonging to an inactive constraint must be zero, while each one belonging to an active constraint must be less than zero, or equal to zero if the constraint boundary passes through the (unconstrained) minimum point. Together, these are the KKT conditions [29]. If the KKT conditions are satisfied and, in addition, the curvature of the Lagrangian for all feasible directions is positive, the control is guaranteed to be optimal.

A commonly employed method for solving the direct-optimisation problem is sequential quadratic programming (SQP). With SQP, an approximation to the optimal controls is a starting point and the objective and constraint functions are expanded in Taylor series for the vicinity of the assumed controls. The series are curtailed so that the cost is quadratic and the constraints are linear. An improved control, insofar as the approximate problem represents the real one, is readily computed, so that the full solution can be found by solving a sequence of quadratic sub-problems. Often, in fact, the improved control is used as a search direction, and an empirical parameter is introduced to determine the learning rate. The optimisation becomes one-dimensional and a line search algorithm will suffice. In calculating the improved control, the Hessian of the Lagrangian is needed. It is usually preferred to estimate the Hessian by extrapolation of the results of several prior iterative steps as opposed to evaluating it directly. SQP wisdom is embedded in many computer programs, some gaining advantage from sparseness (SNOPT, http://www.ccom.ucsd.edu), others restricting the search according to feasibility of the controls (CFSQP, http://www.ece.umd.edu).

2.4. *Model predictive control*

Useful textbook support for MPC can be found in [32]. References [36–39] also contain valuable review and tutorial material. MPC has much in common with nonlinear programming insofar as it requires the calculation of a control that will cause a plant to substantially follow a particular trajectory over a specified time ahead. Discrete-time and repetitive computations are involved, with each control value being held constant over a time step. At each time step, a finite problem is implied, with the number of variables involved being influenced by the prediction horizon, the independently chosen control horizon and the time step. If the control horizon is shorter than the prediction horizon, the control input between the two must be presumed. Only the first-step control is utilised, allowing the system state to be updated and the horizons to be advanced by one step, after which the control calculations are repeated. A plant model provides the basis for the calculation of the optimal controls. For application to a real plant, real-time capability is essential, so that speed of computation is important. Employing a relatively simple plant model with a large time step, using short prediction and control horizons, linearising the plant model to accommodate only small perturbations from the current operating conditions at each time step and dealing efficiently with constraints are all important in relation to enabling real-time performance with plants of sufficient bandwidth for this to be problematic. Any of these features favouring fast calculations may compromise accuracy of the trajectory tracking or even stability of the controlled plant.

2.5. *Robust control*

Robust control textbooks include [40–43]. Robust control is a form of optimal control prioritising low sensitivities to variations in the controlled plant and to any disturbances. It applies to linear uncertain systems and involves minimising the maximum gain of closed-loop system frequency response functions relating inputs to outputs. Plant modelling errors are addressed directly and worst-case conditions are presumed in the control design. Outputs that can vary freely within known bounds can be incorporated into a robust control scheme by including a generalised frequency response aspect into the design [17,44]. Loop-shaping is a common feature of robust control designs, allowing stability to be enhanced without prejudice to responses.

3. Selected applications

3.1. *Automotive suspension systems*

3.1.1. *Background*

Active suspensions for ground vehicles were reviewed by Hedrick and Wormley [45]. Other relevant reviews include [7,46–51]. The focus here is on those developments that depend on the application of some form of the optimal-control theory. Terms used are aligned with those defined by Sharp and Crolla [47].

An automotive suspension system is primarily for vibration isolation with limited suspension working space, or rattle space, due to packaging considerations. Rattle space limits are invariably set by elastic stops, which stiffen the suspension markedly for large deflections. Vibration isolation performance and working space provision are bound to be in conflict to some extent. Compliances introduced for vibration isolation have implications for attitude control under load variation and for handling qualities. A single track of an undamaged road surface can be characterised as a Gaussian random processes with approximately constant velocity spectral density over the frequency range that influences suspension design [52–55]. Two typical tracks can be constructed by superposition of two uncorrelated random processes, one representing the centreline roughness, the other representing the cross-level [53–58]. Real roads may be damaged, a typical feature being a pothole. Roughnesses vary widely in terms of amplitude [52,54], such that practical systems must be biased towards worst-case running conditions, or must be nonlinear or adaptive. Accepting the presence of pneumatic tyre, unsprung mass and sprung mass, the system will be capable of body resonance, wheel resonance and tyre resonances, all excitable by different wavelengths in the road spectrum. To avoid excessive resonant responses, damping is necessary. The required damping cannot come from the tyre, since its low rolling resistance depends on it having minimal energy dissipation, so the damping must be provided by the main suspension units joining the wheel to the body. Damping is normally hydraulic and for high-frequency, small-amplitude excitation, an automotive damper will typically become an effective motion transmitter [59]. To provide isolation against such an excitation, dampers are invariably mounted through rubber bushes, to break the transmission path referred to.

Variations in the normal contact force between tyre and ground reduce the capacity of the tyre to develop shear forces for longitudinal and lateral manoeuvring, so that minimising such variations is part of the objective. The effect of the variations is nonlinear in terms of amplitude, becoming pronounced when the tyre leaves the ground, and being strongly dependent on frequency [60]. Deterioration in shear force development will be felt by personnel only when the car concerned is manoeuvring.

When traversing a pot-hole, the support for the tyre is removed temporarily at the hole leading edge and the wheel will accelerate downwards into the hole, effectively driven by a force equal to the static weight supported by the wheel. Subsequent impact of the tyre leading edge with the far edge of the hole causes a potentially strong longitudinal input to the unsprung mass. The isolation system therefore has a longitudinal part, in addition to the obvious vertical part. It may be that this is the main motivation for the common strong bias towards damping in rebound [47,61].

To simplify the picture, the cross-level excitation can be neglected, reducing the car suspension problem to a planar one with a single input, applied first to the front wheel and, a little later, to the rear wheel. To a large extent, when the front and rear suspensions act separately, the front of the car responds to the front excitation while the rear of the car responds to the rear excitation [47]. Making the rear suspension somewhat stiffer than the front leads to the suppression by interference of the body pitching resonance at typical vehicle travel speeds [62–64]. A further important step in simplifying the situation is to ignore this wheelbase filtering action, considering one end of the car to act on its own. The problem reduces to the quarter-car problem in which the frequency range of interest is restricted to cover body and wheel but not tyre resonances, and the tyre is treated as making point contact with the rigid road surface and as being undamped. The quarter-car is an abstraction of the real problem and, in studying its behaviour, it should be remembered that there are several aspects of the real system that are omitted from consideration. In particular, the influence of the suspension arrangements on the handling qualities on smooth roads is outside such studies. In the background is the idea that an active suspension will be capable of levelling and roll control, and that those problems are relatively simple and can be solved separately. It might also be imagined that an active suspension will be capable of adaptation to the running conditions but such ideas need to be tested in some way.

Applications of optimal-control theory to automotive suspensions presume at the outset that the main passive elements of a conventional passive system will be augmented or replaced by actuators. Most often, such actuators are considered to have bandwidth covering the full range of interest and to be capable of giving a force between body and wheel tracking a demand from a controller, notwithstanding that this is far from straightforward [65–70]. Alternatively, the active system can be considered to have a bandwidth limited to around 5 Hz, with the wheel-hop motions being controlled passively. In this case, actuators are displacement producers and it is evident that they are easier to realise and to maintain. The earliest applications actually assumed a one-mass system [18,71], to some extent oversimplifying the real problem [72], but they established an idea that has persisted through subsequent developments, the idea that a sky-hook damper is close to ideal as a suspension element. In terms of the quarter-car, the strict sky-hook is unrealisable since there is no anchor for it, but, as far as the wheel is concerned, a normal passive damper provides such a sky-hook for conditions in which the car body is effectively seismic. Quarter-car results provide a better perspective on the sky-hook damper.

3.1.2. *Full-bandwidth active suspension*

A large proportion of the theoretical work done on active automotive suspensions uses the constant-speed, quarter-car model with an undamped tyre, a road surface whose displacement spectral density function is like that of integrated white noise and a quadratic cost function to represent body vertical acceleration, dynamic tyre load variation from the mean value and suspension working space. In this 'standard' model, root mean square (rms) body acceleration is taken to be a measure of passenger discomfort, tyre load variation indicates prejudice to tyre grip of the road and the three aspects of performance are weighted against each other to show the priorities in the design. Modest variations from this standard occur; for example, the body acceleration may be frequency-weighted to align the discomfort parameter with the ISO recommendations [73], the road spectrum may be bounded for low frequencies to increase the realism in the modelling [11,12], the road vertical displacement spectrum slope may be greater than 2 [16], or the unsprung mass velocity may replace the tyre load variation in the cost function [74] or appear in addition to it [72]. Such variations or realistic variations in sprung mass, unsprung mass and tyre stiffness do not significantly change the conclusions that can be reached from the studies.

The representation of tyre load variations and rattle space usage by squared terms in the cost function is not well aligned with the real problem as described above. In reality, the rattle space would be better included as a power-4 or power-6 term or as a constraint and the penalty on tyre load variations would be improved by incorporation of frequency dependence. However, the 'standard' model fits the needs of the LQR/LQG theory and it is easy and quick to generate optimal controls corresponding to different design priorities. The performance potential of high-bandwidth active suspensions, for running straight on ordinary undamaged roads, can be established well by such means [10,12,72,74–77].

The optimal LQR control is of full-state-feedback form. The manner in which the state variables have to be selected to satisfy detectability and stabilisability conditions implies the need for measurement of the displacements of the body and wheel relative to the ground but sub-optimal systems which avoid this need suffer little performance loss [16,78]. Better than these systems is one in which only the suspension deflection is measured and a Kalman filter is employed to estimate the remainder of the state variables [72]. Full-state-feedback optimal controls typically feature (i) gains on displacements that are spring-like, that is they are of opposite sign and of roughly equal magnitude, and such a feedback can be realised with a conventional spring, (ii) a gain on wheel velocity that implies conventional damper-like forces and (iii) a gain on body velocity that is of opposite sign to that on wheel velocity and is three to five times greater in magnitude.

Performance possibilities in terms of rms values of body acceleration, rattle space and dynamic tyre load variation are shown in many places [11,16,48,49,74,76]. Assessment of these possibilities usually depends on comparisons with passive systems and preferably pays heed to practicality. Passive systems are limited by the need to deflect little under sustained loading from payload change or manoeuvring, while active arrangements have no such restrictions. Further, Chalasani employs a minimum-wheel-hop-damping-factor criterion and both [11,16] effectively use equal-working space comparisons to assist focusing on realisable systems. Chalasani's conclusion that a 15% reduction in body acceleration by changing a passive system to an active one, without demanding more working space and without deteriorating the tyre-to-road contact, is possible, under the more or less standard assumptions, is typical. The main source of difference between an effective active system and a passive one is provided by the body-velocity feedback, which adds damping to the body mode without causing high transmissibility above the wheel-hop frequency as a passive damper would. However, body-velocity feedback leads to a constant low-frequency suspension–deflection asymptote, which would be problematic for positive and negative road ramps. The tyre-deflection transfer function benefits from the suppression of body resonance around 1 Hz but suffers a little towards the wheel-hop frequency. If a design were to prioritise wheel load variation, a control like that of a passive system with fairly stiff spring and unusually high damper coefficient would be necessary [16,79] and this would prejudice the body-acceleration transfer function for higher frequencies [72]. Yue *et al.* also show that measuring only suspension deflection, using a Kalman filter to estimate the other states, costs little in terms of performance and actually brings advantage with respect to the low-frequency asymptote of the suspension-deflection-to-road-velocity frequency–response function.

A fixed-parameter linear-system design must be biased in favour of worst-case operating conditions, to accommodate the working space constraint and to keep the tyre on the ground. Such a fixed-parameter design would not fully utilise its rattle space on good-quality roads and the general inverse relationship between discomfort and working space usage apparent in many results, for example [49], suggests that an adaptive control may be beneficial. A smooth-road design differs from a rough-road one mainly with respect to the effective spring stiffness, which is lower for the former [16]. A very compliant system would need levelling against payload change and roll and pitch controls, but these aspects are implicit in the background to all the present discussions. The suggested basis for an adaptive control is that the on-board suspension-system measurements be used to identify the road qualities and that those be employed to choose between stored control schemes or even to compute a control scheme on-line [80–82]. The possibility for ride comfort enhancement on smooth roads does not seem so exciting anyway and identifying the road in the manner suggested is inevitably a slow and somewhat insecure process, so there are obvious difficulties at the outset. Indeed, fast adaptation would prejudice the quality of the LQR-based control designs [83]. Gordon *et al*. compare their fully researched adaptive scheme against a nonlinear alternative, arguing convincingly that a fixed nonlinear scheme is much to be preferred to an adaptive one of the type envisaged. A fixed nonlinear control can do the same as a standard passive one insofar as it can pay attention to the working space restriction problem when the circumstances demand it, but otherwise need not do so. The findings of Yu and Crolla support this idea.

The key ideas that come from studying the quarter-car with perfect force-tracking actuator are that (i) sufficient information on which to base control action resides in the suspension displacement, (ii) the feedback of body velocity to control force is the biggest source of advantage of an active system in comparison with a passive one, (iii) displacement-dependent control forces should be spring-like, but there is some advantage in making the spring stiff for rough-road running and conversely (iv) a fixed-parameter nonlinear design is better than a linear-system based adaptive one and (v) the lessons from the quarter-car are very likely to apply to real cars. With respect to practicality, making a near-perfect force-tracking actuator is difficult and the actuator dynamics should be accounted for. With the current technology, such an actuator would have to be a servo-valve driven, precision-made, super-clean, energy-hungry hydraulic strut, which would cost too much to make and maintain. Harshness also appears to be an intrinsic property of such systems [84]. Two alternatives are suggested, themselves, namely, a limited-bandwidth active system and a semi-active (dissipative) system.

3.1.3. *Limited-bandwidth active suspension*

Recognising that the main advantage of an active suspension over a passive one stems from the body-velocity feedback to actuator force providing body mode damping, without harming the high-frequency vibration–transmission properties, it is evident that an actuation system with a restricted bandwidth is capable of providing the same advantage. The bandwidth would need to be around 5 Hz and an appropriate actuator could be electrical [13], electro-hydraulic or electro-pneumatic [14], avoiding some of the practical disadvantages of full-bandwidth systems. Highly evolved means for looking after higher-frequency (wheel hop) motions and more or less conventional suspension limit stops appear suitable for dealing with working space constraints. Low-bandwidth suspensions have been produced commercially by Toyota [85] and Nissan [86], developed as prototypes [87], and they are currently standard on several Mercedes cars. Performance estimates relying on output-feedback optimal controls for the actuator and employing equal-working space comparisons [13,14] confirm the capability anticipated, but adaptation of actuator and damper laws to the operating conditions has been shown to be of little value [88]. Detailed energy consumption calculations have shown quite a marked difference in favour of placing an irreversible displacement-tracking actuator in series with a spring/damper parallel combination as opposed to having it in series with the spring, with that combined element in parallel with the damper [88,89]. In the preferred case, the energy consumption appears to be tolerable. Realisation issues in the context of production cars are discussed in [87,90].

