Optimal Control of a Formula One Car on a Three-Dimensional Track—Part 2: Optimal Control

D. J. N. Limebeer
Department of Engineering Science,
University of Oxford,
Parks Road,
Oxford OX1 3PJ, UK
e-mail: david.limebeer@eng.ox.ac.uk

G. Perantoni
Department of Engineering Science,
University of Oxford,
Parks Road,
Oxford OX1 3PJ, UK
e-mail: giacomo.perantoni@gmail.com

The optimal control of a Formula One car on a three-dimensional (3D) track is studied. The track is described by its geodesic and normal curvatures, and its relative torsion. These curvature parameters are obtained from noisy measurement data using the optimal estimation technique described in Part 1. The optimal control calculations presented are based on the aforementioned track model and a vehicle model that is responsive to the geometric features of a 3D track. For vehicle modeling purposes, the track is treated as a plane tangent to a nearby point on the track's spine. This tangent plane moves under the car and is orthogonal to the principal normal vector m at the nearby spine point. Results are presented that compare two-dimensional (2D) and 3D minimum-lap-time results, with the two compared. The Barcelona Formula One track studied in Part 1 is used again as an illustrative example. [DOI: 10.1115/1.4029466]

1 Introduction

Minimum-lap-time optimal control problems for Formula One cars have thus far been studied in two dimensions only. These 2D studies are reviewed in Refs. [1–3] and the citations therein. While the "flat track" assumption is often a valid and useful idealization, it neglects a number of influences that cumulatively may become important on some race circuits.

It is clear from everyday experience that a vehicle's top speed increases on downhill sections, while the opposite is true on ascending inclines. It is also known from this everyday experience that extended braking distances are required on descending gradients, while the opposite true on ascending ones. High-speed corners on modern roads are often banked (superelevated) in order to improve drainage and road holding, on the other hand, adverse cambering undermines road holding and reduces the maximum achievable cornering speeds. Slightly less obvious is the fact that a vehicle tends to become "light" if it is driven fast over the brow of a hill, while it will "squat" down on the suspension system if driven fast through a dip in the road. This variation in apparent mass is the result of a centripetal force, and has an effect similar to that produced by aerodynamic lift or down force. As with all changes in down force, this centripetal force can impact the top speed of the vehicle. Three dimensionality also introduces gyroscopic moments that change the lateral and longitudinal load transfer between the tires. As with all load transfer effects, these moments may change the vehicle's top speed and handling characteristics on some tracks. The aim of this paper is to study the significant influences on the minimum achievable lap time of a Formula One car resulting from 3D road geometries. Our focus will be on general-purpose analysis techniques and conceptual understanding, rather than on arriving at specific conclusions that relate to particular cars on specific tracks.

In Part 1, the 3D characteristics of the track are modeled using three curvatures rather than the single curvature used for 2D circuits. In the case of a flat (horizontal) track, the curvature, or

indeed the geodesic curvature, is used to describe the tightness of the bends. In the "flat-track" setting, geodesics are straight lines, or curves with no geodesic curvature. Once the track takes on 3D features, such as banked corners, elevation changes, and camber variations, the geometric and dynamic behaviors of the vehicle becomes significantly more complicated. If a car is driving along a straight road that then goes over the brow of a hill, a second curvature variable is required. This is called the normal curvature, which describes the curvature of the car's path in the vertical plane. As the reader will appreciate, normal curvature will produce centripetal forces that will change the car's apparent weight. The third curvature variable, the relative torsion, describes the way in which the road changes camber, or how it can adopt a "cork screw" motion, while going straight ahead. This cork screwing motion also produces centripetal and gyroscopic forces that may influence the dynamics and operating envelope of the car. It is self-evident that the importance, or otherwise, of 3D influences will depend on the geometry of the track under study. We aim to make these observations precise.

Section 2 describes the car model used in this study; the vehicle's heave, roll, and pitch freedoms have been neglected. Formula One cars are designed to have very stiff suspension systems; there are two reasons for this. First, a stiff suspension is a prerequisite for good road holding with the suspension minimizing variations in the tire normal loads. Second, road holding is improved substantially by aerodynamic down force. To achieve this, the car's skid block and front wing must be held close to the road, with minimal heave, pitch, and roll variations. It is assumed that the vehicle travels on a tangent plane that is orthogonal to the Darboux frame's normal vector m as defined in Part 1. The kinematics of this plane are studied in Sec. 2.1. This analysis is followed by a study of the vehicle's motion on the tangent plane. The car's equations of motion are derived in Sec. 2.2, with the vehicle data available in Appendix A. The tire load transfer is considered in Secs. 2.3 and 2.3.1, with tire model given in Appendix B. The aerodynamic drag and down force models are given in Sec. 2.4. The solution methodology for the minimum-lap-time optimal control problem is outlined in Sec. 3. The results are given in Sec. 4, with the conclusions summarized in Sec. 5.
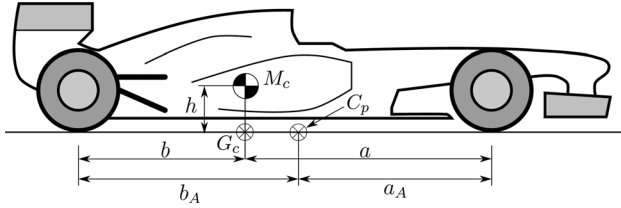
Fig. 1 Side view of a Formula One car showing its mass center M_c, its geometric center G_c, and its center of pressure C_p
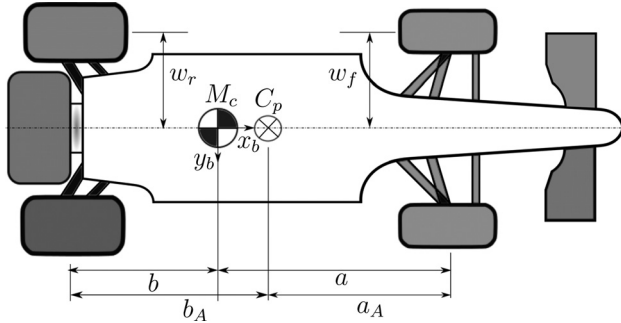

Fig. 2 Plan view of a Formula One car with some of its basic geometric parameters

2 Car Model

The track and car kinematics are modeled using the differential-geometric framework developed in Part 1. In the analysis given here we assume that the road surface is a plane orthogonal to the normal vector m as defined in Part 1 Sec. 2.3. This plane will move with the car, thereby generating a road surface that heaves, pitches, and rolls under the vehicle. The basic parameters of the car model are illustrated in Figs. 1 and 2. The side view shows the car's geometric center G_c and its aerodynamic center of pressure C_p in the road plane, as well as the vehicle's mass center M_c. The plan view shows the body-fixed axes x_b and y_b in the road plane with origin G_c, and the wheel-based parameters w_r and w_f.

A track segment is illustrated in Fig. 3, which shows an inertial axis system, a point s (the curvilinear abscissa) that travels with the car along the spine of the track and a vector n, which is perpendicular to the tangent vector t and that points toward G_c. The road normal m is perpendicular to t and n, and defines the moving road tangent plane. The magnitude of n is the perpendicular distance between the track's spine and the car's geometric center G_c. The angle ξ is the yaw angle of the car relative to the spine of the track. As was explained in Part 1, the track model is based on measured data with its curvature representation found as the solution of a subsidiary optimal control problem. The track's spatial curvature vector is given by $\boldsymbol{\Omega} = [\Omega_x \ \Omega_y \ \Omega_z]^T$ in which Ω_x is the relative torsion, Ω_y is the normal curvature and Ω_z the geodesic curvature. The track's angular velocity vector is given by $\boldsymbol{\omega} = \boldsymbol{\Omega}(ds/dt)$, where $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]^T$; in the unit-speed case $\boldsymbol{\omega} = \boldsymbol{\Omega}$, since $(ds/dt) = 1$.