3.1.4. *Semi-active suspension*

Another possibility that arises from an understanding of the active system results is that a purely dissipative suspension with a practical level of spring stiffness may provide good vibration performance, while being adequate with respect to sensitivity to load change and handling qualities [16,79]. The semi-active systems considered in these papers involve the assumption that the damper force can track that of an active counterpart, except when the active system supplies energy by virtue of the force and the velocity across the damper being aligned. Then the damper force goes to zero. This control is commonly referred to as ‘clipped optimal’ and the idea that it is of good quality is supported by the observation that the semi-active damper very rarely switches to the ‘off state’ in simulations of random road traversal. In practice, the behaviour cannot be achieved, since real dampers have both low and high coefficient limits and response delays, but it can be approximated. Comparisons of the results show that such a semi-active suspension of adequate stiffness and fixed design can give a roughly 20% lower discomfort parameter, while being similar in terms of wheel load control under random road excitation. Body-velocity feedback is again the source of the performance improvement. If the cost-function weights are adjusted to accentuate ride comfort, on the assumption that working space is plentiful, semi-active designs have only modest capability to respond [79]. What they do is predictable from active-system results. They add some suspension-spring cancelling to the damper-force feedback, encouraging the damper to fight the spring and making the spring appear less stiff dynamically than it does statically. Adaptation of damper-force laws to the running conditions appears not to bring much advantage, for the same reasons as above.

The optimal-control law, assuming unlimited damper capability and minimising a quadratic cost, has been established by Tseng and Hedrick [91] to involve the solution of a time-varying LQR-type problem. The computations were numerical, iterative and necessarily offline. An example showing a 10% advantage with respect to clipped-optimal control is included but the difference between the two is shown to be situation dependent. A ‘steepest gradient’ method, that is operable on-line, is shown to give nearly the same behaviour as the true optimal control, especially when the damper force goes to zero regularly. A valuable feature of semi-active suspension is that the low-frequency asymptote of the working-space-to-road-velocity transfer function cannot be problematic, since the damper is bound to give zero force for zero velocity. Later work [92] allowed the optimal semi-active quarter-car suspension with quadratic cost and damper upper force limit, damper power dissipation limit and a passivity constraint, to be represented as a receding horizon, hybrid MPC problem. The optimal control was shown to be clipped optimal if the prediction horizon was only one time step but this may be improved upon significantly, depending on the circumstances. The optimal control was shown to be of state-feedback form, with switching between gain sets becoming very complex as the number of prediction steps rises, as the operating regime of the suspension changes. Numerical optimisations are discussed below, when road preview is also an issue.

3.1.5. *Inclusion of road preview*

With preview of the road disturbance in front of the car, optimal controls can be established by extensions of LQR methods as described above, as long as sufficient preview is available. With restricted preview, MPC methods could be employed to make the best possible use of what is available. Questions arising are (i) how do the controls relate to those of the corresponding non-preview system; (ii) how much can be gained from how much preview; (iii) in what form does the advantage come; (iv) is it feasible to gather the preview information and (v) in the case of an active suspension, what are the power consumption implications of a preview control. For the full-bandwidth quarter-car, [15,21] apply, but most of the results available relate to limited-bandwidth systems. Hac, for the continuous-time case (see also [22]), showed that the optimal control consists of the same state-feedback as for the corresponding non-preview systems, with a feedforward term of integral form which must be evaluated on-line after observation of the road profile over the preview horizon. The integral contains an exponential weighting which indicates that the stability of the non-preview closed-loop system implies diminishing returns from additional preview. Strictly, the control is optimal only if the preview is infinite, as indicated above, but this diminishing-returns character shows that the control is effectively optimal if sufficient but finite preview is available. In discrete time [23,92], the optimal preview control again involves the same state-feedback as in the non-preview case, with feedforward in the form of a summation of gain-times-error terms, having the same significance as in the continuous-time case. The gain sequences reflect the inverse dynamics of the feedback-controlled plant and they can be found offline, since they are independent of the previewed information.

The most comprehensive performance results appear in [24,93], the first of which adds a state estimator to the system. They show that, with an active-system bandwidth extending beyond the body frequency in a quarter-car case, comfort can be further enhanced by the preview by even greater suppression of the body resonance and that this also further reduces the need for rattle space. This improvement extends to the low-frequency asymptote, with the preview acting in the same way, in this respect, as the estimator of Yue *et al*. [72]. If the bandwidth does not extend to wheel-hop frequency, the preview cannot enhance performance beyond the body-frequency range, so that the wheel hop has to be dealt with passively. As the preview sensor noise increases, so the system performance decays to that of the corresponding one without preview, exactly as might be expected. Preview gain sequences show that, if the bandwidth is sufficient to cover both body and wheel modes, the sequence close to the car reflects the faster dynamics of the wheel mode, while that further from the car shows the slower dynamics of the body mode. When the bandwidth is sufficient only to cover the body mode, the wheel-hop pattern is absent. The preview control can only contribute to improving the body mode, see also [15,94]. Correspondingly, more preview is needed for the improvement of a low-frequency mode, typically 0.25 s for a 1 Hz mode and 0.1 s for a 10 Hz mode. In the correct circumstances, quite substantial cost reductions occur, but obtaining sufficient and accurate road-preview data seems to be out of reach currently. Another deterrent to trying to use preview to improve performance is that energy consumption increases systematically as the cost reduces [83]. It is evident that with the more practical limited-bandwidth suspensions, even wheelbase preview, not having much cost implication, is of little value for usual vehicle speeds.

In the semi-active-suspension case, energy consumption is not an issue. The relationships between hardware specifications, amount of preview available and performance advantage are of concern and these are exposed in [95]. The system considered is a quarter-car with fixed practical spring rate and damper in series with a spring representing a rubber mount for high-frequency body isolation. The damper has a realistic force/velocity map, limited in terms of both high- and low-forces, and its dynamic responses are governed by a low-pass Butterworth filter with 40 Hz undamped natural frequency. Open-loop optimal controls minimising the usual quadratic cost are found by the application of Pontryagin's principle, with some economies of computation from discretisation of the control sequence and approximation of the final-time boundary state. Through simulations with white-noise road vertical velocity excitation, the method is shown to give the same results as known independently for linear active suspensions with either no preview or infinite preview. Bar charts show cost- and cost-component reductions as functions of preview time for full active, semi-active optimised and semi-active clipped-optimal-controlled systems and the optimally controlled damper is shown to spend a greater proportion of its time at the soft and hard limits than the 'clipped' one. Benefits from preview are not as great as with an active system with similar responsiveness. They come mainly from the first 0.1 s of preview and the improvement is primarily in wheel-load control. Clipped-optimal control loses ground in comparison with the true optimal control as the preview increases but it always gives behaviour that is not much worse than the best that can be done.

3.1.6. *Optimal nonlinear suspensions*

It should be clear that a real suspension system must be nonlinear, the imperative arising from the working space constraint. A lesser factor is the tyre leaving the ground. It may be that a substantially linear system design with conventional elastomeric limit stops is near optimal, since such a system will apply forces to limit suspension travel only when they are needed. It appears not unlikely that doing the same sort of thing through the control law will bring little benefit. Evidence on the issue is not available from the literature, so that nonlinear suspension design methods remain of interest. Gordon's approaches, in one of which different LQR designs are mapped into a fixed nonlinear scheme [25], and which has already been mentioned, and another involving the use of learning automata [96] are available. Another obvious possibility is that a control scheme can be designed on a heuristic basis and parameter optimisations can be employed to iteratively improve performance [97]. Devising nonlinear controls by means of the SDRE technique is illustrated in [98].

3.1.7. *Summary*

Even though suspension system descriptions by randomly excited linear system theory are only approximate, such a theory has been instrumental in yielding a good understanding of the possibilities of advantages available from active systems and of the main sources of advantage. Spin-off in the form of a greater appreciation of traditional passive designs has also occurred. The major benefits have come from quarter-car-level studies, with more elaborate treatments largely just confirming the ideas that quarter-car studies have established. Simple measurements are likely to suffice. Due to practical considerations, slow-active and semi-active systems are more important than full-bandwidth arrangements in relation to commercial exploitation. Even for semi-active suspensions, linear quarter-car theory has been important, but it has also been established that clipped-optimal control can be improved upon, but not by much. The potential benefits from using disturbance preview have been shown but, to get enough preview to gain significantly, requires sensing the road surface with the ability to distinguish between hard and soft surfaces some distance ahead of the vehicle, which cannot be done in usual operating circumstances. An effective basis for adaptation of control laws as conditions change has not been discovered. Nonlinear designs with amplitude dependence, especially dealing with working space limitations, are likely to suffice without adaptation.

3.2. ***Worst-case manoeuvring***

3.2.1. *Background*

Given time and numbers, automotive customers will find how to provoke the most extreme behaviour of a vehicle, sometimes inadvertently. Consequently, it is desirable that manufacturers themselves understand that behaviour and how to get it. Historically, trial-and-error methods were employed to find and define manoeuvres which would fully test a product. Notable failures have occurred [99] and been much publicised. More recent examples can be found on the Internet (http://www.autozine.org), (http://en.wikipedia.org/wiki/Rollover). The US National Highway Safety Administration supported work to devise and test methods for finding the worst-case controls in the late 1960s [100], leading to a 'drastic steer and brake manoeuvre'. A fundamental difficulty with this methodology is that the worst-case control inputs depend on the vehicle and possibly the road characteristics, so that work naturally focused on contemporary vehicles, may not provide long-standing solutions. A more general, practical approach is needed.

3.2.2. *Analysis and prediction*

Analysis of worst-case manoeuvring aims to answer the following question: given a dynamic system, what is the worst-possible disturbance input that maximises a given cost function. Solving this problem is useful, for example, in identifying rollover propensity of a vehicle or weakness of a vehicle stability control system. Despite this problem appearing quite similar to typical optimal-control problems – many of which are discussed and reviewed in this paper – there are several key differences:

(1) Instead of identifying optimal-control signals, in this new problem the aim is to find optimal disturbance signals. The dynamic system that the disturbance is 'sabotaging' may or may not also be subject to control inputs.
(2) Instead of minimizing a cost function, the goal is maximizing, e.g. to destabilise a vehicle. When the cost function to be maximised is quadratic, the problem becomes concave instead of convex. Therefore, the maximizing solution might not be unique, and the solutions are either at infinity (for unbounded disturbances) or at the input bounds (for bounded disturbances). The possibility of a non-unique solution as well as the fact that the solutions typically lie on the boundaries make the worst-case problem much harder to solve numerically than typical optimal-control problems. This is because convex problems have smooth gradients leading to a unique global solution, which is not the case for the worst-case problems. Also the problem is inherently constrained rather than unconstrained.
(3) Because of the non-convexity issue discussed above, worst-case manoeuvres involve large excitations. Thus, the underlying vehicle dynamics must be accurate even in nonlinear regions. The requirement on model uncertainties is also greater than for typical control problems.

Due to the fundamental differences described above, worst-case manoeuvring problems are still in their infancy and few papers exist in the literature. Researchers in the vehicle dynamics and control community, however, must take the leadership role in this important field. We will elaborate on the importance of the problem by reviewing the limited literature in the following.

The safety performance of vehicles, such as rollover propensity, is usually assessed by government agencies through well-defined standard tests. The safety performance is typically scored in the form of clear quantitative schemes, such as the star ratings system used in the US New Car Assessment Program (NCAP) [100], Euro NCAP (http://www.c-ncap.org.cn/C-NCAP/index.htm) or the China C-NCAP (http://www.c-ncap.org.cn/C-NCAP/index.htm). As new safety control systems such as antilock brake systems (ABS) and electronic stability control (ESC) are introduced to vehicles, government agencies such as NHTSA and Euro NCAP are forced to come up with new methods to evaluate performances of vehicles equipped with those systems. The standard tests developed to evaluate the performance of vehicles without the control systems are not adequate for vehicles with these new control systems. NHTSA and Euro NCAP now face a new and difficult challenge in designing simple, repeatable and reliable ways to assess the safety performance of vehicles with smart chassis control systems.

The need for a new assessment tool can be understood through examining the current test procedure. In the USA, rollover propensity is assessed through a 5-star rating system based on the static stability factor (SSF) of the vehicle along with modification based on a dynamic test result. The fishhook manoeuvre is selected for dynamic testing by NHTSA. With the advent of active safety control systems such as ESC, these standard tests and approaches to identify vehicle rollover propensity face a new set of circumstance as follows: Starting in the late 1990s, ESC systems quickly penetrated the market as an active safety device. Automotive companies soon realised that ESC is a relatively cost-effective way to boost the rollover star ratings of SUV's and light trucks. If wheel lift-off occurs during the given sine-with-dwell test, the safety score will be reduced. Instead of redesigning the chassis or repackaging, vehicle developers can simply calibrate the ESC systems to prevent wheel lift-off under the test conditions and thus improve the rollover star rating. The Euro NCAP system also uses the sine-with-dwell manoeuvre for its ESC performance test.

The fundamental problem of this evaluation process is analogous to assessing the learning of students. Traditional standard test procedures are similar to announcing the exam questions ahead of time. It is possible that some 'students' will do a great job answering the exam questions but otherwise learn very little about the rest of the course material. As a side note, it is interesting to see that all cars with ESC received a 'pass' grade in the Euro NCAP test in 2009 because they 'prepare well for the test questions'. The evaluators, NHTSA or Euro NCAP, need to find a new way to assess vehicle safety without relying on their performance on a single and pre-announced test.

The worst-case evaluation concept was suggested as a possible alternative way to assess the performance of dynamic systems (including vehicles) equipped with active control systems under a broad range, instead of only a small set, of manoeuvres [102,103]. The proposed method uses computer simulations to systematically identify worst-case scenarios, i.e. potential cases when the active safety systems fail to perform satisfactorily. This problem is essentially a trajectory optimisation problem to solve for a sequence of disturbance inputs (e.g. braking and steering) that maximises a cost function (e.g. roll angle). Several prior research programmes investigated mathematical approaches for optimal disturbance generation. When the dynamic system is linear, the worst-bounded inputs can be obtained easily through the signage of its impulse response [104]. For nonlinear systems, a solution of the Hamilton–Jacobi–Bellman equations is necessary to discover the optimal trajectory [105].

The work presented in [102,106] focuses on rollover and spinout of vehicles using the worst-case generation methodology. A key technical challenge is to obtain good initial guesses so that the subsequent numerical search can successfully find the local optimum. Possible approaches include starting from standard manoeuvres such as a fishhook path and sine-with-dwell inputs, taking initial conditions from 1-player optimal-control solutions or using two-player game theory [107], or from linear systems theory [104,108].

Instead of generating good initial guesses and then using a numerical approach to search for a local minimum in the current neighbourhood, it is also possible to solve the worst-case disturbance problem using *bona fide* global searching methods. In principle, brute-force search methods such as dynamic programming can be used to guarantee global optimality. In practice, however, due to the large number of states and input signals, it is typically not feasible to solve vehicle worst-case problems using contemporary computing machines. In [109], the iterative dynamic programming (IDP) technique was suggested as a feasible alternative. In IDP, very coarse grids are used and the initial ranges of states and inputs cover the full problem space.

The search space is subsequently reduced to focus on regions with high potential based on the initial results. By using this iterative approach, a *bona fide* brute-force search can be used to find a global optimum at reduced computation cost. However, because of the coarse grids, the optimality of the obtained results is not guaranteed. In [110] two local search methods, Mesh Adaptive Direct Searching and SQP, are used to find worst-case manoeuvres for an integrated chassis control (ICC) system. Despite the fact that the global optimality of the results cannot be guaranteed, it is encouraging that the obtained worst-case manoeuvres consistently result in vehicle instability (e.g. rollover) at much lower steering inputs than those identified by other methods, and weakness of the targeted ICC system can be systematically identified.