2.1 Kinematics. The Euler angles underlying the track description will be needed to orientate the gravity vector. This information is computed as part of the optimal control problem solution described in Part 1 in the unit-speed case. Using Eq. (19) in Part 1 (we will use the notation (Part 1–19)), and $S(\boldsymbol{\omega}) = R^T \dot{R}$, we can find differential equations that describe the evolution of the Euler angles

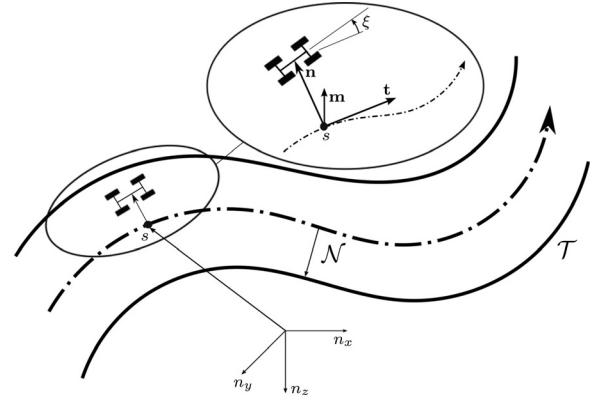$$\dot{\phi} = \omega_x + (\sin\phi\,\omega_y + \cos\phi\,\omega_z)\tan\mu \tag{1}$$

Fig. 3 Differential-geometric description of a track segment $\mathcal{T}$. The independent variable s represents the elapsed centerline distance traveled in the direction indicated. The track half-width is $\mathcal{N}$, with ξ the car's yaw angle relative to the spine's tangent direction. The inertial reference frame given by n_x, n_y, and n_z.

$$\dot{\mu} = \cos\phi\,\omega_y - \sin\phi\,\omega_z \tag{2}$$

$$\dot{\theta} = (\sin\phi\,\omega_y + \cos\phi\,\omega_z)\sec\mu \tag{3}$$

The angular velocity vector $\boldsymbol{\omega}$ is expressed in the Darboux frame with the dots representing time derivatives.

The next kinematic relationships we will require relate to the way in which the car progresses down the road. Suppose that the absolute velocity of the car in its body-fixed coordinate system is $v = [u\ v\ w]^T$; the longitudinal velocity component u is determined by the throttle/brakes, the lateral component v is determined by the steering, while the vertical component w is determined by the car's position on the track. The distance between the car's geometric center and the origin of the Darboux frame is $n = [0\ n\ 0]^T$ in Darboux coordinates. The angular velocity of the Darboux frame is $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]^T$.

We can now express the absolute velocity of the car in the Darboux frame as

$$\begin{bmatrix} \dot{s} \\ \dot{n} \\ 0 \end{bmatrix} = \boldsymbol{n} \times \boldsymbol{\omega} + R_z(\xi)v = \begin{bmatrix} n\omega_z + u\cos\xi - v\sin\xi \\ v\cos\xi + u\cos\xi \\ w - n\omega_x \end{bmatrix} \tag{4}$$

where

$$R_z(\xi) = R(\mathbf{e}_z, \xi) = \begin{bmatrix} \cos\xi & -\sin\xi & 0 \\ \sin\xi & \cos\xi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

represents the yawing of the car relative to the spine of the track. The first term in Eq. (4) derives from the angular velocity of the Darboux frame, while the second is the velocity of the car's geometric center expressed in Darboux coordinates. The first row of Eq. (4) gives the speed of the origin of the Darboux frame in its tangent direction and can be rewritten as

$$\dot{s} = \frac{u\cos\xi - v\sin\xi}{1 - n\Omega_z} \tag{6}$$

since $\omega_z = \dot{s}\Omega_z$ in the variable-speed case. This equation is used to generate $S_f(s) = (dt/ds)$ that transforms "time" as the independent variable into the "elapsed distance" as the independent variable. In order to fulfill this function, $S_f(s)$ and its inverse must be nonzero everywhere on the spine curve. The second row of Eq. (4) describes the way in which the vehicle moves normal to the spine

$$\dot{n} = u \sin \xi + v \cos \xi \tag{7}$$

The third row of Eq. (4) describes the absolute vertical velocity of the car's geometric center as a result of the road camber changes, and follows from the fact that the car cannot have a vertical velocity component in the m direction. This gives

$$w = n\omega_x \tag{8}$$

If the "distance traveled" is selected as the independent variable, then a complete set of kinematics equations includes Eq. (6), which generates $S_f(s)$

$$\frac{d\phi}{ds} = S_f \dot{\phi} \tag{9}$$

$$\frac{d\mu}{ds} = S_f \dot{\mu} \tag{10}$$

$$\frac{d\theta}{ds} = S_f \dot{\theta} \tag{11}$$

which come from Eqs. (1)–(3)

$$\frac{dn}{ds} = S_f(s)(u \sin \xi + v \cos \xi) \tag{12}$$

which comes from Eq. (7). Suppose $\bar{R} = R_z(\xi)R$, where R is defined in (Parts 1–19), then $S(\bar{\omega}) = \bar{R}^T \dot{\bar{R}}$ gives

$$\bar{\omega} = \begin{bmatrix} \bar{\omega}_x \\ \bar{\omega}_y \\ \bar{\omega}_z \end{bmatrix} = \begin{bmatrix} \cos \xi \omega_x + \sin \xi \omega_y \\ \cos \xi \omega_y - \sin \xi \omega_x \\ \omega_z + \dot{\xi} \end{bmatrix} \tag{13}$$

which is the absolute angular velocity of the car in its body-fixed reference frame. In the sequel, we will use the vehicle dynamics equations to find $\bar{\omega}_z$, with the car yaw angle ξ deduced from Eq. (13) by integrating

$$\frac{d\xi}{ds} = S_f(s)\bar{\omega}_z - \Omega_z \tag{14}$$

2.2 Dynamics. The equations describing the dynamics of the car can be derived using either the Newton–Euler, or the Lagrange equations. These approaches are similar, but the second is preferred here.

It follows from Eq. (8) that the absolute velocity of the car's mass center (expressed on the vehicle's coordinate system) is given by

$$v_B = v + \bar{\omega} \times h \tag{15}$$

$$= \begin{bmatrix} u \\ v \\ n\omega_x \end{bmatrix} + \begin{bmatrix} \bar{\omega}_x \\ \bar{\omega}_y \\ \bar{\omega}_z \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ -h \end{bmatrix} = \begin{bmatrix} u - h\bar{\omega}_y \\ v + h\bar{\omega}_x \\ n\omega_x \end{bmatrix} \tag{16}$$

where $\times$ denotes the cross product.

The car's kinetic energy is given by

$$T = \frac{1}{2}\left(Mv_B^T v_B + \bar{\omega}^T I_B \bar{\omega}\right) \tag{17}$$

in which M is the car's mass and I_B its inertia in the body-fixed coordinate system.

The car's potential energy is

$$U = -Mgz \tag{18}$$

where z is the height of the car's mass center relative to a datum in the inertial reference system.

The equation of motion are given by the Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_B}\right)\bigg|_I = F_B + \frac{\partial L}{\partial x_B} \tag{19}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \bar{\omega}}\right)\bigg|_I = M_B \tag{20}$$

where F_B and M_B are external forces and moments, respectively, acting on the car, and

$$L = T - U \tag{21}$$

is the system Lagrangian. The time derivatives in Eqs. (19) and (20) make use of the time-derivative operator

$$\frac{d\cdot}{dt}\bigg|_I = \frac{d\cdot}{dt}\bigg|_B + \bar{\omega} \times \cdot$$

which describes the transformation of the time derivative between inertial and body-fixed reference frames; see Chap. IV in Ref. [4].

Solving Eqs. (19) and (20) gives

$$M(\dot{v}_B + \bar{\omega} \times v_B) = F_B + MgR^T e_z \tag{22}$$

$$I_B \dot{\bar{\omega}} + \bar{\omega} \times (I_B \bar{\omega}) = M_B \tag{23}$$

where the car's inertia matrix is assumed diagonal $I_B = \mathrm{diag}(I_x\, I_y\, I_z)$, with $F_B = [F_x\, F_y\, F_z]^T$ and $M_B = [M_x\, M_y\, M_z]^T$ the external force and moment. The gravitational acceleration acting on the car's mass center is

$$gR^T e_z = R_z^T(\xi)R_x^T(\phi)R_y^T(\mu)\begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix}$$

$$= g\begin{bmatrix} \sin \xi \sin \phi \cos \mu - \cos \xi \sin \mu \\ \sin \xi \sin \mu + \cos \xi \sin \phi \cos \mu \\ \cos \phi \cos \mu \end{bmatrix} \tag{24}$$

The car's equations of motion can now be assembled from Eqs. (22)–(24) as follows:

$$\dot{u} = (v + h\bar{\omega}_x)\bar{\omega}_z - n\omega_x\bar{\omega}_y + h\dot{\bar{\omega}}_y$$
$$+ g(\sin \xi \sin \phi \cos \mu - \cos \xi \sin \mu) + F_x/M \tag{25}$$

$$\dot{v} = n\omega_x\bar{\omega}_x - (u - h\bar{\omega}_y)\bar{\omega}_z - h\dot{\bar{\omega}}_x$$
$$+ g(\sin \xi \sin \mu + \cos \xi \sin \phi \cos \mu)) + F_y/M \tag{26}$$

$$\dot{\bar{\omega}}_z = \left((I_x - I_y)\bar{\omega}_x\bar{\omega}_y + M_z\right)/I_z \tag{27}$$

in which F_x, F_y, and M_z are, respectively, the longitudinal and lateral tire and aerodynamic forces, and the z-axis tire moment acting on the car. These quantities are given by

$$F_x = \cos \delta(F_{frx} + F_{flx}) - \sin \delta(F_{fry} + F_{fly}) + F_{rrx} + F_{rlx} + F_{ax} \tag{28}$$

$$F_y = \cos \delta(F_{fry} + F_{fly}) + \sin \delta(F_{frx} + F_{flx}) + F_{rry} + F_{rly} \tag{29}$$

$$M_z = a\left(\cos \delta(F_{fry} + F_{fly}) + \sin \delta(F_{frx} + F_{flx})\right) - b(F_{rry} + F_{rly})$$
$$+ w_f\left(\sin \delta(F_{fry} - F_{fly}) + \cos \delta(F_{flx} - F_{frx})\right) + w_r(F_{rlx} - F_{rrx}) \tag{30}$$
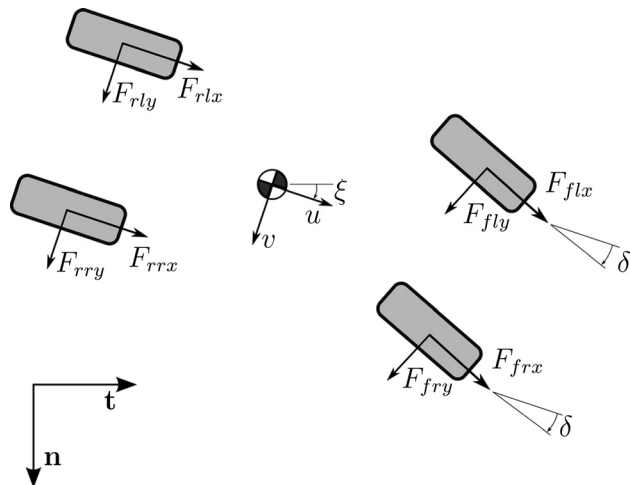
Fig. 4 Tire force system. The car's yaw angle is ξ with respect to the Darboux frame, which is defined in terms of the vectors t and n.

The aerodynamic drag force F_{ax} is given in Eq. (41), with the tire forces illustrated in Fig. 4 and described in detail in Appendix B.

The equations of motion (25), (26), and (27), expressed in terms of the elapsed arc length, are as follows:

$$\frac{du}{ds} = S_f \dot{u} \tag{31}$$

$$\frac{dv}{ds} = S_f \dot{v} \tag{32}$$

$$\frac{d\bar{\omega}_z}{ds} = S_f \dot{\bar{\omega}}_z \tag{33}$$

The angular acceleration of the Darboux (road) frame is neglected in the dynamic equations for the examples given here, since its influence is negligibly small.

2.3 Load Transfer. In order to compute the tire loads, we balance the forces acting on the car normal to the road, and then balance moments around the body-fixed x_b- and y_b-axes; see Fig. 2. These calculations must recognize the gravitational, inertial, centripetal, and aerodynamic forces acting on the car, as well as the road-related gyroscopic moments. The vertical force balance gives

$$n\dot{\omega}_x = (F_{rrz} + F_{rlz} + F_{frz} + F_{flz} + F_{az})/M \\ + g\cos\phi\cos\mu + (u - h\omega_y)\bar{\omega}_y - (v + h\omega_x)\bar{\omega}_x \tag{34}$$

in which the $F_{..z}$'s are the vertical tire forces acting on each of the four wheels, F_{az} is the aerodynamic down force acting on the car, the sixth term is the force due to gravity, while the seventh and eighth terms are centripetal forces.

Balancing moments around the car's mass center in the x_b direction gives

$$I_x\frac{d}{ds}\bar{\omega}_x = w_r(F_{rrz} - F_{rlz}) + w_f(F_{frz} - F_{flz}) - hF_y + (I_y - I_z)\bar{\omega}_z\bar{\omega}_y \tag{35}$$

The first two terms are the roll moments produced by the vertical tire forces, the third term is the roll moment produced by the lateral tire forces F_y (see Eq. (29)), while the fourth term is a gyroscopic moment acting in the x_b direction; see Fig. 2.

Balancing moments around the car's mass center in the y_b direction gives

$$I_y\frac{d}{ds}\bar{\omega}_y = b(F_{rrz} + F_{rlz}) - a(F_{frz} + F_{flz}) + hF_x \\ + (a_A - a)F_{az} + (I_z - I_x)\bar{\omega}_z\bar{\omega}_x \tag{36}$$

The first two terms represent the pitching moments produced by the vertical tire forces, the third term is the pitching moment produced by the longitudinal force F_x (see Eq. (28)), the fourth term comes from the aerodynamic down force, while the fifth term is a gyroscopic moment acting in the y_b direction.

Equations (34)–(36) are a set of linear equations in four unknowns. A unique solution for the tire loads can be obtained by adding a suspension-related roll balance relationship, in which the lateral load difference across the front axle is some fraction of the whole

$$F_{frz} - F_{flz} = D(F_{frz} + F_{rrz} - F_{flz} - F_{rlz}) \tag{37}$$

where $D \in [0, 1]$. If this fourth equation is recognized, and Eqs. (34)–(37) are linearly independent, a unique solution results.

2.3.1 Non-Negative Tire Loads. The forces satisfying Eqs. (34)–(37) are potentially both positive and negative. Negative forces are indicative of vertical reaction forces, while positive forces are fictitious "forces of attraction." Since the model being used here has no pitch, roll, or heave freedoms, none of the wheels is free to leave the road, while simultaneously keeping faith with Eqs. (34)–(37); an approximation is thus required.

To cater for the possible "positive force" situation within a nonlinear programming (NLP) environment, we introduce the tire normal load vector $\bar{F}_z$, which may contain positive components. These are set to zero by defining $F_z = \min(\bar{F}_z, 0)$, where the minimum function $\min(\cdot, \cdot)$ is interpreted element-wise. It is clear that $\bar{F}_z$ and F_z will be equal unless at least one entry of $\bar{F}_z$ is positive (i.e., nonphysical). Since the model must respect the laws of mechanics, Eqs. (34)–(36) must be enforced unconditionally. In contrast, we assume that the solution to Eq. (37), which is only an approximate representation of the suspension system, can be "relaxed" in the event of a tire normal load sign reversal.