3.2.3. *Summary*

The need for worst-case vehicle assessments in the interests of passenger safety is established. Comprehensive assessments of this kind are almost impossible by testing and very difficult by simulation methods. Numerical techniques from optimal-control theory have to be employed. They have to be associated with quite detailed and accurate vehicle models which, in a contemporary context, implies the inclusion of electronic chassis control system representations. Successful trials have been reported, but further work is needed to refine the procedures and to bring them to everyday use.

3.3. ***Minimum-time manoeuvring***

3.3.1. *Background*

Vehicle dynamics by simulation is a key technology for all advanced circuit racing teams. The ultimate objective is to predict with accuracy the outcome of testing and to some extent racing itself, avoiding the need for extensive and expensive experimental work, which may be limited by the rules, and to aid understanding of how to design and operate the fastest feasible vehicle, with proper regard for the track. The tools commonly employed are clearly useful but improvements are both needed and possible.

3.3.2. *Studies of minimum-time manoeuvres*

It is likely that the first treatment of minimum-time manoeuvring as an optimal-control problem was by Fujioka and Kimura [111], although the possibility of such a treatment had been mentioned before [112]. They used the Sequential Conjugate Gradient Restoration Algorithm (SCGRA) [113,114]. The method involves the minimisation of a cost functional and possibly parameters, while ensuring satisfaction of differential or non-differential constraints and boundary conditions, through alternating between gradient descent and restoration. SCGRA is a direct method, producing a sequence of feasible solutions each with cost lower than that of any prior solution.

The manoeuvring problem was set up such that the car decelerates with tyre-friction-limited rate in a straight line initially and accelerates correspondingly after the curve, passing through specified speeds at the beginning and end of the curve. The path was unbounded in the curve, having a large angle between starting and finishing lines, so that the car was obliged to travel in a U-shape to satisfy its boundary conditions. The optimisation objective was to minimise the time for the whole manoeuvre. The car had just two wheels and longitudinal, lateral and yaw freedoms, apparently with no longitudinal load transfer. Tyre forces were limited by a ‘friction circle’ treatment with a specified constant friction coefficient. Cases ranging from four-wheel-drive and four-wheel-steer to two-wheel-drive and front-steer only were computed, with variations in the longitudinal location of the mass centre between front-heavy and balanced. Steer angles were reverse-engineered on the basis that the tyre forces were proportional to sideslip angle until they reached their friction limit, modified by a curious allowance for longitudinal force that reduced the angle for driving force, while increasing it for braking force. Simulated behaviour was reasonable with modest differences between cases but the assumptions made seem to preclude gaining much understanding of real vehicles.

Hendrikx *et al*. [115] made major improvements to the prior work. Their car had the same three degrees of freedom, longitudinal, lateral and yaw, but had an engine torque model, nonlinear tyre forces, body-aerodynamic forces, load transfer by statics according to longitudinal and lateral accelerations of the mass centre, and a track defined by its centre-line and width. The optimisation was indirect, relying on Pontryagin's principle, and involving the setting up and solution of a two-point-boundary-value problem in states and co-states. The latter equations give the history of the Lagrange multipliers, which can be used to iteratively lower the cost by improving the controls. Specifically, the method was the Gradient Method, described in [5,30].

Tyre forces were represented by a pure-lateral-slip Magic Formula [60] with a 'friction ellipse' modification to deal with simultaneous longitudinal and lateral forces. Longitudinal forces were distributed evenly between the right and left sides, with examples assuming a rear-wheel-drive Formula One-type car and a front-wheel-drive medium-sized European saloon car. Braking forces were distributed between front and rear according to a fixed ratio, in each case. The problem was transformed so that the independent variable became distance along the track centre-line instead of time, to facilitate a fixed-distance, minimum-time optimisation. Controls consisted of longitudinal force, limited by the powertrain or the tyre friction, assumed limited by a known friction coefficient and the calculated normal load on each tyre and the steer rate with range corresponding to a driver's capacity. Rather than imposing the path restrictions as constraints, they were incorporated into the cost functional, such that the square of the distance of the car mass centre from the track centre-line multiplied by a step function having the value zero when the car is within the track bounds but otherwise having a significant non-zero value, was added to the time increment. Lane-change manoeuvres were optimised and simulated, with verbal descriptions of the motions but no information on the important issue of computation times.

Allen considered a problem similar to that of Hendrikx *et al*. but used a direct-optimisation method, the nonlinear programming, direct shooting method of Kraft [116], to obtain solutions [117]. The car model was the same as that of Hendrikx with minor variations with respect to lateral load transfers, but a complete flat and level circuit was designed using straight sections, circular arcs, elliptic sections and error functions. The requirement for the vehicle to remain on the track was retained as a constraint and, although in principle the control constraints would include drive-torque limiting, the speeds and path conditions simulated were such that the engine was always sufficiently powerful to spin the driving wheels or the brakes strong enough to lock the wheels. Therefore, the tyre friction limit always provided the bounds on longitudinal forces. Allen also used a time-to-distance transformation, with control values specified for equal-distance intervals along the track and cubic-spline interpolation to give intermediate values. A car-design parameter vector was included in the problem formulation but was not actually used to generate results.

With direct shooting, the initial conditions and an estimate of the controls are used to find the trajectory of the car over the period of interest, by numerical integration of the equations of motion. Allen needed a few iterations to obtain each tyre force, since the force depends on the load and the load depends on the car accelerations, representing a circular dependency. The cost functional, the finishing state and the constraint violations are evaluated. Computation of road boundary constraints depends on comparison of the current car position with the road edges, involving some interpolation between road-data points. In turn, each control is varied by a small amount and the numerical integration is re-run, from one track point before the variation to the end point, and the evaluations are repeated. Forward differencing gives the gradients of the cost functional, constraint violations and final state values to variations in the controls. Hessians are estimated from successive values of the gradients, once a few iterations have been done. On completion of these trials for all the control values, a SQP scheme is used to improve the controls and the whole procedure is repeated to convergence within a specified tolerance.

Running the same minimum-time lane-change case as Hendrikx, with the same car data also yielded very similar results and a half-track was treated with apparent success, but attempts to complete a full lap failed, despite that optimisations of each track segment separately worked well. Casanova [31] later found the probable cause to be extreme sensitivities of track-constraint violations to control variations for track points far ahead of the vehicle, which a real driver would treat as irrelevant to current control decisions.

Optimal-control theory was applied to motorcycle manoeuvring in [118]. An unusual feature of the work is that the motorcycle equations of motion were obtained in an implicit form, making the application non-standard. Inequality constraints corresponding to staying on the track were considered easier to deal with as penalties included in the cost functional. An IMSL library program BVPFD, now apparently extinct (http://www.vni.com), was initially employed but the necessary reduction in the motorcycle equations to state-variable form was found expensive, irrespective of whether it was done once symbolically or repeatedly by numerics. Better performance came from using SOLVDE from [33], which will accommodate the implicit form of the equations of motion. The motorcycle model had only one body with longitudinal and lateral, yaw and roll freedoms and three controls, the rate of change of rear tyre driving force, the rate of change of front tyre lateral force and the braking force front to rear ratio for braking phases. The cost function finally contained terms to maximise distance travelled in a given time, with penalties on path errors, tyre forces and the first two controls. Results for a short section of the Mugello racing circuit were shown. This work was extended later [35], with emphasis on difficulties that arise if the equations of motion do not come in a state-variable form, computational algebra tools for aiding the preparation of the problem, capitalising on sparsity in the matrices and employment of a more realistic motorcycle model. However, the model tyre force treatment remains quite rudimentary. The influences of changing the starting state of the motorcycle in a hairpin manoeuvre are shown and a whole circuit of the Adria racing circuit is optimised.

The method was developed with a view to real-time performance [119] and applied to robotic and driver-assistance problems. Again, differential equations in implicit form were assumed and inequality constraints were moved into the cost functional using barrier functions. Three forms of barrier function, each using parameters to control the detailed workings, were discussed. The main point of departure from the prior method was in the solution of the transformed problem by a finite difference scheme and algorithms are described for aiding successful outcomes.

Casanova [31] found minimum-time controls for a model of a then-current Formula One car in a double-lane-change manoeuvre and for the Barcelona and Suzuka racing circuits. The car model had longitudinal, lateral and yaw freedoms for the body and a wheel-spin freedom for each of the four wheels. A combined-slip Magic Formula tyre model [60] and a standard treatment of aerodynamic drag and down-forces were employed. Tyre loads were computed by statics, allowing for weight, down-force and longitudinal and lateral accelerations, with lateral-load-transfer sharing between front and rear axles allowing for simple suspension geometric and stiffness properties. The model included engine torque as a numerical function of speed and throttle opening with interpolation between table points, discrete gear ratios and a torque-sensitive, limited-slip differential gear in the rear axle. The front to rear braking torque ratio was fixed by a parameter. The longitudinal control was confined to the range (-1 to +1) with negative values implying braking and positive ones implying driving the rear wheels. The lateral control was the assumed-equal steer angles of the front road wheels. The model was built in state-variable form with the aid of VehicleSim Lisp (formerly AutoSim, see http://www.carsim.com).

The relative merits of indirect and direct methods for minimum-time problems were considered, recognising that constraints that may be difficult to deal with in an indirect approach can easily be translated into additional terms in the cost functional and they need not prejudice indirect methods. Nevertheless, he found a mismatch between the assumed continuity of indirect methods and the accelerating to braking switch that is a feature of preparation for a corner on a typical racing circuit and saw a chronological trend towards direct methods for large-scale problems, with much support in the form of computational suites. Having decided on a direct approach, he sought advantage in distance-based independent control decision grids for longitudinal and lateral controls, with close spacing where the control activity is likely to be high and conversely, capable of minimising the number of controls to be optimised, but requiring some good judgements to be applied by the analyst. Linear interpolation between control-decision points was used to refine the control histories. He applied a transformation to change the independent variable from time to distance, so that the time taken for a specified distance could be the main constituent of the cost function and he also applied track-boundary constraint conditions which operate only on short sections of track, where they are likely to become active. These boundary constraints were of the continuous and twice differentiable form $\{(2d_{max}/w_t)^2-1\}$ where d_{max} is the maximum lateral separation of the car's reference point and the track centreline on the section and w_t is half of the track width in the section. Again, with good judgement in the problem definition, some economy of computation, without affecting the optimisation, can be obtained. To avoid two problems in particular, the first that the sensitivity of track-boundary-constraint violations to variations in controls tend to grow artificially with the longitudinal distance between them, the second that a variation in a control with a race car near to its cornering limits may initiate a spin and spoil the calculations if the spin is allowed to develop, the race circuit was segmented with a rough rule that each segment would be traversed in about 1 s. Segment boundaries were chosen to be near corner apexes. The difference between the end state for segment (i) and the beginning state for segment (i+1) was then constrained to be zero at the solution point.

Parallel shooting [120,121] was identified as a matching strategy, with potential for solution by parallel computation, although this was not attempted. Repetitive simulation to find gradients by finite differences was found to use the bulk of the computation time and a much-preferred alternative in the form of automatic differentiation (AD) was found [122,123]. SQP solutions were obtained through the use of the code SNOPT [124], which focuses especially on exploiting sparseness to minimise the computations needed. SNOPT requires the employment of smooth functions in cost and constraints. Numerical trials for a 300 m long track including a single 70° turn with longitudinal and lateral control decision points at 12 m intervals and using a somewhat simplified car model showed AD to be more than 10 times faster and more accurate than the finite-difference alternative.

Many numerical trials were needed to establish the circumstances under which good solutions could be obtained. Trials involving the integration step lengths and tolerance settings were numerous and skill in interpretation of SNOPT's iteration log, with data on the learning step length and the conditioning of Lagrangian and Hessian functions was needed. A benefit from track segmentation was that fast segments could be integrated with a coarser step than slow ones. Track segmentation details were non-critical but if control-decision-points were separated by more than about 5 m, the best lap time achievable started to grow. Variations in yaw inertia over a large range were explored in a double-lane-change manoeuvre [125], the sensitivity to mass variations of the lap-time for complete circuits of Barcelona and Suzuka was successfully computed [126,127] and optimisation of the longitudinal location of the car mass centre for each of the two circuits was studied [128]. Study of a rear-drive rally car, with appropriate parameter value changes to reflect the lower engine power, the tyre forces on a typical loose surface, reduced aerodynamic down-force etc., performing a double-lane-change showed the characteristic oscillatory pre-steering, 50 m ahead of the main geometric feature of the turn, and later opposite-lock steering control, often observed in reality. Imperfections of the method include a modest spread of results from repeated computations with different starting points, the length of the calculations and the problem preparation time but results obtained compared well with corresponding measured ones. An interesting feature is that the theoretical steering input spectrum is much richer in components above 0.5 Hz than the practical ones, which show real driver control largely inactive for higher frequencies. Results suggested that linear interpolation between control points is restrictive and that the problem naturally tends to contain many local minima, as racing-line changes are compensated by control variations.

Kelly's quite recent work [129,130] was directed towards the treatment of the vehicle concerned as a black box, the optimiser having access to only those model states which are likely to be common to all models, namely (i) the absolute position coordinates of the car mass centre; (ii) longitudinal and lateral velocities of the car mass centre; (iii) yaw rate; (iv) lateral acceleration; (v) longitudinal slip ratio at each wheel and (vi) load on each tyre. The car and its controls were similar to those of Casanova and the approach was direct. However, more constraints were used to control the search space, avoiding looking where a knowledge of the problem would indicate the solution not to lie. An interior-point code, Feasible Sequential Quadratic Programming (FSQP), was employed (http://www.ece.umd.edu) as a shell, with forward finite differencing being used to generate gradients, like Allen. A solution for short manoeuvres was developed first and then a way of extending the manoeuvre to an unlimited degree was devised through a receding-horizon strategy.

The manoeuvre is defined by suitably spaced waylines, each being perpendicular to the track centreline and having a defined length, representing the width of the track. At each wayline, longitudinal and lateral controls had to be found and track-boundary constraints had to be satisfied. Distance-based linear interpolation gave control values between waypoints. Constraints in the form $\{(z/z_{lim})^2 - 1\} \leq 0$ were applied to tyre longitudinal and lateral slips, a so-called understeer angle, being the difference between the steer angle and that required from a purely kinematic point of view, and the wheel loads. At each iterative step, (i) FSQP constructs a local quadratic model for the cost function using the computed gradients and an approximate Hessian through monitoring the course of the gradients through prior iterations; (ii) Newton's quadratic approximation method [29] is used to generate a search direction. It may be that this search direction is infeasible due to constraints; (iii) a feasible direction is generated; (iv) the two directions are combined to yield a feasible descent direction and a line search in this direction is followed until a minimum cost or a constraint boundary is reached.

Solutions for a 90° right turn followed by a 90° left turn showed diminishing returns for waypoint separations less than 3 m and for horizon length greater than 200 m. A nice example showing too much speed early and too little late as a consequence of too short a horizon was run, with a longer horizon giving a better solution. A complete lap of the Jerez Formula One circuit was optimised, with realistic results. One hundred repetitions of a 250 m long, 90° turn optimisation [128] revealed two solutions, with 64 runs leading to the better solution and 36 to the worse one. Within a group, the spread was 0.02% but between the groups, the difference was 0.2%, confirming the existence of local minima in the minimum-time problem. Using a Fourier Transform basis for the controls led to the same results. Calculating the sensitivity of the manoeuvre time to vehicle mass, using 1 kg steps between 600 and 700 kg, clearly demonstrated two distinct solutions for the majority of mass values but similar computations for the chicane on the Jerez circuit and Beckett's corner at Silverstone showed only one solution for each mass setting. Two distinct solutions also occurred for the corner manoeuvre with horizon between 250 and 270 m but not outside this range. Reducing the friction level of the rear tyres in this manoeuvre caused a systematic increase in the spread of solutions.