Equations (34)–(36) are arranged in matrix form as

$$A_1 F_z = c \tag{38}$$

while Eq. (37) can be written as

$$A_2 F_z = 0 \tag{39}$$

The entries in the matrices A_1 and A_2, and the vector c can be assembled from Eqs. (34)–(37). In order to deal with the "light wheel" situation, we combine Eqs. (38) and (39) as

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} F_z \\ \bar{F}_z \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix} \tag{40}$$

in which F_z in Eq. (39) has been replaced by $\bar{F}_z$. If there is a "light wheel," the mechanics Eqs. (34)–(36) will be satisfied by the nonpositive forces F_z, while the roll balance equation is satisfied by the now fictitious forces $\bar{F}_z$ that contain a "force of attraction." It is clear that the tire normal loads have to satisfy the nonlinear circularly dependent relationship (40), which will be solved by a NLP algorithm. The light wheel modeling is covered in more detail in Ref. [2].

2.4 Aerodynamic Loads. The external forces acting on the car come from the tires and from aerodynamic influences. The aerodynamic force is applied at the center of pressure, which is located in the vehicle's plane of symmetry. The drag and lift forces are given by
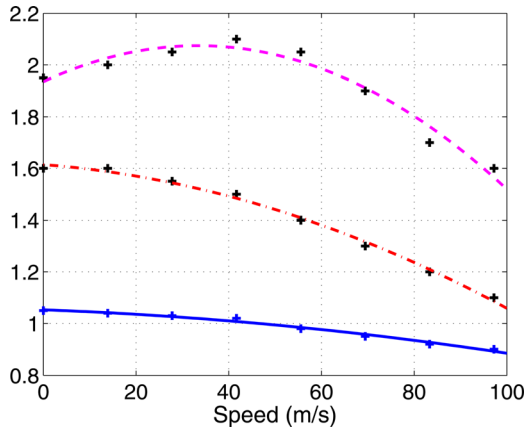
Fig. 5 Car aerodynamic maps. The drag coefficient C_D is the solid (blue) curve; the down-force coefficient C_L is the dotted–dashed (red) curve. The aerodynamic center of pressure is given by the dashed (magenta) curve in meters from the front axle. The "+" symbols represent measured data points.

$$F_{ax} = -0.5 C_D \rho A u^2 \qquad (41)$$

and

$$F_{az} = 0.5 C_L \rho A u^2 \qquad (42)$$

respectively. The speed-dependent drag and down-force coefficients, and the speed-dependent location of the aerodynamic center of pressure are given in Fig. 5.

2.5 Other Effects. In order to optimize the vehicle's performance one needs to control the torques applied to the individual road wheels. In current Formula One race cars the braking system has to be designed so that equal pressure is applied to the brake callipers of each axle, with the braking pressures between the front and rear axles satisfying some brake balance ratio. The drive torques applied to the rear wheels are controlled by a differential mechanism. Since the braking system and the differential can be modeled as if the road were flat, we refer the reader to Refs. [2,3] for an account of these features of the model.

3 Optimal Control

General purpose numerical optimal control problem solvers are becoming common place, but skill is still required in the use of these tools if reliable solutions are to be obtained. When solving new problems for the first time, it is prudent to be sceptical about the solutions obtained until they have been thoroughly checked from multiple initial conditions and if possible against measured data. One should also verify that parameter changes produce the trends expected and if possible consistency against alternative optimization algorithms should be checked. As has been point out elsewhere: ".just as there is no fool proof method, there is no fool proof advice [5]." Other important issues relate to: (a) ensuring that the problem is properly posed and that it is compatible with the theory underpinning the solution methodology; (b) discretization techniques; (c) scaling; (d) the treatment of nonsmooth problem features; (e) problem regularization; (f) the computation of gradients, and (g) mesh refinement. For several reasons, we have decided to use a direct numerical method based on an orthogonal collocation to solve the problem presented here. In the remainder of this section, we will review salient features of the theory used and present brief notes on each of the five topic alluded to.

3.1 Optimal Control Problem. A wide class of optimal control problems can be posed in Mayer–Pontryagin form,

recognizing that the more familiar *Bolza* formulation can be converted into this form if so desired [6].

Suppose the system to be controlled is described by the state-space model

$$\frac{dx}{d\tau} = f(x(\tau), u(\tau), \tau), \quad x(-1) = x_0 \qquad (43)$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, with x_0 given. The optimization interval $[-1, 1]$ is particularly well suited to pseudospectral methods, but it can be transformed to the more general interval $[t_0, t_f]$ with the affine transformation $t = (t_f - t_0)\tau/2 + (t_f + t_0)/2$. The control vector u may be subject to control constraints with $u \in \mathcal{U}$, where $\mathcal{U}$ is some region in the control space. The problem may also contain path constraints, but these have been omitted in this brief summary. The optimization problem is to find a trajectory (i.e., a solution of Eq. (43) consistent with the boundary and control constraints) and that minimizes

$$\Phi(x(1)) \qquad (44)$$

for some scalar function $\Phi(\cdot)$ of the terminal state.

In accordance with any of the standard texts such as Refs. [7–9], first-order necessary conditions for optimality are given by the Pontryagin minimum principle (PMP)

$$u^* = \arg \min_u \mathcal{H}(x, \lambda, u, \tau) \qquad (45)$$

where $\mathcal{H}$ is the Hamiltonian function

$$\mathcal{H}(x, \lambda, u, \tau) = \lambda^{\mathrm{T}} f \qquad (46)$$

The control u^* is an extremal control with λ an n-dimensional costate vector that is a continuous function of time. In this notation, Eq. (43) and λ are given by

$$\dot{x} = \nabla_\lambda \mathcal{H} \quad x(-1) = x_0 \qquad (47)$$

$$\dot{\lambda} = -\nabla_x \mathcal{H} \quad \lambda(1) = \nabla \Phi(x(1)) \qquad (48)$$

Equations (47) and (48), in combination, form a two-point boundary value problem in which the initial state and the terminal costate are specified. In the case of arcs for which u is in the interior of $\mathcal{U}$, assuming the existence and continuity of the relevant partial derivatives, Eq. (45) implies that

$$\nabla_u \mathcal{H} = 0 \qquad (49)$$

with the weak Legendre–Clebsch condition

$$\nabla_{uu} \mathcal{H} \geq 0 \qquad (50)$$

also satisfied. Condition (50) is often overlooked, but is nonetheless important. The singularity of $\nabla_{uu} \mathcal{H}$ may be a significant issue in minimum-time and/or minimum-fuel problems.

3.1.1 Discretization. The minimum-lap-time optimal control problem will be solved using GPOPS-II [10], which is direct pseudospectral method based on Legendre–Gauss–Radau (LGR) collocation and Radau's integration formula; see page 103 in Ref. [11]. In this framework, the state and controls are treated as parameters on a finite collocation, or implicit integration mesh that is determined by N Radau points $(\tau_0, \ldots, \tau_N)$. These points are the roots of $P_{N+1}(-\tau) + P_N(-\tau)$, in which $P_N(\tau)$ is the Nth-degree Legendre polynomial; these roots lie in $[-1, 1)$ with one root at $\tau_0 = -1$ always. The Radau integration formula is exact for polynomials of degree $2N - 2$ and lower. In this discretization process, the state is approximated in a basis of Lagrange polynomials

$$\mathcal{L}_i(\tau) = \prod_{\substack{j \neq i \\ j=0}}^{N} \frac{\tau - \tau_j}{\tau_i - \tau_j} \quad i = 1, \ldots, N \tag{51}$$

The point $\tau_0 = -1$ is not collocated and the time derivative of the state is not approximated there. The state in Eq. (43) is approximated by

$$x(\tau) \approx \begin{bmatrix} \mathcal{L}_0(\tau) & \cdots & \mathcal{L}_N(\tau) \end{bmatrix} \begin{bmatrix} x(\tau_0) \\ \vdots \\ x(\tau_N) \end{bmatrix} \tag{52}$$

in which the state vector is treated as a *row vector* $\begin{bmatrix} x_1(\tau) & \cdots & x_n(\tau) \end{bmatrix} \in \mathbb{R}^n$. This means that

$$\begin{bmatrix} \dot{x}(\tau_1) \\ \vdots \\ \dot{x}(\tau_N) \end{bmatrix} \approx \begin{bmatrix} \dot{\mathcal{L}}_0(\tau_1) & \cdots & \dot{\mathcal{L}}_N(\tau_1) \\ \vdots & \ddots & \vdots \\ \dot{\mathcal{L}}_0(\tau_N) & \cdots & \dot{\mathcal{L}}_N(\tau_N) \end{bmatrix} \begin{bmatrix} x(\tau_0) \\ \vdots \\ x(\tau_N) \end{bmatrix} \tag{53}$$

The $N \times (N+1)$ matrix on the right-hand side of Eq. (53) is a constant differentiation matrix. If the first column of the differentiation matrix in Eq. (53) is removed, the resulting $N \times N$ matrix is nonsingular with its inverse a LGR integration matrix; LGR collocation can be viewed as an implicit numerical integration scheme. Using this discretization, the optimal control problem given in Sec. 3.1 can be transformed into a NLP problem with an established correspondence between the Karush–Kuhn–Tucker optimality conditions for the NLP and the first-order necessary conditions [12]. The convergence rate for the LGR collocation method used in GPOPS-II is established in Ref. [13], with this analysis extended to *hp*-adaptive schemes in Ref. [14]. In the case of unconstrained problems, convergence of the NLP on a LGR mesh ensures convergence of the state, the costate, and the controls with error estimate of the form $\mathcal{O}\big(l/N^{(l-(5/2))}\big)$, where l is the number of continuous derivatives in the solution and N is the degree of the approximating polynomials in the collocation scheme. Convergence conditions for general nonlinear optimal control problems on LGL meshes are given in Ref. [15] and the references therein.

3.2 Scaling. Scaling can have a significant influence on the performance of optimization algorithms. Since convergence tolerances, and other criteria, are necessarily based on some notion of "small" and "large" quantities, problems with unusual or unbalanced scaling may cause difficulty. One notion of scaling is to transform the variables from their original representation, which may reflect the physical nature of the problem, to dimensionless quantities that have desirable properties in terms of optimization. In our work, we take the length of the car to be the fundamental unit of length—after scaling, the car has a dimensionless length of one. In the same way, the fundamental unit of mass is the mass of the car—after scaling, the car has a dimensionless mass of one. In order to scale time we give the car unit weight and so a time scale of $\sqrt{g/l_0}$ is used in which l_0 is the car's length. If the car's mass is m_0 kg, then its weight is $m_0 g$ N, which after scaling becomes unity. During racing, the car's speed might be 90 m/s; after scaling, this becomes 15.58 in dimensionless units. During firm acceleration and fast cornering, the car's longitudinal and lateral tire forces might reach 8000 N; this becomes 1.236 in dimensionless units. The scaling scheme should force the NLP's decision variables into a spherical space rather than a highly elongated hyperellipsoid.

3.3 Nonsmooth Features. Interior point NLP algorithms such as IPOPT require first- and second-order derivative information for the functions defining the cost and constraints. For this reason nonsmooth problem features have to be approximated in a way that does not change significantly the problem's solution. Functions such as $\min(x,0)$ and $\max(x,0)$ have undefined derivatives at $x=0$, and are therefore approximated using

$$\max(x,0) \approx \frac{x + \sqrt{x^2 + \epsilon}}{2} \quad \text{and} \quad \min(x,0) \approx -\frac{-x + \sqrt{x^2 + \epsilon}}{2} \tag{54}$$

in which ϵ is a small constant. Since

$$\frac{\partial}{\partial x}\left(\frac{1}{2}\left(x + \sqrt{x^2 + \epsilon}\right)\right) = \frac{1}{2}\left(\frac{x}{\sqrt{x^2 + \epsilon}} + 1\right) \tag{55}$$

one sees that the derivative approximation for $\max(x,0)$ at $x=0$ is now well defined. One may also make use of $\max(a,b) = a + \max(0, b-a)$. In the case of $|x|$, one might use

$$|x| \approx \sqrt{x^2 + \epsilon}$$

in this case

$$\frac{\partial}{\partial x}\sqrt{x^2 + \epsilon} = \left(\frac{x}{\sqrt{x^2 + \epsilon}}\right)$$

with the derivative approximation at $x=0$ again well defined. These approximations become more accurate as the value of ϵ is reduced, with values in the range of $10^{-5} \leq \epsilon \leq 10^{-2}$ typical for the results given here.

3.4 Regularization. When the controls enter the system dynamics and performance index linearly, the possibility of bang–bang controls and/or singular arcs exists. This type of problem occurs, for example, when one seeks to vary the thrust magnitude of a rocket in order to achieve maximum altitude [7]. Bang–bang controls also appear in car racing if the drive/braking torque is varied in order to minimize transit times (p. 247 [8]). If the control Hamiltonian ceases to be an explicit function of the control variable u, no information about the optimal control can be gleaned from the PMP; these problems are referred to as singular. In other words, if $\mathcal{H}$ does not depend upon u explicitly, the usual procedure of selecting u^* so as to maximize $\mathcal{H}$ breaks down and the PMP cannot determine directly a unique optimal control as a function of the state and costate variables. Instead, the optimal control must be determined by a new set of necessary conditions [16–19]. All the general-purpose optimal control software known to the authors assume the strong form of the Legendre–Clebsch necessary condition (i.e., $\nabla_{uu}\mathcal{H} > 0$), and so these codes cannot be used to solve problems containing singular arcs.

In order to avoid oscillatory solutions and nonconvergent NLP behavior in these problems, it is sometimes useful to introduce into the performance index small terms that are quadratic in the controls. These terms ensure, at least in theory, that the PMP can be used to find the optimal controls (for the perturbed problem). If these terms are small relative to the elapsed time, they do not change the problem in a significant way. In our case, we used an index of the form

$$J = \int_0^T \left(1 + \sum_{i=1}^{m} \epsilon_i u_i^2\right) dt \tag{56}$$

in which the ϵ_i's are small constants with the u_i's the controls. This is clearly no longer a pure minimum-time problem, but is close to one if the ϵ's are chosen sufficiently small. Instead of using J as a measure of the minimum lap time, the "true" lap time is computed separately as an auxiliary state. In our case, the minimum value of J is of the order 80 s, while the computed true lap

time is typically 0.1 s smaller; these differences are comparable with the changes obtained with variable numbers of mesh refinements.

3.5 Computing Gradients. Gradient-based methods for solving NLPs require the derivatives of the objective function and constraints with respect to the states and the controls. The most obvious way for computing these derivatives is analytically, either by hand, or by using a computer algebra package such as Maple™. While the exactness of this approach is appealing, and generally results in faster rates of convergence, it is sometimes not practical to compute derivatives this way. Another approach is to compute derivatives numerically using finite differencing. In the case of forward differencing, one uses

$$\frac{df}{dx} \approx \frac{f(x + \Delta) - f(x)}{\Delta}$$

in which Δ is the step length. The difficulty with this approach is that Δ must be chosen small enough to provide a good approximation to the derivative, but not so small that round-off errors occur when calculating the difference $f(x + \Delta) - f(x)$. Another approach is *complex-step differentiation*, which is both accurate and efficient [20,21]. In the case of higher-order derivatives, a more general version of the complex-step derivative approximation, called the multicomplex-step derivative approximation, can be used as a way of computing Hessians without truncation errors [22]. Another powerful approach is automatic differentiation that computes function derivatives to the precision that one would achieve using analytic or symbolic differentiation. While any of these approaches can be used, in our experience automatic or symbolic differentiation produce the most reliable results.