3.3.3. *Automatic differentiation*

The calculation of derivatives of the cost function and constraints with respect to controls in the direct methods is the most time-consuming part of the minimum-time process and the indications are that AD is a key technology in connection with efficiency. Therefore, an outline of how AD works and of how to use it is now included. Forward and reverse accumulations are alternative strategies for AD. With forward accumulation, each variable carries with it a gradient vector, specifying its partial derivatives with respect to all the independent variables involved in the minimisation. In the evaluation of a cost or constraint function, new variables are formed from old ones by function operations, e.g. sin or log, which are unary, or by combining old variables, e.g. by multiplication, which are binary. The rules by which the individual gradients change in these elementary operations are 'known' to an AD program, so that the gradient information is just updated every time an operation occurs. Finally, when the cost or constraint is evaluated, so are its derivatives. The rules come from symbolic differentiation, e.g. of sin or log, but all the evaluations in the forward accumulation process are numerical. Forward accumulation is quite straightforward but is extravagant with respect to processing a large proportion of zeros, representing a complete lack of dependence of one variable on another one.

In reverse accumulation, as a function evaluation proceeds, a computational graph keeping a record of which variables participated in each operation is constructed. The graph has a node for every value held by a program variable and an adjoint vector containing the partial derivatives of the value with respect to all of the prior nodes is associated with the node. In fact, since each value is the result of either a unary operation on one prior value or a binary operation on two prior values, only one or two derivatives will be non-zero. Adjoint vectors are computed in reverse order with respect to the function evaluation, employing the chain rule to go back from the function at the end of the chain through the steps that led to the function. Construction of the computational graph is relatively difficult and storing the full graph may be problematic. In [31], the computational graph was constructed track-segment by track-segment due to this difficulty. However, processing irrelevant zeros has become unnecessary. AD software will normally contain the capability to construct and record the computational graph [123,131], see also http://www.Autodiff.org.

3.3.4. *Summary*

Quite realistic minimum-time lap problems can be solved but the solution times are sufficiently long to restrict the usefulness of the approach. This is true despite that the vehicle models used compromise significantly between detailed accuracy and fast simulation. Ideally, a contemporary racing car model would include bounce, pitch and roll freedoms, suspension geometric properties, a comprehensive-force and moment aerodynamic map and a thermodynamic tyre [132,133], such that track irregularities, the aerodynamic influences of attitude and steer angle changes and the heating and consequent friction variations of racing tyres could be represented. Including such features would add substantially to the necessary computation times, so that the search for the fastest accurate solution possible is far from over. Solutions that have been obtained give the racing line and realistic minimum-time controls which are impressive but they also indicate the sometimes presence of local optima and often strong dependency of the computation time on features in the problem representation like constraint details. The smoother the computational problem can be made, the better it is from this point of view.

Indirect methods require the problem to be described as sufficiently smooth that derivation of the adjoint equations is feasible. Representing track position constraints as costs employing barrier functions seems not to present any difficulties and indeed gives opportunity to choose descriptions of the constraints that best align with the real situation. In the method, substantial advantage seems to accrue in that the gradients, which are crucial to the optimisation procedure, are calculable very efficiently. Standard procedure involves the control update step being the same as the discrete-time step for solution of the two-point boundary value problem and normally learning by steepest descent is employed. It may be possible to improve performance by varying these practices, using a coarser control grid and SQP-based iteration, for example. An alternative to solving the two-point boundary value problem is to use a finite difference scheme but it is not known whether this is better or worse in the minimum-time manoeuvre context.

To be competitive in terms of speed of computation, it appears almost essential in direct methods to use AD to evaluate derivatives, but it seems not beyond the bounds of possibility that the adjoint equations of indirect methods can be used in conjunction with a direct approach. Also, derivatives can be found by integrating sensitivity equations but the computational implications are not known. Greater freedom in representing the minimum-time problem mathematically exists with direct optimisation but it remains likely that advantage will come from making the theoretical problem as smooth as possible, without spoiling it. Linear interpolation to find controls between decision points is possibly restrictive, causing the control grid to be finer than it would need to be with, say, cubic-spline interpolation. Use of non-obvious constraints, on front and rear axle sideslip ratios, for example, to avoid searching in areas where the analyst will know the solution not to lie, and using a feasible search or interior point methods for iterative improvement seems advantageous, representing part of a general philosophy that one should not oblige the optimisation calculations to find, possibly repeatedly, what one already knows about the solution.

In both indirect and direct cases, the way that the track boundaries are portrayed is interesting. It is understood that long track sections cannot be treated straightforwardly, due to the high sensitivity of late track-boundary constraint violations to early controls. Segmentation of the track and repetitive receding-horizon calculations have provided solutions but these are possibly not the best available. If a snapshot of the vehicle motion is taken, it should be clear that any control changes at the present time cannot affect excursions from the track at the present location, but only ahead of the present location. Also, control decisions now will not sensibly be based on the track geometry a long way ahead. There comes a point where the track ahead has diminishing importance for control decisions affecting the present time. Driver modelling with track preview suggests that track-boundary constraints should be modelled with a weighting function shaped like a sine wave stretched out using a stretch function similar to that used by the basic Magic Formula for tyre forces, such that the track 0.1 s ahead of the car (or 0.6 s for a motorcycle) is the most important and the track a long way ahead has no importance at all. Inclusion of such a weighting may make the optimisation of a whole lap without segmentation or repetition possible. The evidence is strong that a minimum-time lap problem will possess local minima close to the global minimum, which represents a fundamental difficulty with respect to making the calculations converge quickly to a reliable best solution. It may transpire that decoupling the calculation of the racing line from the calculation of the optimal speed profile will provide the route to fast calculation of minimum-time manoeuvres.

3.4. *Driver/rider modelling*

3.4.1. *Background*

Review papers on driver modelling have been written by Guo and Guan [134], MacAdam [135] and Plöchl and Edelmann [136]. The first makes 36 references, the second includes 157, while the last contains 215. The subject in total has become quite large. Here, the concern is with the contributions to modelling techniques and to the understanding of driving and riding which have come from the application of optimal control theory. This implies a focus on highly skilled driving and little involvement with errors and accidents. Review material specific to motorcycles can be found in [137].

Effective driving and riding are made possible by learning, the most likely form being reinforcement learning. The learner simply repeats those actions that bring a desired result and avoids repetition of the converse. It is reasonable to associate serious and sustained learning with optimal control [138]. Both longitudinal and lateral controls are necessary, although for simplicity it may well suffice to consider only steering control initially. Relevant vehicles include cars, motorcycles, bicycles and unicycles. Commercial vehicles, combinations such as cars and caravans, trucks and trailers, are outside the present scope. For simplicity, the discussion here will concentrate on cars. At low speeds, cars are kinematic devices with few parameters governing the differences between one and another but as they travel faster, the dynamics become more important and the car driver must learn the changing dynamic responses and adapt to them. The driver must be able to see the road ahead and, with a road of any complexity, it is advantageous to know the road so that viewing it provides a reminder and not new information. The faster one travels, the further one needs to see ahead to be effective. Effectiveness is associated with accurate path following and low effort, and drivers learn to balance accuracy against effort, according to circumstances. Cars driven near to their physical limits caused by tyre force saturation can become unstable or uncontrollable, making effective driving very much more difficult than it normally is. Under some operating regimes, vehicles are oscillatory and one should expect the driving to reflect that behaviour. Examples include the rally car, running at high speed on loose surfaces, and the high-performance motorcycle running at high speeds and typically possessing a lightly damped weave mode. Over most of the speed range, motorcycles and bicycles need stabilising by the rider. Unicycles need such stabilisation at all speeds, both longitudinally and laterally. Guidance by single-point preview is known to be possible in some cases. A drawbar-steered full trailer (with a steering dolly) has mechanical steering which involves effectively a proportional control of steer angle to lateral position error of the drawbar at the pintle-pin, as seen from the trailer. Such steering gives good tracking and oscillatory behaviour when the drawbar is short and conversely [139,140]. Single-point preview is clearly sufficient in some circumstances but is unlikely to be universally capable, even with fuller use of the error [134].

3.4.2. *Models*

The application of optimal-control theory effectively starts with MacAdam [141,142], who, apart from recognising the possibility of good-quality steering control being seen to involve optimisation, saw the crucial role of look-ahead or preview information in driver control of a road vehicle. The majority of previous approaches to driver modelling came out of aircraft pilot and other man–machine studies in which the pilot's task was seen to consist in regulating against disturbances which could not be anticipated. The driver would typically be represented as a transfer function relating lateral position error to steer angle or by nonlinear input–output blocks [143–146]. Some studies suggested that drivers maintain an internal representation of the vehicle and use plant inversion to arrive at a control [147,148].

MacAdam's method depends on a simple form of MPC. The vehicle must be linear, the control is presumed constant over the prediction time and the cost function is an integral over the prediction time (the preview horizon), or a corresponding summation, of the weighted squares of differences between the intended path and the predicted path of the vehicle reference point. The path prediction is the superposition of the free vibration decay (assuming the vehicle is stable) and the forced vibration response to the change in steering input. Equating the differential of the cost with respect to the control input to zero allows the control that minimises the cost to be found quickly. A special case of the theory arises when the weighting function is the Dirac delta function, causing only one path error to contribute to the cost and constituting single-point preview. MacAdam applied his single-point-preview control to the calculation of frequency response results for comparison with similar ones from the literature. Quite good agreement was achieved but the optimal controls were calculated without considering any driver lags and then a time delay was added, as if the driver would not know about his own response limitations when he/she learned the optimal control, and then both the preview time and the time delay were chosen after the fact, to obtain a best fit between the two sets of results. A 10-preview-point version of the theory with equal weighting on each of the 10 tracking errors was similarly applied to the explanation of a low-lateral-acceleration lane-change manoeuvre which was conducted experimentally, with reasonable agreement between measured and computed results, but again, the same features were involved. Furthermore, the car model used was quite simplified (no roll freedom, for example) compared with the real car and it should not be concluded that the experiments validate the theory. In the low-frequency limit, a manoeuvre becomes a circular turn with constant controls. Then it is obvious that there is no loss of capability associated with the assumption of a constant control over the prediction time. However, the more dynamic a manoeuvre is, the more likely it becomes that the 'optimal' control will vary, depending on the control horizon assumed and the assumption made about the control beyond the control horizon and before the prediction horizon. Testing a steering controller on paths that require only very low frequency inputs may imply a bias in favour of that controller.

For several years, MacAdam's method represented the conventional wisdom regarding effective modelling of car steering control and an updated and extended version of the model has been developed under General Motors sponsorship [135,149]. Strangely, the model retains the feature that the driver delay is applied as an appendage to the controls calculated assuming that it does not exist.

A more elaborate optimal-control approach to car steering, with a view to automating the driver's function, was applied by Peng and Tomizuka [150]. The road lateral curvature and superelevation over a certain preview horizon were assumed to be known perfectly, with the road ahead of the previewed section assumed bounded and of known statistical properties, generated when needed as an exponential decay. A frequency-weighted quadratic cost function, including terms that represent lateral acceleration (ride comfort); lateral tracking error; attitude angle error; a lateral position integral term, mainly to eliminate steady-state errors when only feedback control is being used; and the steer angle input, was shown equivalent to a standard LQR cost function. The optimal control was derived by nonlinear programming, requiring the steady-state solution of a Riccati equation, exactly as with LQR, to give the same state-feedback control as is optimal in the absence of preview. The optimal control had two preview parts, one corresponding to the previewed road section, the other corresponding to the non-previewed road ahead of the previewed section, represented by the exponential decay. These feedforward parts were associated with a combination of the inverse vehicle dynamics and the road-ahead profile. Lateral tracking error and lateral acceleration frequency responses to effective road curvature (accounting for superelevation) input showed the value of preview in reducing low-frequency tracking errors and high-frequency discomfort, the latter presuming sufficient bandwidth in the steering actuator. Simulations of a complex car at 32 m/s speed with first-order-lag dynamics of time constant 0.15 s traversing a track consisting of two straights joined by a circular curve were carried out in discrete time with 0.01 s interval. Markers 1 m apart on the road centre line were taken to provide path data intermittently and specific frequency weighting parameters for lateral acceleration, tracking error and attitude angle error were adopted. Tyre cornering stiffnesses were estimated on-line, so that local values could be employed in the control computations. Road curvature beyond the previewed section was assumed constant. The influence of preview in providing control in anticipation of the road curvature, smoothing the control, reducing the discomfort and improving the tracking were shown, with only small differences between cases with 1 and 0.3 s preview horizons. It should be noted that the control will be truly optimal if the preview information is sufficient, due to the diminishing-returns feature of preview control or if the presumed disturbance beyond what can be observed is correct for the manoeuvre considered. Otherwise, the calculated control will not be optimal and it will not have a known relationship with that which is optimal.

The structure of Casanova's driver steering model [151] was based on knowledge of optimal linear discrete-time preview control deriving from research on active suspensions. It was realised that the full control has state-feedback and preview components, each preview gain being multiplied by a preview error and added to a sum to yield the total preview control contribution. The preview gain sequence falls to zero far enough ahead of the vehicle, reflecting the diminishing-returns nature of preview control, and the sequence reflects the inverse dynamics of the vehicle in some way. If the vehicle is oscillatory, then so will be the gain sequence and the frequencies of the variations will match. This reflects the idea that the control in such a case will operate on the lightly damped mode to obtain the desired response with little expenditure of control effort. If the vehicle is well damped, the preview gain sequence will be smooth and decaying. On this basis, the control was designed to obtain maximum performance from a Formula One car, especially recognising tyre force saturation as a crucial factor. The steering controller included feedback of attitude angle error and eight unevenly spaced preview gains with gains tuned heuristically according to the guiding principles above, and saturation levels on each preview gain and on the total control. The similarity between this structure and that of a neural network controller [152] was noted and many simulations demonstrated excellent car control at near-limit running conditions.

A formal optimal linear discrete-time preview control was derived in [153], independently of [149], with strong similarities but some important differences. At the outset, the problem was seen as a DLQRP one, amenable to numerical solution by standard software if suitable assumptions were made. These embrace linearity of the system, choice of a discrete-time step, a quadratic cost function, infinite preview and a road disturbance consisting in sample values from a white-noise process. If the model of vehicle and road are referenced to a fixed ground, the road dynamics are those of a shift register or delay line and the road and vehicle are related to each other only through the cost function. The cost contains tracking error and steer angle terms. The tracking error had lateral position and attitude angle parts but it is now realised that putting a cost on attitude angle may not be a good idea, since a car with no rear steering system needs to run at an attitude angle to develop the rear axle forces required to follow a curved path. The weights used in the cost function determine the balance between tracking accuracy and control effort that the driver utilises, emphasising that a whole range of optimal controls exists, depending on the control tightness. From this form of the theory, the optimal controls for a given car, speed and control tightness, are revealed as a set of state-feedback gains and an infinite sequence of preview gains but, corresponding to the diminishing-returns nature of preview control, the preview gains converge on zero and become negligible, at which point 'full' preview is presumed. Each control is optimal only in an 'overtaking' sense [154,155], so that a shortened preview-gain sequence, that might be thought optimal for restricted preview circumstances, is not in fact optimal. Also, if the disturbance is obtained by low-pass filtering a white-noise sequence and full preview is available, the same preview control gains are obtained [19], so that a randomly profiled road fits to the theory nicely.

The controls need transforming to a driver's view to allow general manoeuvres to be simulated and it is shown that, if the reference axes are relocated to align with the vehicle with their origin at the vehicle reference point at each time step, the loss of lateral displacement and attitude angle feedback terms, due to the errors becoming zero, is exactly compensated by the changes to the preview errors, again provided that full preview is available [153]. Path-tracking simulations employing both global and local approaches demonstrated the effectiveness of the method and showed the anticipatory and smoothing effect of the preview control if the path demands contain high-frequency components. A car driver was concluded to need between 1 and 1.5 s preview for full performance, the greater at the higher speed, and the optimal preview control sequence for car steering tends to become more 'oscillatory' as the speed increases.

In [156], the focus is on car driving with modest skill levels. Driving is divided into a two-stage activity, the first of which involves working out ideal control inputs, both longitudinal and lateral, the second of which involves trying to realise the ideal controls with PI longitudinal and PID lateral controllers. Either of two car models could be used for stage one, the first, a simple point-mass model corresponding to a driver with little appreciation of the vehicle dynamics, the second, a more realistic longitudinal, lateral, yaw model with combined-slip tyre forces based on [157], corresponding to a more expert driver. Nonlinear Constrained MPC theory was employed for deriving the 'ideal' controls with standard SQP procedures. A number of alternative 'ideal' driving styles were described, with the possibility that they could be normalised and combined into one performance objective as expressed in a single cost function. Maximum distance, minimum lateral acceleration, minimum braking, minimum tracking error, optimal engine utilisation and most even speed maintenance criteria were defined. Some limits could also be imposed as constraints.

Test runs on a special track with position marking provision were compared with predictions, presumably based on the more realistic car model. Eight ordinary drivers were used, with the track restricting the possible manoeuvres. Drivers were instructed to prioritise different aspects of their performance, according to the alternatives described above. There were obvious similarities between the actual paths and the predicted ones but not precise agreements. A standard double lane change at three different speeds was simulated, with predictions based on the simple car model. Unsurprisingly, for the lowest speed at which the manoeuvre could be completed without exceeding 3 m/s^2 lateral acceleration, the car stayed nicely within the road limits, for the medium speed, it overshot the path on the exit a little but for the highest speed, it went well off-track. The results illustrate the difficulties associated with understanding driving of only moderate quality, especially in connection with optimal-control theory.

A steering control model, also aimed at mimicking real driving of various characters, was described in [158]. On-line identification is used to obtain a locally valid linear model of a nonlinear vehicle, enabling the linear model to be employed in an MPC fashion to minimise a lateral-position-and-velocity-based cost function of sliding mode form over a preview horizon. The cost is an extension of MacAdam's cost, which can be reduced to the MacAdam form through choice of parameters. The time constant and the order of the sliding surface are adjustable to represent different driving styles and skill levels. Simulations of a TruckSim version of a Sport Utility Vehicle 'Moose' test gave almost the same behaviour as MacAdam's model when set up with the same objective but, by varying the cost function details, it was possible to alter the quality of the driving.

Definition of the character of the driver through the choice of cost function is also a feature of [159], which brings together driver modelling and minimum-manoeuvre-time problems. Longitudinal and lateral controls are included, limited by acceleration and steer velocity bounds. Direct shooting is employed to find the nonlinear optimal system under equation-of-motion equality constraints, inequality constraints relating to path boundaries and equality constraints relating to initial and final states. Special attention is paid to plants described by differential algebraic equations of the semi-explicit form and to dealing with discontinuous and piecewise-defined dynamics. The basic optimisation is by SQP with Newton's method to guide the iterations, requiring a good approximation to the solution in the beginning to achieve convergence. An interesting feature is the calculation of gradients by integration of the sensitivity equations, derivable from the vehicle-state equations, alongside the state equations themselves. The difficulty caused by long manoeuvres, as also reported in [31], associated with high sensitivities of track-constraint violations a long way ahead to steering inputs, made long-manoeuvre optimisations by single shooting impossible and the solution was found in a receding-horizon strategy, in which only the first control is actually used. The remainder of the computed control history for a given time step was used as an approximation to the optimal solution for the following step, to avoid the computing effort being wasted.

Simulation of a complete short circuit traversal by an elaborate virtual car with a cost function balancing speed and steering effort and using 11 s preview and 1.1 s control steps with linear interpolation of the controls is illustrated. The circuit contains a short narrow section and a hairpin bend. Apparently, the starting state was chosen so that the final state was almost the same. The car completed the manoeuvre but some of the behaviour is not so convincing from a human driver point of view. With lateral accelerations no greater than 7 m/s^2, the slip angles rose to about 0.9 rad. at the hairpin and the controls appear to vary much too rapidly and too much to correspond with a real driver.

The work of [153] was extended in [160]. A low-pass filter was added to the road description to control the bandwidth as desired and numerical trials showed that the optimal preview controls are unaffected by the filter properties, provided that full preview is available, as could be expected from [19]. Correspondingly, in this case, the filter states are not used for feedback. Costs were associated only with tracking error and steering wheel angle and tight and loose controls were illustrated. It was suggested that the preview needed for full benefit, being a measure of how far ahead the driver needs to look, is a handling-quality metric and the sensitivities of the necessary preview to car-design parameters were calculated from a simple linear yaw-sideslip car model. The rear to front steer ratio for an all-wheel steering configuration was similarly studied. Simulations with parameters chosen to represent a rally car showed the anticipated oscillatory steering in preparation for a direction-change, in agreement with common observations.

Ungoren and Peng [161] provide an extension and elaboration of the work reported in [158]. The TruckSim vehicle model gives behaviour that is taken to represent actual vehicle operation, but the model complexity is considered too great to use as the virtual driver's internal model. Therefore, on-line identification of the vehicle dynamics was employed. The linear dynamics were identified over a period of 'rich' excitation and then fixed, making the internal model relate to the linear range of the vehicle. An extended MacAdam-style optimisation, employing control that can switch half way to the prediction horizon and an elaborated cost function, was carried out, and the resulting driver–vehicle system was subject to the 'Moose' test at 23.6 m/s speed. Attitude angle, lateral position and lateral velocity errors with variable weightings were included in the cost and it was argued that higher derivatives could be included if desired. As discussed in [160] however, minimising the attitude angle of a vehicle without rear axle steering is in conflict with accurate path tracking because the rear axle needs to generate forces by side-slipping to follow a curved path, so that it is evident that tracking quality can be prejudiced by putting a high weighting on attitude errors and a low one on lateral position errors. Short, medium and long hyperbolic-tangent-function weights were applied to the preview window, causing the calculated control to be relatively tight in the first case and relatively loose in the last one. As in [160], the tight, high-authority control can be characterised by vigorous control actions, accurate tracking and high risk of control breakdown. In some cases, the vehicle reaches lateral accelerations beyond the scope of the linear model used for the path-tracking task. Varying model parameters indeed do allow the performance of the closed-loop system to change, as intended at the outset.

It was confirmed in [162] that the same optimal controls as are obtained by LQR with preview [153] come from linear unconstrained MPC with the same objectives, when the control horizon is the same as the prediction horizon and they are sufficient for full preview. MPC has the advantage that it will yield an optimal control with a shortened prediction horizon but it has the disadvantages that the controls need to be calculated on-line and they contain an inseparable combination of disturbance and control, making it unlikely that much intuitive appreciation of the optimal controls will be derived.

In [163], a simple yaw-sideslip car with saturating front and rear tyre forces is employed as a basis for a multi-model steering controller. It is argued that real drivers learn and store multiple optimal controls, each with its own local domain of applicability, recalling them as appropriate to cover the whole feasible range of operation. The car model is linearised for small perturbations from states defined by front- and rear-axle lateral slips, with the most general description having a 10-by-10 grid of evenly spaced conditions. Optimal control sets from linear unconstrained MPC as in [162] were computed offline and selected in tracking simulations, going up to $9\,\mathrm{m/s^2}$ lateral acceleration, according to which set corresponds most closely to the current front and rear slip values. The anticipated advantage of multiple controls in comparison with a fixed straight-running control was demonstrated. It appears unlikely that such offline preparations could be made in the case of a general nonlinear vehicle model, since the nonlinear features would be too numerous and impossible to isolate, but linearising about trim states offline or about operating states on-line are feasible ways in which the multi-model notion can be applied generally.

The methods of [153,160] were applied to a high-fidelity motorcycle representation at a given speed with steer torque and rider lean-torque controls [164,165], to a bicycle with only steer torque control [166,167] and to speed control of motorcycles [168] and cars [169] in straight running. A motorcyclist was shown to need much more preview than a car driver for similar speeds by a factor of about six, a feature well-known in rider-training circles, and lean-torque control was shown to add only marginally to the steer-torque control that is fundamental with single-track vehicles. The optimal preview controls had a strong connection with the open-loop dynamics of the machine, such that, when the motorcycle has a lightly damped weave mode for example, the preview gain sequence shows oscillations at the weave frequency. A new feature in the bicycle work was the calculation of the rider-controlled (closed-loop) system frequency responses, demonstrating perfect tracking up to a frequency (around $1\,\mathrm{rad/s}$), increasing with the control tightness. Many accurate path-tracking simulations are included in these papers. MPC, similar to that of [162], was applied to the straight-running motorcycle in [170] with the advantage over the DLQR method that restricted previews can be dealt with effectively but the disadvantage that the controls must be worked out on-line. Complementary results can be found in [171].

Several publications are directed towards autonomous cars, not considering driver limitations like neuro-muscular delays, but still with some relevance to driver modelling. An example is [172], in which MPC with multiple models, here called linear time varying (LTV), is one option. First, a nonlinear MPC is devised for a longitudinal, lateral, yaw car with 'Magic Formula' tyres but apparently with no load transfer. The quadratic cost function contains lateral position and attitude angle errors and steer angle increment terms. The prediction horizon is always greater than the control horizon and the control is presumed constant beyond its horizon. Second, a sub-optimal MPC, employing linearisation of the plant at each time step is created and third, a reduced version of the second controller with only one control step is explored. A sampling interval of 0.05 s is used. The second and third controllers run much faster than the first, this being their main purpose, and the loss of performance for given prediction and control horizons is compensated by the practicality associated with lengthened horizons. Improved results are obtained from the sub-optimal controllers by including soft constraints on axle slip angles in the formulation. Each controller was implemented on a production car with electric servo-motor steering and comparisons between experimental and simulation results are included. Results quoted suggest that the horizons utilised were often rather short. Further work is reported in [173] where the conditions for stability of the closed-loop system receive special attention and controlled manoeuvring of the test car on a snow-covered surface is demonstrated.

Neuro-muscular system actuation properties allied to successive double-lane-change manoeuvres were studied particularly in [174]. Steering wheel angle was the output from a DLQR-based path-following controller, with reflex action and muscle co-contraction providing complementary mechanisms by which the actuation can occur. Minimisation of a cost function is used to generate control inputs similar to those observed in driving-simulator experiments. As the subjects learned the dynamics of the simulator, the extent of co-contraction reduced, indicating the improvement of the subjects' internal models over time. Performance in the simulator was robust to changes in the torque properties of the steering system but not to the displacement properties, reinforcing the notion that people drive cars employing position (fixed) control, being not so sensitive to the feel of the system.

Imperfections were added to such an 'optimal' system in [175], in the form of limited understanding of the vehicle, time delays and again the neuro-muscular system. Controls and lane-change performance were compared for the closed-loop system with and without imperfections.

The focus in [176] is on intermittency of attention as a source of imperfect driving. Road curvature over a preview horizon was considered known and a general predictive control strategy applied to finding a control law, the resulting control being utilised in intermittent fashion. Comparisons with real driving in a lane change were made and a range of intermittence periods in which satisfactory performance occurs was suggested.

Adaptation by gain scheduling was added to the DLQR-based scheme discussed above [177]. Still with reference to a yaw-sideslip car at constant speed, saturating tyre forces by Magic Formula replaced the former linear ones, and trim states were generated by simulation for all feasible levels of lateral acceleration. Optimal feedback and preview controls were generated offline for several levels of front-axle sideslip and stored for use in simulation. The preview controls showed decreasing gains and the need for lengthening preview as the cornering effort rises, corresponding to reducing control authority as the tyre forces are used up for the trim state. When the front-axle forces go beyond their peak value, the preview gains reverse, so that the closed-loop system 'knows' to reduce the steer angle if the car runs beyond saturation. Frequency response computations, for small perturbations from a trim state of course, showed perfect tracking up to around 3 rad/s frequency, with the bandwidth reducing as the cornering effort increases. Tracking simulations on tracks designed to be fully challenging for the 30 m/s speed showed excellent behaviour of the system with adaptation, when the non-adaptive system goes out of control. The effect of control switching was reduced by including a rate limit to the steering wheel. At the limit, the control could hold the front-axle force at its peak value, enabling the car to get back to the track if it ran wide through going too fast.

Similar techniques were applied to a rear-limiting car in [178]. Such a car is prone to yaw instability and, in the real world, would be much more difficult to control than a normal vehicle. It is of interest in connection with competition, since variations in running conditions can cause a racing car to become rear-limiting. Rear-axle sideslip replaced that for the front as scheduling variable, and results were qualitatively similar. The controller ‘understands’ the need for ‘opposite-lock’ steering when the rear axle limits. The closed-loop system bandwidth was particularly prejudiced as the rear tyre forces were used for the relevant trim state.

The approach was further extended in [179,180] to relate to a much more elaborate multi-body car model with variable speed and throttle and steering controls, both longitudinal and lateral tracking requirements and driver-bandwidth limitations built into the problem. Trim states covering wide variations in speed and lateral acceleration were used for control design, but only a small number of results are shown. As might be expected from the lateral symmetry of the straight-running condition, longitudinal and lateral problems decouple for this case, but, when the symmetry is lost by cornering, the two sets of controls become interdependent and the influence of throttle control on the lateral behaviour, and the converse, can be seen in the optimal controls. The driver-bandwidth limitation does not have much influence until he/she becomes slow relative to the car. Then, the closed-loop system capability deteriorates, evidenced partly by frequency–response results. Path-following simulations showed speed and steering control operating together but were restricted in terms of manoeuvre severity in recognition of the need for scheduling over many control sets if the full range of a vehicle is to be exploited. Relating to the same car model, preview controls for an arbitrarily chosen speed of 40 m/s but with a full range of lateral accelerations are shown in [181]. A lane change is simulated to illustrate bilinear gain scheduling over speed and lateral acceleration.

Within the same context as above, the special problems of severe braking were considered in [182]. The difficulty arises because rapid deceleration and the constant speed of the trim state are incompatible. If a true trim state is used as the basis for small perturbation controls, the real perturbations may be too large on sudden and severe braking for the controls to be of good quality. As a palliative, an artificial horizontal gravity was included in a multibody car model developed from that described above, allowing pseudo-trim states with severe braking to be simulated. These states can be used in the normal way to generate optimal controls corresponding to the car being pitched nose-down, most of the weight supported by the front tyres, etc. and gain scheduling can then be employed to ensure that locally valid controls are installed. Simulations of straight-running, near-limit braking of the car were run, first with linear interpolation over trim states and control sets according to car speed and second, with bilinear interpolation over speed and deceleration. In this latter case, computation of the preview controls requires careful treatment to represent the vehicle path well, under the trim controls only. Excellent use of the available tyre forces is demonstrated, when the bilinear scheduling is employed.

This basic scheme has also been shown to work well when applied to the difficult stabilisation and path-following control of a unicycle [183], with its rider having full upper-body rotational freedoms and band-limited control of the muscles governing relative yaw, pitch and roll and pedalling. Longitudinal and lateral frequency responses are shown for different tightness levels, showing perfect tracking for sufficiently low frequencies, and a lane-change type of manoeuvre with a fixed set of gains designed for straight running is simulated. A full set of preview gains for a turning trim state is included and following a clothoid path at 2 m/s speed with gain-scheduled switching according to the unicycle roll angle going up to 7 degrees is illustrated.