3.6 Mesh Refinement. The numerical solution of optimal control problems using direct transcription methods involves three key steps: (a) transcribing the optimal control problem into a NLP problem using a numerical integration scheme; (b) solving the NLP; (c) reviewing the accuracy of the solution and, if necessary, refining the mesh and resolving the (refined) problem. The accuracy and efficiency of this process can be influenced by many things including the choice of integration scheme, the solution mesh, the mesh refinement strategy (if one is used), and the problem characteristics themselves (cf. singular arcs, stiff and/or nonsmooth dynamics, and/or rapidly varying path constraints).

When considering transcription methods in general, the factors favoring the use of simple low-order integration schemes such as the trapezoidal rule include sparse Jacobian matrices, sparse Hessians, and fewer right-hand-side evaluations as compared with complex implicit integration schemes. Set against these advantages is the fact that these simple integration schemes have relatively poor accuracy and so a fine integration mesh is required.

Global orthogonal collocation methods lie at the other end of the spectrum. In these cases the state and costate are approximated by polynomials over the entire optimization interval, with increased accuracy coming from increasing the degree of the approximating polynomials used. For problems with smooth solutions pseudospectral (or orthogonal collocation) methods converge rapidly as is evident from the standard error bounds for Radau integration

$$E_N = \frac{N 2^{(2N-1)} [(N-1)!]^4}{[(2N-1)!]^3} f^{(2N-1)}(\eta) \tag{57}$$

in which N is the degree of approximating polynomial, $f^{(2N-1)}$ is the $2N - 1$ th derivative of the integrand and $-1 < \eta < 1$ [11,23]. The exponential nature of this error bound is illustrated in Fig. 6. For problems with smooth solutions, convergence rates of $\mathcal{O}(e^{-cN})$ are routinely achieved; N is the order of the
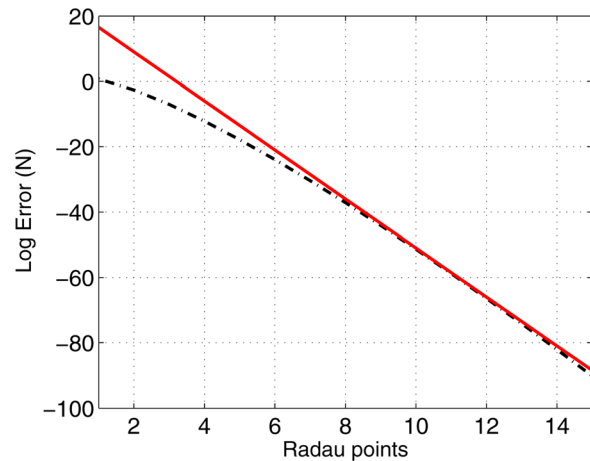

Fig. 6 Radau integration error. (Black) dotted–dashed line in the error bound (57) with $f^{(2N-1)}(\eta) = 1$, with the (red) line given by $\log(E_N) = 24 - 7.5\,N$.

approximating polynomial and c is a constant. In the case of nonsmooth problems, the convergence rate of global pseudospectral methods will be slow (as the degree of the approximating polynomial increases), and a poor approximation may result even if high-degree polynomials are used. A second limitation of global orthogonal collocation methods is that high-degree polynomials result in a dense NLP.

An alternative, which seeks the best of both of these worlds, is to segment the problem (as with conventional integration algorithms), and then employ global orthogonal collocation techniques within each segment. GPOPS II uses a two-tiered grid refinement strategy that refines both the problem segmentation and the orthogonal polynomial orders [24]. If the error across a particular segment has a uniform behavior, the number of collocation points is increased. If the error at an isolated point within a segment is significantly larger than the errors at other points within the segment, the segment is subdivided. The interested reader can refer to Figs. 5 and 6 in Ref. [3] for an example of adaptive grid refinement in a race car application.

3.7 Problem Setup. In order to solve the minimum-lap-time optimal control problem, one has to assemble a performance index, which in our case is the regularized lap time (56), the constraints associated with the problem dynamics, the path constraints, and a constraint on the engine power.

The key kinematic equations required are: Eq. (6) that describes the relationship between the elapsed time and the elapsed distance, Eqs. (9)–(11) that describe the evolution of the Darboux frame Euler angles, and Eq. (12) that describes the evolution of the car's position relative to the spine of the track. A path constraint on $n(s)$ is used to ensure that the car remains on the track; the track width may be variable and thus a function of s. Equation (14) is used to describe the evolution of the vehicle's yaw angle relative to the spine.

The key dynamic equations include the longitudinal and lateral force balance equations (31) and (32), which require Eqs. (28) and (29), and the yaw moment balance equation (33). The tire force constraint equations include the normal load balance condition (34); the roll and pitch load transfer conditions (35) and (36); and the roll balance constraint (37). To complete the tire force description, one also requires the normal load constraint (40) and the tire force description given in Appendix B. Path constraints are included in the problem setup to ensure that tire normal loads do not change sign, thereby producing "nonphysical" forces of attraction between the tire and the road. The tire slips are constrained to remain within a region of realistic tire wear.
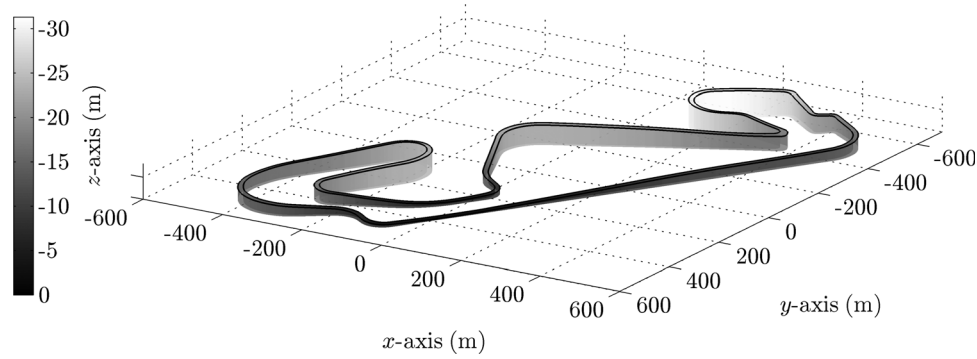
Fig. 7   3D view of the "Circuit de Catalunya" with all its dimensions given in meters
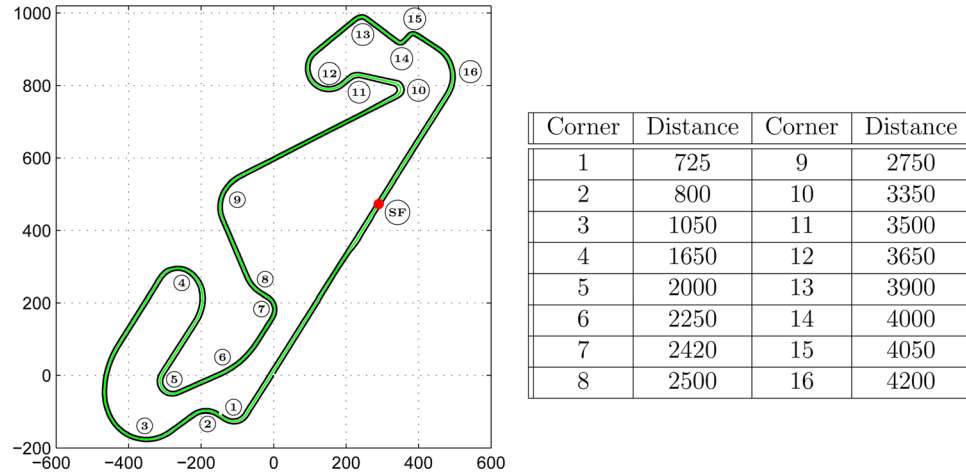

Corner	Distance	Corner	Distance
1	725	9	2750
2	800	10	3350
3	1050	11	3500
4	1650	12	3650
5	2000	13	3900
6	2250	14	4000
7	2420	15	4050
8	2500	16	4200

Fig. 8   Approximate distances to mid corner on the Circuit de Catalunya track (in meters from the start–finish line ⓢⒻ)

The aerodynamic forces are given by Eqs. (41) and (42). One also needs to include longitudinal tire force constraints that are associated with the brakes and the differential—one possible set of constraint equations can be found in Ref. [2]. Another path constraint ensures that the delivered engine power is restricted to that physically available.