The work of [184] is directed towards an understanding of drivers’ internal plant models and of the implications of their being imperfect. It is related in technical terms to the earlier research of the same authors [163]. A yaw-sideslip-roll car model, simplified with respect to issues like lateral load transfer so that it has saturating axle forces that are just functions of sideslip, is constructed from experimental steering-response data. Axle forces are linearised for small perturbations from given slip values and a grid of linear model results for combinations of front- and rear-axle slips that cover the feasible operating range of the car. The linear models are prepared offline and cover a 25-by-25 grid at their most elaborate. Controllers based on the linear models can be devised using the MPC methods established earlier, such that a transient manoeuvre, namely, the standardised 'Elk' test, can be carried out on the basis that the driver knows all of the linear models, and can select that most appropriate at intervals during the test or he/she can employ only one controller, or can be anywhere in between the extremes. Thus, different levels of appreciation of the vehicle dynamics can be built into the driving function. Conventionally with MPC, at a computation step the control for the whole control horizon is calculated, the first control is installed for the next step and the remainder of the control is discarded. Preview computations for the next step involve comparing the predicted vehicle path over the preview horizon, assuming that the base control will be the current one, against the ideal path. The computed control gives changes from the base control to optimise the motions over the prediction horizon. A novel feature introduced into these computations is that the base control uses all the information from the previous-step calculations, making the differences between the predicted path and the ideal path smaller and enhancing the validity of the small-perturbation assumptions involved in the linearisation process. The authors call their model levels 'variable-model-preview', 'single-model-preview' and 'fixed-model-preview', corresponding crudely to expert, average and novice drivers. Frequency responses for each of four linearisations around trim states representing zero and three non-zero levels of lateral acceleration are shown, with the expected feature, for fixed weights in the MPC calculations, that the closed-loop system bandwidth decreases as the tyre forces are needed by the trim state and the control authority therefore diminishes. Elk test simulation results show the relative merits of the control strategies aligned with expectations, with lateral accelerations in the range to $7\,\mathrm{m/s^2}$ in the best-controlled case. The controller choices made in the variable-model-preview case are also shown, with a concentration of the chosen controllers not too far from those that belong to trim conditions.

A parametric study centred on traversing 2 km of random roadway at 18.06 m/s speed. The density of the driver's internal model array, the control tightness employed by the driver and the driver's knowledge range, in the form of the range of tyre slip angles over which he/she has learned the vehicle dynamics, were involved. With respect to the internal model density, diminishing returns, measured primarily by smoothness of motions, were obtained beyond a 10-by-10 model grid. Tighter controls improved the tracking quality and reduced the amount of preview needed for full benefit and increasing the knowledge range improved the tracking quality right up to a 20-by-20 grid of linear models, somewhat in conflict with the model density finding.

The latest work on the motorcycle [185] includes a detailed multibody machine model, throttle displacement, steer torque and rider lean torque as controls, generates optimal controls offline by the DLQR methods described earlier for a full range of trim states covering speed and lateral acceleration ranges, and uses bilinear interpolation in a gain-scheduling scheme for tracking simulations. Rider controls are not restricted in terms of frequency. Samples of optimal feedback and preview controls are shown and some frequency–response characteristics are illustrated. Good control of the motorcycle in manoeuvres that require rapid speed change and roll angles up to 50° is demonstrated, with some short bursts of throttle displacement and steering torque that are at too high a frequency to be realistic. The steer torque bursts are closely related to the oscillatory properties of the machine at high speed and would be improved by changing the motorcycle, by restricting the rider-control frequencies or by employing looser controls.

3.4.3. *Summary*

MPC, with its strong focus on using preview information which is central to effective driving or riding, has become a main resource for the modelling of drivers and riders. All kinds of vehicles have been shown controllable under wide-ranging conditions. Adaptive control by gain scheduling, LTV or multi-model methods, constituting different names for the same thing, are necessary for the widest variations in running conditions. In general terms, DLQR methods allow offline determination of control schemes while MPC methods require on-line calculations, although there is some overlap between the methods. In the on-line case and especially for real-time applications there is a premium on computational load and compromises may be necessary to balance quality against speed. Linearisation appears to be best in terms of gaining speed without losing too much quality but there is likely to be some dependency on the particular application. In the offline case, the enemy is proliferation of cases for which linear models need to be prepared. Not knowing in advance where the vehicle will go obliges the analyst to cover all eventualities and much of the preparatory work may be wasted. The preparation load can get out of hand if several scheduling variables and a fine grid of conditions need to be covered. The fineness of the grid required is related to the quality of the interpolation between grid points that can be employed. MPC is most naturally aligned with perfect driving but average or poor driving can be associated with variations of the cost function, slow driver activity, inattentiveness, poor appreciation of the vehicle dynamics, not enough path preview, etc. These features have been included in models but it remains difficult to make the necessary connections with real human actions. Research so far completed has contributed greatly to improving knowledge and understanding of driving and riding of both high and moderate quality but many issues are still quite open. Even the relatively simple issues of vehicle model complexity, manoeuvre choice, prediction and control horizons, cost function terms and weights and sampling interval leave considerable scope for the generation of new results and understanding.

4. Conclusions

Applications of optimal-control theory to automotive problems have been wide ranging and informative. The focus here has necessarily been on a few specific areas, namely, suspension systems, worst-case manoeuvring, minimum-time manoeuvring and vehicle driving. In connection with suspensions, the much-used linear, quadratic-cost, randomly excited quarter-car model has revealed a great deal of what is possible and how to gain advantage. More elaborate models, while often being more realistic, have tended to add to the understanding much more sparingly. However, some features that are commonly built into the abstracted optimal-control problem preclude certain characteristics in the resulting system designs, which characteristics may be almost universal in real systems. Notable in this context are the up-down asymmetry that is usual in automotive suspension dampers and the hardening-spring property caused by the inclusion of elastic limit stops in real systems. There is currently no theory to explain the asymmetry and no conventional wisdom relating to ideal amplitude-dependence of force–displacement or force–velocity relationships. Of course, it may be that these ideals depend on specific aspects of an application and must be decided by value-judgement and compromise. If there are more things to be understood by the application of optimal-control theory, it seems likely that it will be necessary to widen the terms of reference of the theory relative to what has gone before, to open up the possibility that new features, previously precluded, will appear in the solutions.

Vehicle behavioural properties which encourage loss of control by the driver in unusual circumstances are dangerous and they need to be avoided. Traditionally, the evolutionary process provided some insurance against the occurrence of such behaviour but as vehicles change, methods for finding worst-case control inputs and running conditions need to change in sympathy. Using proving ground facilities and normal test procedures makes the determination of worst-case conditions almost impossible. The application of optimal-control theory in a virtual world has been shown to make the necessary process feasible and practical. The computations should become routine for new vehicles that have any chance of exhibiting behavioural problems, before designs are finalised.

The theory of minimum-time manoeuvring of realistic vehicles over realistic tracks implies computational complexity. It is necessary to consider how to solve the problem and how to achieve a solution quickly on ordinary computers together, for the results to be really useful. The more representative of the real world such a study is made, the more likely it seems to be that the best solutions will be by nonlinear programming. There are several software packages which provide a framework for solving nonlinear-programming problems and it is likely that using one of these packages, probably incorporating SQP will be preferred. A large cost is associated with calculating gradients, the default method for doing this being by finite differences, but AD is much more efficient and the software tools for enabling AD to be employed by non-specialists are improving in-line with developments of computer memory. It appears essential in the future to use AD to achieve a suitable combination of accuracy and speed of solution. Another source of efficiency is to avoid having the computer discover by calculation things that are already known to the analyst. Among these factors may be restricting the solution-search space according to vehicle stability and controllability criteria and applying track-boundary constraints only where they are at all likely to become active. Parallel computation may contribute strongly in the future. The ultimate solution may involve the decoupling of the racing-line calculation from the optimal-speed calculation, with a sequence of much simpler computations replacing the conventional complex parallel one. A high-quality driver model would be required to provide a path- and speed-tracking capability in this case.

MPC in a few different forms has been applied to driving and riding, as described in a rapidly expanding literature. In any application, several technical choices need to be made, relating to plant model, manoeuvres, constraints, objectives, prediction and control horizons and sampling interval. Many studies are therefore necessary to obtain anything approaching a complete picture. Those reported so far give a good framework on which to build. Vigorous manoeuvres of cars and motorcycles have been simulated under good control and there is no reason to think that nonlinear MPC will not replicate the high skills of real drivers and riders. However, nonlinear MPC is too heavy computationally for some purposes and the necessary and effective compromises, which will allow good control and fast computation together, have not been fully developed. Several suggestions relating to vehicle control of lower quality have been made. Demonstrating connections between quality-reducing features and human characteristics, priorities and judgements is difficult and relatively immature.

Acknowledgements

The authors are pleased to acknowledge the contributions to the paper from Michael Valasek, relating to the SDRE technique and to active and semi-active suspensions, and to Tim Gordon for contributions relating to indirect optimisation.

References

[1] M.A. Athans and P.L. Falb, *Optimal Control*, McGraw-Hill, New York, 1966.
[2] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.
[3] A.E. Bryson and Y.C. Ho, *Applied Optimal Control*, 2nd ed., Hemisphere Publishing Corporation, Washington DC, 1975.
[4] B.D.O. Anderson and J.B. Moore, *Optimal Control: Linear Quadratic Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[5] D.E. Kirk, *Optimal Control Theory: An Introduction*, Dover Publications, Mineola, NY, 2004.
[6] A.G. Thompson and B.R. Davis, *Optimal active suspension design using a frequency-shaping PID filter*, Veh. Syst. Dyn. 21 (1992), pp. 19–37.
[7] M. Tomizuka and J.K. Hedrick, *Advanced control methods for automotive applications*, Veh. Syst. Dyn. 24 (1995), pp. 449–468.
[8] A. Gelb, *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.
[9] O.L.R. Jacobs, *Introduction to Control Theory*, 2nd ed., Oxford University Press, Oxford, UK, 1993.
[10] A.G. Thompson, B.R. Davis, and F.J.M. Salzborn, *Active suspensions with vibration absorbers and optimal output feedback control*, SAE 841253, SAE International, Warrendale, PA, (1984).
[11] A. Hac, *Suspension optimisation of a 2-dof vehicle model using a stochastic optimal control technique*, J. Sound Vibration 100 (1985), pp. 343–357.
[12] D.A. Wilson, R.S. Sharp, and S.A. Hassan, *The application of linear optimal control theory to the design of active automotive suspensions*, Veh. Syst. Dyn. 15 (1986), pp. 105–118.
[13] R.S. Sharp and S.A. Hassan, *On the performance capabilities of active automobile suspension systems of limited bandwidth*, Veh. Syst. Dyn. 16 (1987), pp. 213–225.
[14] R.S. Sharp and J.H. Hassan, *Performance predictions for a pneumatic active car suspension system*, Proc. IMechE. D, J. Automob. Eng. 202 (1988), pp. 243–250.
[15] A. Hac, *Optimal linear preview control of active vehicle suspension*, Veh. Syst. Dyn. 21 (1992), pp. 167–195.
[16] R.S. Sharp and S.A. Hassan, *The relative performance capabilities of passive, active and semi-active car suspension systems*, Proc. IMechE. D, J. Automob. Eng. 200 (1986), pp. 219–228.
[17] H. Du and N. Zhang, *Designing H1/GH2 static-output feedback controller for vehicle suspensions using linear matrix inequalities and genetic algorithms*, Veh. Syst. Dyn. 46 (2008), pp. 385–412.
[18] E.K. Bender, *Optimum linear preview control with application to vehicle suspension*, Trans. ASME, J. Basic Eng. 90 (1968), pp. 213–221.
[19] M. Tomizuka and D.E. Whitney, *Optimal discrete finite preview problems (why and how is future information important?)*, Trans. ASME, J. Dyn. Syst. Meas. Control 97 (1975), pp. 319–325.
[20] M. Tomizuka, *Optimal linear preview control with application to vehicle suspension – revisited*, Trans. ASME, J. Dyn. Syst. Meas. Control 98 (1976), pp. 309–315.
[21] A.G. Thompson, B.R. Davis, and C.E.M. Pearce, *An optimal linear active suspension with finite road preview*, SAE 800520, SAE International, Warrendale, PA, (1980).
[22] N. Louam, D.A. Wilson, and R.S. Sharp, *Optimisation and performance enhancement of limited bandwidth active suspensions for automobiles under preview of the road*, Veh. Syst. Dyn. 21 (1992), pp. 39–63.
[23] N. Louam, D.A. Wilson, and R.S. Sharp, Optimal control of a vehicle suspension incorporating the time delay between front and rear wheel inputs, Veh. Syst. Dyn. 17 (1988), pp. 317–336.
[24] G. Prokop and R.S. Sharp, *Performance enhancement of limited bandwidth active automotive suspensions by road preview*, IEE Proc., Control Theory Appl. 142 (1995), pp. 140–148.
[25] T.J. Gordon, C. Marsh, and M.G. Milsted, *Control law design for active and semi-active automobile suspension system*, VDI–Ber. 816 (1990), pp. 537–546.
[26] T. Cimen, *State-dependent Riccati equation control: A survey*, Proceedings of the 17th World Congress IFAC, July, Seoul, 2008, pp. 3761–3775.
[27] P.E. Gill, W. Murray, and M.H. Wright, *Practical Optimization*, Academic Press, London, 1981.
[28] R. Fletcher, *Practical Methods of Optimization*, Wiley-Interscience, Chichester, UK, 1987.
[29] A. Antoniou and W.S. Lu Practical, *Optimization – Algorithms and Engineering Applications*, Springer, New York, 2007.
[30] A.P. Sage, *Optimum Systems Control*, Prentice-Hall, New Jersey, 1968.
[31] D. Casanova, *On minimum-time vehicle manoeuvring: The theoretical optimal lap*, Doctoral thesis, Cranfield University, 2000.
[32] J.M. Maciejowski, *Predictive Control with Constraints*, Pearson Prentice Hall, Harlow, England, 2002.
[33] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge University Press, Cambridge, UK, 2007.
[34] M.C. Best and T.J. Gordon, *Simultaneous Optimisation of Vehicle Parameter and Control Action to Examine the Validity of Handling Control Assumptions*, Proceedings of the 6th International Symposium on Automotive Control (AVEC 2002), September 9–13, SAE of Japan, Hiroshima, 2002, p. 6.
[35] E. Bertolazzi, F. Biral, and M. da Lio, *Symbolic-numeric indirect method for solving optimal control problems for large multibody systems: The time-optimal racing vehicle example*, Multibody Syst. Dyn. 13 (2005), pp. 233–252.
[36] J.B. Rawlings, *Tutorial overview of model predictive control*, IEEE Trans. Control Syst. Mag. (2000), pp. 38–52.