Cyclic constraints are placed on all of the vehicle states to ensure that they do not change across the start-finish line. In other words, while the initial and final vehicle states are free within the physical limits of the problem, they must be equal (as they cross the start-finish line).

4   Results

In order to gain a proper insight into the 3D behavior of the optimally controlled car, it is necessary to understand the relevant features of the track. A 3D view of the Barcelona track is shown in Fig. 7. The corner numbering scheme and the approximate distances to each corner from the start-finish line are given in Fig. 8.

Figure 9 shows the track elevation and geodesic curvature as a function of the distance from the start–finish line. In the sign convention used here, positive geodesic curvature is associated with right-hand corners. The z-axis points downward and so points of high elevation have large negative values. In order to determine if a corner has a positive or negative camber, the product of the geodesic curvature and the camber angle is computed and plotted in Fig. 10. Since positive values of geodesic curvature are associated with right-handed corners, and positive values of camber angle indicate clockwise cambering, a positive product indicates positive camber, while a negative value for the product indicates adverse cambering.

Adverse cambering limits the cornering speed for the reason set out in Fig. 11 and the following single-axle analysis for steady-state cornering in bend of radius r. Balancing forces normal to the road gives

$$F_{zl} + F_{zr} = m\left(g\cos\beta + \frac{u^2}{r}\sin\beta\right) \quad (58)$$

while balancing forces parallel to the road gives

$$F_{yl} + F_{yr} = m\left(g\sin\beta - \frac{u^2}{r}\cos\beta\right) \quad (59)$$

Taking moments around the right wheel ground contact point gives

$$F_{zl} = \frac{m}{2}\left(g\cos\beta + \frac{u^2}{r}\sin\beta\right) - \frac{mh}{2w}\left(g\sin\beta - \frac{u^2}{r}\cos\beta\right) \quad (60)$$

and so

$$F_{zr} = \frac{m}{2}\left(g\cos\beta + \frac{u^2}{r}\sin\beta\right) + \frac{mh}{2w}\left(g\sin\beta - \frac{u^2}{r}\cos\beta\right) \quad (61)$$

These equations show that on a level road the tires' lateral forces act toward the center of curvature of the turn and that the car tends to roll out of the corner making the right-hand tire normal load lower (higher) than that of the left-hand tire in a right- (left-) hand turn. Equation (58) shows that positive camber tends to increase
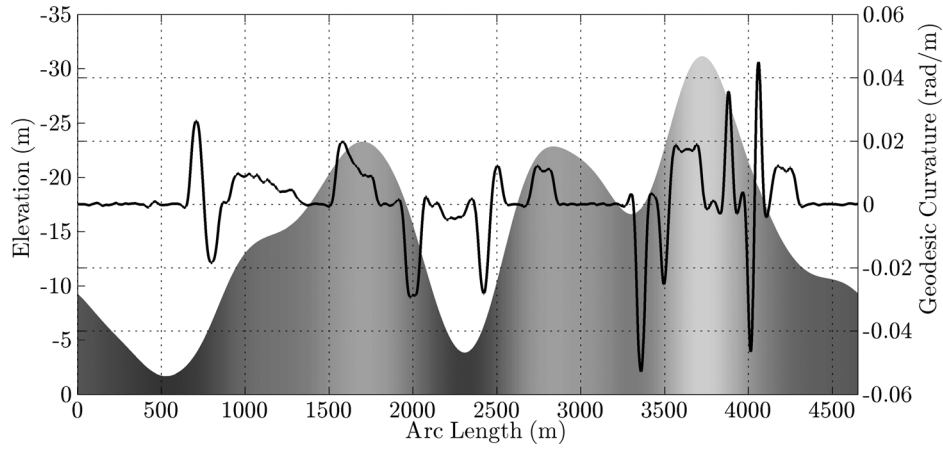
Fig. 9 Elevation and geodesic curvature of the track. Distances are given in meters, while the curvature is given in radians per meter. Positive curvature is associated with right-hand bends.
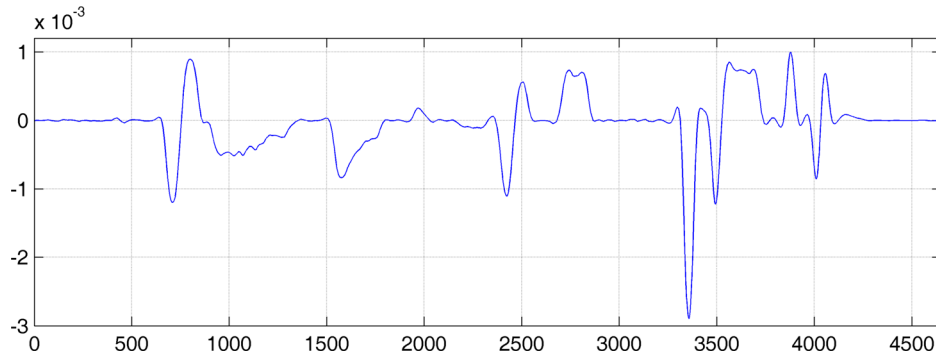

Fig. 10 Product of the geodesic curvature and the track camber angle. Positive values are indicative of positive cambering, while negative values are associated with adverse cambering.

the normal load on both tires, Eq. (59) shows that positive camber reduces the lateral forces required to hold the corner and Eqs. (60) and (61) show that positive camber tends to redistribute the normal tire loads so that they are more nearly equal. The performance improvement derived from positive cambering is thus threefold, and the speed of the car can be increased above that achievable on a flat road. The opposite effects occur in a corner
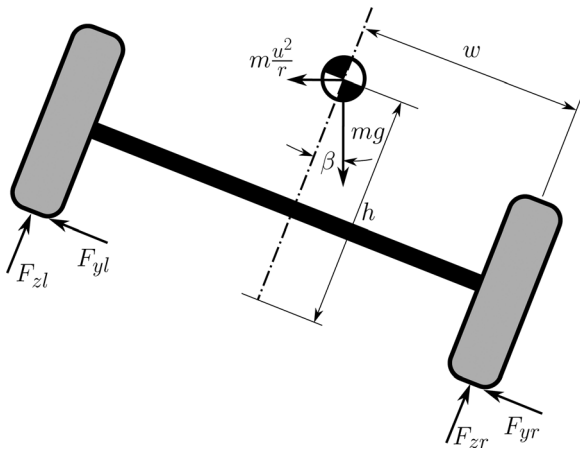

Fig. 11 Forces acting on a cornering car at speed *u*, turn radius *r* and road camber angle *β*. The car is traveling into the page and turning right.

with adverse camber. The presented optimal control results were checked against different meshes, different NLP solvers, different gradient calculation techniques, different initial conditions, as well as the solutions obtained with a different transcription process [2].

Speed profiles of the car on the 3D track and a 2D simplification, which was obtained by zeroing the track elevation data, are shown in Fig. 12. This figure also shows the differences in the times required to reach each point on the track. Positive values indicate that the car on the 3D track is faster. In this case, the optimal lap time for the 3D track is 80.25 s, while that for the 2D track is 79.58 s, giving a difference of −0.67 s. This figure shows that there are speed differences of up to ±10 mi/h at various points on the circuit. These differences come primarily from three influences: a gravitational force component in the vehicle's direction of travel, positive or adverse cambering, and a centripetal force component normal to the road surface. At 90 m/s, a 100 N gravitational force component in the direction of travel is equivalent to a 9 kW variation in engine power.