[37] R. Findeisen and F. Allgöwer, *An introduction to nonlinear model predictive control, in Summerschool on 'The Impact of Optimization in Control'*, C.W. Scherer and J.M. Schumacher, eds., Dutch Institute of Systems and Control, DISC, Netherlands, 2001, pp. 3.1–3.45.
[38] R. Findeisen, L. Imsland, F. Allgöwer, and B.A. Foss, *State and output feedback nonlinear model predictive control: An overview*, Eur. J. Control 9 (2003), pp. 179–195.
[39] S.J. Qin and T.A. Badgwell, *A survey of industrial model predictive control technology*, Control Eng. Pract. 11 (2003), pp. 733–764.
[40] M. Morari and E. Zafiriou, *Robust Process Control*, Prentice-Hall, Upper Saddle River, NJ, 1989, ISBN 0-13-782153-0.
[41] M. Green and D.J.N. Limebeer, *Linear Robust Control*, Prentice-Hall, Englewood Cliffs, NJ, 1995, ISBN 0-13-102278-4.
[42] K. Zhou and J.C. Doyle, *Essentials of Robust Control*, Prentice-Hall, Upper Saddle River, NJ, 1999, ISBN 0-13-525833-2.
[43] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, Wiley, New York, 2005, ISBN 0-470-01167-X.
[44] A. Akbari and B. Lohmann, *Output feedback H1/GH2 preview control of active vehicle suspensions: A comparison study of LQG preview*, Veh. Syst. Dyn. 48 (2010), pp. 1475–1494.
[45] J.K. Hedrick and D.N. Wormley, *Active suspension for ground transport vehicles – a state of the art review*, *Proceedings of the ASME Symposium on Mechanics of Transportation Suspension Systems*, AMD Vol. 15, Houston, TX, 1975, pp. 21–39.
[46] K.N. Morman and F. Giannopoulos, *Recent advances in the analytical and computational aspects of modelling active and passive vehicle suspensions*, *Proceedings of the ASME Symposium on Computational Methods in Ground Transportation* Vol. 50, Phoenix, AZ, 1982, pp. 75–115.
[47] R.S. Sharp and D.A. Crolla, *Road vehicle suspension system design – a review*, Veh. Syst. Dyn. 16 (1987), pp. 167–192.
[48] D. Hrovat, *Applications of optimal control to advanced automotive suspension design*, Trans. ASME, J. Dyn. Syst. Meas. Control 115 (1993), pp. 328–342.
[49] D. Hrovat, *Survey of advanced suspension developments and related optimal control applications*, Automatica 33 (1997), pp. 1781–1817.
[50] E.M. Elbeheiry, D.C. Karnopp, M.E. Elaraby, and A.M. Abdelraaouf, *Advanced ground vehicle suspension systems a classified bibliography*, Veh. Syst. Dyn. 24 (1995), pp. 231–258.
[51] M. Valasek and W. Kortuem, *Semi-active suspension systems 11*, in *Mechanical Systems Design Handbook*, Y. Hurmuzlu and O. Nwokah, eds., CRC Press, Boca Raton, 2002, pp. 230–247.
[52] C.J. Dodds and J.D. Robson, *The description of road surface roughness*, J. Sound Vibration 31 (1973), pp. 175–183.
[53] D. Cebon and D.E. Newland, *The artificial generation of road surface topography by the inverse FFT method*, in *Proceedings of the 8th IAVSD Symposium on Dynamics of Vehicles on Roads and on Railway Tracks*, J.K. Hedrick, ed., Swets and Zeitlinger, Lisse, 1984, pp. 29–42.
[54] M.W. Sayers, *Characteristic power spectral density functions for vertical and roll components of road roughness*, *Proceedings of the ASME Symposium on Simulation and Control of Ground Vehicles and Transportation Systems*, Anaheim, CA, 1986, pp. 113–139.
[55] D. Ammon, *Problems in road surface modelling*, in *Proceedings of the 12th IAVSD Symposium on Dynamics of Vehicles on Roads and on Tracks*, G. Sauvage, ed. (Supplement to Vehicle System Dynamics) Vol. 20, Swets and Zeitlinger, Lisse, 1992, pp. 28–41.
[56] G. Ruf, *The calculation of the vibrations of a four-wheeled vehicle, induced by random road roughness of the left and right track*, Veh. Syst. Dyn. 7 (1978), pp. 1–23.
[57] G. Rill, *The influence of correlated random road excitation on vehicle vibration*, in *Proceedings of the 8th IAVSD Symposium on Dynamics of Vehicles on Roads and on Railway Tracks*, J.K. Hedrick, ed., Swets and Zeitlinger, Lisse, 1984, pp. 449–459.
[58] R.S. Sharp, *Decoupling in automotive active suspension system design*, in *Proceedings of the IUTAM Symposium on Interaction Between Dynamics and Control in Advanced Mechanical Systems*, D.H. van Campen, ed., Kluwer, Dordrecht, 1997, pp. 363–374.
[59] S.W.R. Duym, *Simulation tools, modeling and identification for an automotive shock absorber in the context of vehicle dynamics*, Veh. Syst. Dyn. 33 (2000), pp. 261–285.
[60] H.B. Pacejka, *Tyre and Vehicle Dynamics*, Butterworth Heinemann, Oxford, UK, 2002, ISBN 0-7506-5141-5.
[61] A.G. Thompson, *Optimum damping in a randomly excited non-linear suspension*, Proc. IMechE. Part 2A 184 (1969), pp. 169–178.
[62] R.S. Sharp and C. Pilbeam, *Achievability and value of passive suspension designs for minimum pitch response*, in *Vehicle Ride and Handling*, MEP, Bury St Edmunds, 1993, pp. 243–259, ISBN 0 85298 861 3.
[63] D.A. Crolla and R.P. King, *Olley's 'Flat Ride' revisited*, in *Proceedings of the 16th IAVSD Symposium on Dynamics of Vehicles on Roads and on Tracks*, R. Fröhling, ed. (Supplement to Vehicle System Dynamics) Vol. 33, Swets and Zeitlinger, Lisse, 1999, pp. 762–774.
[64] R.S. Sharp, *Wheelbase filtering and automobile suspension tuning for minimising motions in pitch*, Proc. IMechE. D, J. Automob. Eng. 216 (2002), pp. 933–946.
[65] A. Alleyne and R. Liu, *On the limitations of force tracking control for hydraulic servosystems*, Trans. ASME, J. Dyn. Syst. Meas. Control 121 (1999), pp. 184–190.

[66] Y. Zhang and A. Alleyne, *A practical and effective approach to active suspension control*, Veh. Syst. Dyn. 43 (2005), pp. 305–330.
[67] S. Chantranuwathana and H. Peng, *Adaptive robust control for active suspensions, Proceedings of the 1999 American Control Conference*, San Diego, CA, 1999.
[68] S. Chantranuwathana and H. Peng, *Adaptive robust force control for vehicle active suspensions*, Int. J. Adapt. Control Signal Process. 18 (2004), pp. 83–102.
[69] X. Shen and H. Peng, *Analysis of active suspension systems with hydraulic actuators*, in *Proceedings of the IAVSD Symposium on the Dynamics of Vehicles on Roads and on Tracks*, Kanagawa, Japan, M. Abe, ed. (Supplement to Vehicle System Dynamics) Vol. 41, Taylor & Francis, London, 2004, pp. 143–152.
[70] A.G. Thompson and P.M. Chaplin, *Force control in electrohydraulic active suspensions*, Veh. Syst. Dyn. 25 (1996), pp. 185–202.
[71] E.K. Bender, D.C. Karnopp, and I.L. Paul, *On the optimization of vehicle suspensions using random process theory*, ASME 67-Tran-12 (1967).
[72] C. Yue, T. Butsuen, and J.K. Hedrick, *Alternative control laws for automotive active suspensions*, Trans. ASME, J. Dyn. Syst. Meas. Control 111 (1989), pp. 286–290.
[73] Mechanical vibration and shock – evaluation of human exposure to whole-body vibration, Part 1: general requirements, International Organisation for Standardisation ISO 2631-1, 2nd ed. 1997-05-01, corrected and reprinted 1997-07-15 ed. (1997).
[74] R.M. Chalasani, *Ride performance potential of active suspension systems – Part 1: Simplified analysis based on a quarter-car model*, in *Proceedings of the ASME Symposium on Simulation and Control of Ground Vehicles and Transportation Systems*, AMD Vol. 80, 1986, pp. 187–204.
[75] A.G. Thompson, *An active suspension with optimal linear state feedback*, Veh. Syst. Dyn. 5 (1976), pp. 187–203.
[76] R.M. Chalasani, *Ride performance potential of active suspension systems – Part 11: Complex analysis based on a full-car model*, *Proceedings of the ASME Symposium on Simulation and Control of Ground Vehicles and Transportation Systems*, AMD Vol. 80, Los Angeles, CA, 1986, pp. 205–226.
[77] B.R. Davis and A.G. Thompson, *Optimal linear active suspensions with integral constraint*, Veh. Syst. Dyn. 17 (1988), pp. 357–366.
[78] A.G. Thompson, *Optimal and sub-optimal linear active suspensions for road vehicles*, Veh. Syst. Dyn. 13 (1984), pp. 61–72.
[79] R.S. Sharp and S.A. Hassan, *Performance and design considerations for dissipative semi-active suspension systems for automobiles*, Proc. IMechE. D, J. Automob. Eng. 201 (1987), pp. 149–153.
[80] A. Hac, *Adaptive control of vehicle suspension*, Veh. Syst. Dyn. 16 (1987), pp. 57–74.
[81] T.J. Gordon, C. Marsh, and M.G. Milsted, *A comparison of adaptive LQG and nonlinear controllers for vehicle suspension systems*, Veh. Syst. Dyn. 20 (1991), pp. 321–340.
[82] F. Yu and D.A. Crolla, *Adaptive strategies for active vehicle suspension control*, in *The Active Control of Vibration*, C.R. Burrows and P.S. Keogh, eds., Mechanical Engineering Publications, London, 1994, pp. 161–168.
[83] C. Pilbeam and R.S. Sharp, *Performance enhancement by control law adaptation in slow-active road vehicle suspensions*, in *Proceedings of the 14th IAVSD Symposium on the Dynamics of Vehicles on Roads and on Tracks*, L. Segel, ed. (Supplement to Vehicle System Dynamics) Vol. 25, Swets and Zeitlinger, Lisse, 1996, pp. 533–546.
[84] R.A.Williams, *Automotive active suspensions Part 1: Basic principles*, Proc. IMechE. D, J. Automob. Eng. 211 (1997), pp. 415–426.
[85] Y. Yokoya, R. Kizu, H. Kawaguchi, K. Ohashi, and H. Ohno, *Integrated control system between active control suspension and four wheel steering for the 1989 CELICA*, SAE 901748, SAE International, Warrendale, PA, (1990).
[86] Y. Aoyama, K. Kawabata, S. Hasegawa, Y. Kaobari, and M. Sato, *Development of full active suspension by Nissan*, SAE 901747, SAE International, Warrendale, PA, (1990).
[87] R.A. Williams, A. Best, and I.L. Crawford, *Refined low frequency active suspension*, in *Vehicle Ride and Handling*, MEP, Bury St Edmunds, 1993, pp. 285–300, ISBN 0 85298 861 3.
[88] C. Pilbeam and R.S. Sharp, *Performance potential and power consumption of slow active suspension systems with preview*, Veh. Syst. Dyn. 25 (1996), pp. 169–183.
[89] R.A. Williams and S.A. Miller, *Power consumption in automotive active suspensions*, *Proceedings of the International Conference on Systems Engineering*, Coventry, 1994.
[90] R.A. Williams, *Automotive active suspensions Part 2: Practical considerations*, Proc. IMechE. D, J. Automob. Eng. 211 (1997), pp. 427–444.
[91] H.E. Tseng and J.K. Hedrick, *Semi-active control laws: Optimal and sub-optimal*, Veh. Syst. Dyn. 23 (1994), pp. 545–569.
[92] N. Giorgetti, A. Bemporad, H.E. Tseng, and D. Hrovat, *Hybrid model predictive control application towards optimal semi-active suspension*, Int. J. Control 79 (2006), pp. 521–533.
[93] G. Prokop and R.S. Sharp, *On the properties of a preview controlled discrete time system with use of a state observer-application to a limited-bandwidth vehicle suspension*, *Proceedings of the IUTAM Symposium on the Active Control of Vibration*, University of Bath, 1994, pp. 141–152.
[94] R.S. Sharp and C. Pilbeam, *On the ride comfort benefits available from road preview with slow-active car suspensions*, *Proceedings of the 13th IAVSD Symposium on Dynamics of Vehicles on Roads and on Railway*

Tracks, Z.Y. Shen, ed. (Supplement to Vehicle System Dynamics) Vol. 23, Swets and Zeitlinger, Lisse, 1994, pp. 437–448.
[95] T.J. Gordon and R.S. Sharp, *On improving the performance of automotive semi-active suspensions through road preview*, J. Sound Vibration 217 (1998), pp. 163–182.
[96] T.J. Gordon, C. Marsh, and Q.H. Wu, *Stochastic optimal control of active vehicle suspensions using learning automata*, Proc. IMechE. I, J. Syst. Control Eng. 207 (1993), pp. 143–152.
[97] M. Valasek, M. Novak, Z. Sika, and O. Vaculin, *Extended Groundhook New Concept of Semiactive Control of Truck's Suspension*, Veh. Syst. Dyn. 27 (1997), pp. 289–303.
[98] M. Valasek and J. Kejval, *New direct synthesis of nonlinear optimal control of semi-active suspensions*, Proceedings of the 5th International Symposium on Automotive Control (AVEC 2000), Ann Arbor MI, 22–24 August, SAE of Japan, 2000, pp. 691–697.
[99] R. Nader, *Unsafe at Any Speed: The Designed-in Dangers of the American Automobile*, Grossman, New York, 1970.
[100] H. Dugoff, R.D. Ervin, and L. Segel, *Vehicle handling test procedures, PF-100-A*, Highway Safety Research Institute, University of Michigan, 1970.
[101] *The New Car Assessment Program: Suggested Approaches for Future Program Enhancement*, National Highway Traffic Safety Administration, Department of Transportation, Washington DC, HS 810 698 (2007).
[102] W.H. Ma and H. Peng, *A worst-case evaluation method for dynamic systems*, Trans. ASME, J. Dyn. Syst. Meas. Control 121 (1999), pp. 191–199.
[103] A.Y. Ungoren, *Worst-case evaluation methods for rollover prevention system*, Doctoral diss., University of Michigan, 2003.
[104] S. Jayasuriya, *On the determination of the worst allowable persistent bounded disturbance for a system with constraints*, Trans. ASME, J. Dyn. Syst. Meas. Control 117 (1995), pp. 126–134.
[105] T.T. Georgiou and I.J. Fialho, *Worst-case analysis of nonlinear systems*, IEEE Trans. Autom. Control 44 (1999), pp. 1180–1196.
[106] W.H. Ma and H. Peng, *Worst-case vehicle evaluation methodology – examples on truck rollover/jackknifing and active yaw control systems*, Veh. Syst. Dyn. 32 (1999), pp. 389–408.
[107] H. Peng and W.H. Ma, *Solutions of continuous-time preview two-player differential games*, Trans. ASME, J. Dyn. Syst. Meas. Control 121 (1999), pp. 326–330.
[108] Y. Kou, H. Peng, and D. Jung, *Worst-case evaluation for integrated chassis control systems*, in *Proceedings of the IAVSD Symposium on the Dynamics of Vehicles on Roads and on Tracks*, J.K. Hedrick, ed. (Supplement to Vehicle System Dynamics) Vol. 46, 2008, pp. 329–340.
[109] A.Y. Ungoren and H. Peng, *Evaluation of vehicle dynamic control for rollover prevention*, Int. J. Automot. Technol. 5 (2004), pp. 115–122.
[110] Y. Kou, H. Peng, and D. Jung, *Development and evaluation of an integrated cassis control system*, JSAE Rev. Automot. Eng. 29 (2008), pp. 425–438.
[111] T. Fujioka and T. Kimura, *Numerical simulation of minimum time cornering behaviour*, JSAE Rev. 13 (1992), pp. 44–51.
[112] H. Hatwal and E.C. Mikulcik, *Some inverse solutions to an automobile path-tracking problem with input control of steering and brakes*, Veh. Syst. Dyn. 15 (1986), pp. 61–71.
[113] A.K. Wu and A. Miele, *Sequential conjugate-gradient restoration algorithm for optimal control problems with non-differential constraints and general boundary conditions, part 1*, Optim. Control Appl. Methods 1 (1980), pp. 69–88.
[114] A.K. Wu and A. Miele, *Sequential conjugate-gradient restoration algorithm for optimal control problems with non-differential constraints and general boundary conditions, part 2*, Optim. Control Appl. Methods 1 (1980), pp. 119–130.
[115] J.P.M. Hendrikx, T.J.J. Meijlink, and R.F.C. Kriens, *Application of optimal control theory to inverse simulation of car handling*, Veh. Syst. Dyn. 26 (1996), pp. 449–462.
[116] D. Kraft and R. Bulirsch, *Computational Optimal Control*, International Series of Numerical Mathematics Vol. 115, Birkhauser, Basel, 1994.
[117] J. Allen, *Computer optimisation of cornering line*, Master's thesis, Cranfield University, 1997.
[118] V. Cossalter, M. da Lio, R. Lot, and L. Fabbri, *A general method for the evaluation of vehicle manoeuvrability with special emphasis on motorcycles*, Veh. Syst. Dyn. 31 (1999), pp. 113–135.
[119] E. Bertolazzi, F. Biral, and M. da Lio, *Real-time motion planning for multibody systems: Real life application examples*, Multibody Syst. Dyn. 17 (2007), pp. 119–139.
[120] P.J. Enright and B.A. Conway, *Simultaneous optimisation of vehicle parameter and control action to examine the validity of handling control assumptions*, *Proceedings of the AIAA/AHS Astrodynamics Conference, Part 2*, August, Portland, 1990.
[121] T.J. Betts and W.P. Huffman, *Trajectory optimisation on a parallel processor*, J. Guid. Control Dyn. 14 (1991), pp. 431–439.
[122] D. Casanova, R.S. Sharp, M. Final, B. Christianson, and P. Symonds, *Application of automatic differentiation to race car performance optimisation*, in *Proceedings of Automatic Differentiation 2000: From Simulation to Optimization*, G. Corliss, C. Faure, A. Griewank, L. Hascot, and U. Naumann, eds., Springer-Verlag, New York, 2000, pp. 113–121.
[123] G. Corliss, C. Faure, A. Griewank, L. Hascot, and U. Naumann (eds.), *Automatic Differentiation of Algorithms: From Simulation to Optimization*, Springer-Verlag, New York, 2002, ISBN 0-387-95305-1.

[124] P.E. Gill, W. Murray, and M.A. Saunders, *SNOPT: An SQP algorithm for large-scale constrained optimisation*, Numer. Anal. Rep. 97-2, Department of Mathematics, University of California, San Diego, La Jolla, CA, 1997.
[125] D. Casanova, R.S. Sharp, and P. Symonds, *Minimum time manoeuvring: The significance of yaw inertia*, Veh. Syst. Dyn. 34 (2000), pp. 77–115.
[126] D. Casanova, R.S. Sharp, and P. Symonds, *On minimum time optimisation of formula one cars: The influence of vehicle mass*, Proceedings of the 5th International Symposium on Automotive Control (AVEC 2000), August 22–24, SAE of Japan, Ann Arbor, MI, 2000, pp. 585–592.
[127] D. Casanova, R.S. Sharp, and P. Symonds, Sensitivity to mass variations of the fastest possible lap of a Formula One car, Veh. Syst. Dyn. 35 (2001), pp. 119–134.
[128] D. Casanova, R.S. Sharp, and P. Symonds, *On the optimisation of the longitudinal location of the mass centre of a Formula One car for two circuits*, Proceedings of the 6th International Symposium on Automotive Control (AVEC 2002), September 9–13, SAE of Japan, Hiroshima, 2002, p. 6.
[129] D.P. Kelly, *Lap Time Simulation with Transient Vehicle and Tyre Dynamics*, Doctoral thesis, Cranfield University, 2008.
[130] D.P. Kelly and R.S. Sharp, *Time-optimal control of the race car: A numerical method to emulate the ideal driver*, Veh. Syst. Dyn. 48 (2010), pp. 1461–1474.
[131] S. Schlenkrich, A. Walther, and A. Griewank, *Application of AD-based Quasi-Newton-Methods to Stiff ODEs, in Automatic Differentiation: Applications, Theory, and Implementations*, Springer, Wien/New York, 2005.
[132] P. Fevrier and G. Fandard, *Thermal and mechanical tyre modelling for handling simulation*, ATZ 110 (2008), pp. 422–431.
[133] B. Durand-Gasselin, T. Dailliez, M. Mössner-Beigel, S. Knorr, and J. Rauh, *Assessing the thermo-mechanical TaMeTirE model in off-line vehicle simulation and driving simulator tests*, in *Selected and Extended Papers from the 21st Symposium of the International Association for Vehicle System Dynamics*, Stockholm, Sweden, M. Berg and A.S. Trigell, eds. (Supplement to Vehicle System Dynamics) Vol. 48, Taylor & Francis, 2010, pp. 211–229.
[134] K. Guo and H. Guan, *Modelling of driver/vehicle directional control system*, Veh. Syst. Dyn. 22 (1993), pp. 141–184.
[135] C.C. MacAdam, *Understanding and modelling the human driver*, Veh. Syst. Dyn. 40 (2003), pp. 101–134.
[136] M. Plöchl and J. Edelmann, *Driver models in automobile dynamics application*, Veh. Syst. Dyn. 45 (2007), pp. 699–741.
[137] A.A. Popov, S. Rowell, and J.P. Meijaard, *A review on motorcycle and rider modelling for steering control*, Veh. Syst. Dyn. 48 (2010), pp. 775–792.
[138] M. Tomizuka and D.E. Whitney, *The human operator in manual preview tracking (an experiment and its modeling via optimal control)*, Trans. ASME, J. Dyn. Syst. Meas. Control 98 (1976), pp. 407–413.
[139] R.S. Sharp, *The steering responses of doubles*, in *Proceedings of the 6th IAVSD Symposium on The Dynamics of Vehicles on Roads and on Railway Tracks,* H.P. Willumeit, ed., Swets and Zeitlinger, Lisse, 1980, pp. 469–484.
[140] D.A. Crolla and R.S. Sharp, *Stability and steering response problems of trucks towing trailers*, in *Proceedings of the IMechE/MIRA Conference on Road Vehicle Handling*, MEP, London, 1983, pp. 201–210, ISBN 0 85298 515 0.
[141] C.C. MacAdam, *An optimal preview control for linear systems*, Trans. ASME 102 (1980), pp. 188–190.
[142] C.C. MacAdam, *Application of an optimal preview control for simulation of closed-loop driving*, IEEE Trans. Syst. Man Cybern. 11 (1981), pp. 393–399.
[143] G.N. Ornstein, *The automatic analog determination of human transfer function coefficients*, Med. Biol. Eng. 1 (1963), pp. 377–387.
[144] A. Phatak and G. Bekey, *Model of the adaptive behavior of the human operator in response to a sudden change in the control situation*, IEEE Trans. Man Mach. Syst. 10 (1969), pp. 72–80.
[145] R.W. Allen, T.J. Rosenthal, and H.T. Szostak, *Analytical modeling of driver response in crash avoidance maneuvering: Vol. 1: Technical Background*, National Highway Traffic Safety Administration, Department of Transportation, 1988, NHTSA, DOT HS 807 270.
[146] R.A. Hess and A. Modjtahedzadeh, *A control theoretic model of driver steering behavior*, Proc. IEEE Control Syst. Mag. 10 (1990), pp. 3–8.
[147] R.C. Miall, D.J. Weir, D.M. Wolpert, and J.F. Stein, *Is the cerebellum a Smith predictor*? J. Motor Behav. 25 (1993), pp. 203–216.
[148] P.R. Davidson, R.D. Jones, H.R. Sirisena, and J.H. Andreae, *Evidence for the formation of internal inverse models in the human motor system*, Hum. Mov. Sci. 19 (2000), pp. 761–795.
[149] C.C. MacAdam, Development of a driver model for near/at-limit vehicle handling, UMTRI, 2001 2001-43.
[150] H. Peng and M. Tomizuka, *Preview control for vehicle lateral guidance in highway automation*, Trans. ASME, J. Dyn. Syst. Meas. Control 115 (1993), pp. 679–686.
[151] R.S. Sharp, D. Casanova, and P. Symonds, *A mathematical model for driver steering control, with design, tuning and performance results*, Veh. Syst. Dyn. 33 (2000), pp. 289–326.
[152] C.C. MacAdam and G.E. Johnson, *Application of elementary neural networks and preview sensors for representing driver steering control behaviour*, Veh. Syst. Dyn. 25 (1996), pp. 3–30.
[153] R.S. Sharp and V. Valtetsiotis, *Optimal preview car steering control*, in *ICTAM Selected Papers from the 20th International Congress of Theoretical and Applied Mechanics*, P. Lugner and K. Hedrick, eds. (Supplement to Vehicle System Dynamics) Vol. 35, Swets and Zeitlinger, Lisse, 2001, pp. 101–117.

[154] Z. Artstein and A. Leizarowitz, *Tracking periodic signals with the overtaking criterion*, Trans. IEEE J. Autom. Control 30 (1985), pp. 1123–1126.

[155] A. Leizarowitz, *Tracking nonperiodic trajectories with the overtaking criterion*, Appl. Math. Optim. 14 (1986), pp. 155–171.

[156] G. Prokop, *Modelling human vehicle driving by model predictive on-line optimisation*, Veh. Syst. Dyn. 35 (2001), pp. 19–35.

[157] J.E. Bernard, L. Segel, and R.E. Wild, *Tire shear force generation during combined steering and braking maneuvers*, SAE 770852, SAE International, Warrendale, PA, (1977).

[158] H. Peng, *Evaluation of driver assistance systems – a human centered approach*, Proceedings of the 6th International Symposium on Automotive Control (AVEC 2002), September 9–13, SAE of Japan, Hiroshima, 2002, p. 8.

[159] M. Gerdts, *A moving horizon technique for the simulation of automobile test-drives*, Z. Angew. Math. Mech. 83 (2003), pp. 147–162.

[160] R.S. Sharp, *Driver steering control and a new perspective on car handling qualities*, Proc. IMechE, J. Mech. Eng. Sci. 219 (2005), pp. 1041–1051.

[161] A.Y. Ungoren and H. Peng, *An adaptive lateral preview driver model*, Veh. Syst. Dyn. 43 (2005), pp. 245–259.

[162] D.J. Cole, A.J. Pick, and A.M.C. Odhams, *Predictive and linear quadratic methods for potential application to modelling driver steering control*, Veh. Syst. Dyn. 44 (2006), pp. 259–284.

[163] S.D. Keen and D.J. Cole, *Steering control using model predictive control and multiple internal models*, Proceedings of the 8th International Symposium on Automotive Control (AVEC 2006), August 20–24, SAE of Japan, 2006, pp. 599–604.

[164] R.S. Sharp, *Optimal linear time-invariant preview steering control for motorcycles*, in *The Dynamics of Vehicles on Roads and on Railway Tracks*, Milan, Italy, (Supplement to Vehicle System Dynamics) Vol. 44, S. Bruni and G. Mastinu, eds., Taylor & Francis, 2006, pp. 29–340.

[165] R.S. Sharp, *Motorcycle steering control by road preview*, Trans. ASME, J. Dyn. Syst. Meas. Control 129 (2007), pp. 373–381.

[166] R.S. Sharp, *Optimal stabilisation and path-following controls for a bicycle*, Proc. IMechE C, J. Mech. Eng. Sci. 221 (2007), pp. 415–428.

[167] R.S. Sharp, *On the stability and control of the bicycle*, Trans. ASME, Appl. Mech. Rev. 61 (2008) 060803.

[168] R.S. Sharp, *Optimal preview speed-tracking control for motorcycles*, Multibody Syst. Dyn. 18 (2007), pp. 397–411.

[169] R.S. Sharp, *Application of optimal preview control to speed tracking of road vehicles*, Proc. IMechE C, J. Mech. Eng. Sci. 221 (2007), pp. 1571–1578.

[170] S. Rowell, A.A. Popov, and J.P. Meijaard, *Application of predictive control strategies to the motorcycle riding task* (Supplement to Vehicle System Dynamics) Vol. 46 (2008), pp. 805–814

[171] S. Rowell, *Optimal control to modelling motorcycle rider steering: Local versus global coordinate systems in rider preview*, Veh. Syst. Dyn. 48 (2010), pp. 429–456.

[172] P. Falcone, F. Borrelli, J. Asgari, H.E. Tseng, and D. Hrovat, *Predictive active steering control for autonomous vehicle systems*, IEEE Trans. Control Syst. Technol. 15 (2007), pp. 566–580.

[173] P. Falcone, *Linear time-varying model predictive control and its application to active steering systems: Stability analysis and experimental validation*, Int. J. Robust Nonlinear Control 18 (2008), pp. 862–875.

[174] A.J. Pick and D.J. Cole, *A mathematical model of driver steering control including neuromuscular dynamics*, Trans. ASME, J. Dyn. Syst. Meas. Control 130(3), (2008), 031004.

[175] A.J. Pick and D.J. Cole, *Application of linear preview control to modelling human steering control*, Proc. IMechE. D, J. Automob. Eng. 223 (2009), pp. 835–853.

[176] R. Roy, P. Micheau, and P. Bourassa, *Intermittent predictive steering control as an automobile driver model*, Trans. ASME, J. Dyn. Syst. Meas. Control 131(1) (2009), 014501.

[177] M. Thommyppillai, S. Evangelou, and R.S. Sharp, *Car driving at the limit by adaptive linear optimal preview control*, Veh. Syst. Dyn. 47 (2009), pp. 1535–1550.

[178] M. Thommyppillai, S. Evangelou, and R.S. Sharp, *Rear-heavy car control by adaptive linear optimal preview control*, Veh. Syst. Dyn. 48 (2010), pp. 645–658.

[179] M. Thommyppillai, S. Evangelou, and R.S. Sharp, *Towards a practical virtual racing-car driver*, *Proceedings of the European Control Conference*, Budapest, Hungary, August 23–26, 2009, p. 6.

[180] M. Thommyppillai, S. Evangelou, and R.S. Sharp, *Advances in the development of a virtual car driver*, Multibody Syst. Dyn. 22 (2009), pp. 245–267.

[181] R.S. Sharp, M. Thommyppillai, and S.A. Evangelou, *Towards a model of a high-performance driver for circuit-racing cars*, Proceedings of the International Symposium on Automotive Control (AVEC 2010), August 23–26, SAE of Japan, 2010, pp. 713–718.

[182] R.S. Sharp, *Application of linear optimal preview control theory to severe braking of a car,* Proc. IMechE, J. Mech. Eng. Sci. 226 (2010), pp. 1–15.

[183] R.S. Sharp, *On the stability and control of unicycles*, Proc. R. Soc., Ser. A 466 (2010), pp. 1849–1869.

[184] S.D. Keen and D.J. Cole, *Application of time-variant predictive control to modelling driver steering skill*, Veh. Syst. Dyn. 49(4) (2011), pp. 527–559.

[185] R.S. Sharp, *Rider control of a motorcycle near to its cornering limits*, Veh. Syst. Dyn. (2011), in review.