The minimum-lap-time optimal control calculations given in Ref. [3] show that switching strategies may be used in the control of the car's braking and engine power delivery; see, for example, Figs. 10, 14, and 16 in that publication. Figure 13 again shows power controls that are reminiscent of bang–bang type switching strategies. To see how bang–bang controls arise, consider the following heuristic argument: Suppose the car is driven in a straight line on a level road with the longitudinal force balance given by

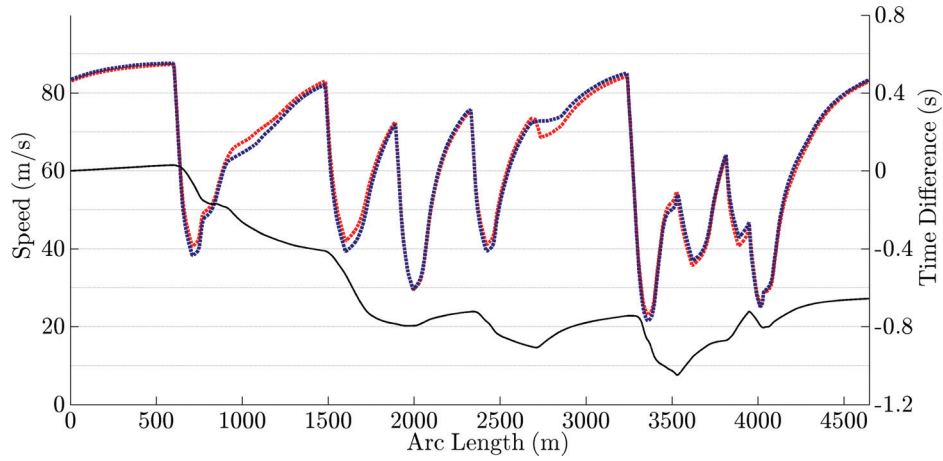$$M\dot{u} = F_{\mathrm{d}} - C_{\mathrm{d}}u^2 \tag{62}$$

Fig. 12 Speeds and time difference for the 2D and 3D track. The (red) curve gives the speed of the car on the flat track, while the (blue) curve is the speed on the 3D track. The solid (black) curve is the time difference; negative values indicate that the car is faster on the 2D track.
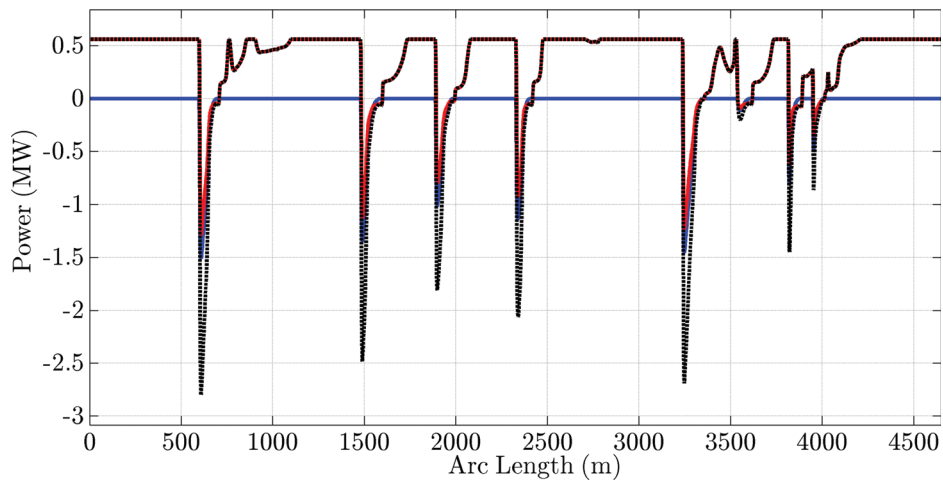

Fig. 13 Engine and braking power as a function of distance from the start/finish line. The (red) curve illustrates the power delivered to the rear wheels, the (blue) shows the power delivered to the front wheels, while the (black) dashed plot is the sum of the two.

In this equation, the externally applied longitudinal drive/braking force is F_d, $C_d u^2$ is the aerodynamic drag, and M is the vehicle's mass. If we multiply this equation by the vehicle speed u, one obtains

$$\frac{M}{2}\frac{d}{dt}(u^2) = P_d - C_d u^3 \qquad (63)$$

in which P_d is the drive/braking power. Importantly Eq. (63) is linear in the control P_d. This model in combination with the minimum-time performance index $J = \int_0^T dt$ and the maximum principle produces bang–bang type optimal controls. In reality, the situation is far more complicated than this, because the tire forces involve nonlinear combined slip and state-dependent saturation effects.

Figure 14 shows that 3D track features can cause the car to change its racing line in order to find the best compromise between the many factors limiting the car's top speed. In turns ② and ③, the vehicle experiences adverse and then positive camber, it is running uphill throughout and the vertical centripetal force is positive and then negative. This combination of effects is causing the car to take a significantly different racing line in the two cases. In turn ④, the car is going over the brow of a hill, it experiences adverse camber for the majority of the corner and the centripetal

force acts to undermine the aerodynamic down force. In this case, however, this combination of influences has little effect on the racing line taken.

We end with a more detailed analysis of the effect of gravitational, centripetal, and camber changes on the vehicle's top speed. Figure 15 shows the speed variations due to 3D influences and the longitudinal gravitational force as a function of distance, while the "centripetal down force" in illustrated in Fig. 16. We will concentrate on three track sections, with others open to a similar study should the reader wish to undertake one.

The first section is the down-hill straight from turn ⑯ to turn ①. The average speed difference on this section is approximately 0.5 m/s. Camber and centripetal force influences appear relatively unimportant on this section of the track. Since the longitudinal gravitation force component is of the order 100 N, which is acting on the car at approximately 85 m/s, this gradient is equivalent to an increase in the engine power of approximately 8.5 kW. An easy calculation using Eq. (41) shows that the peak engine power of 560 kW is capable of supporting a top speed of 85.4 m/s against aerodynamic drag. If the engine power is increased to 569 kW the top speed of the car is limited by aerodynamic drag to 85.9 m/s, which explains this speed difference.

The second track section we will study is the long turn ③ from 500 m to 1500 m. The track elevation increases by approximately
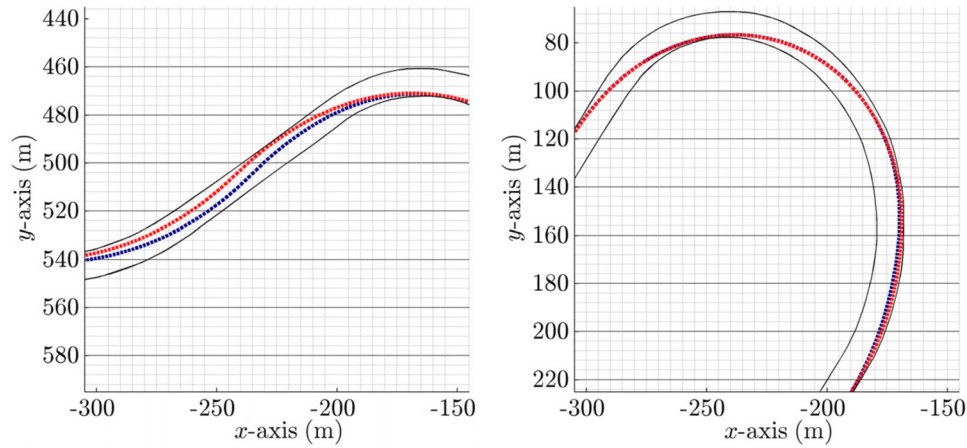
Fig. 14   Racing line on a flat (red) and a 3D (blue) track. The left-hand diagram shows turns ② and ③, while the right-hand diagram is for turn ④.
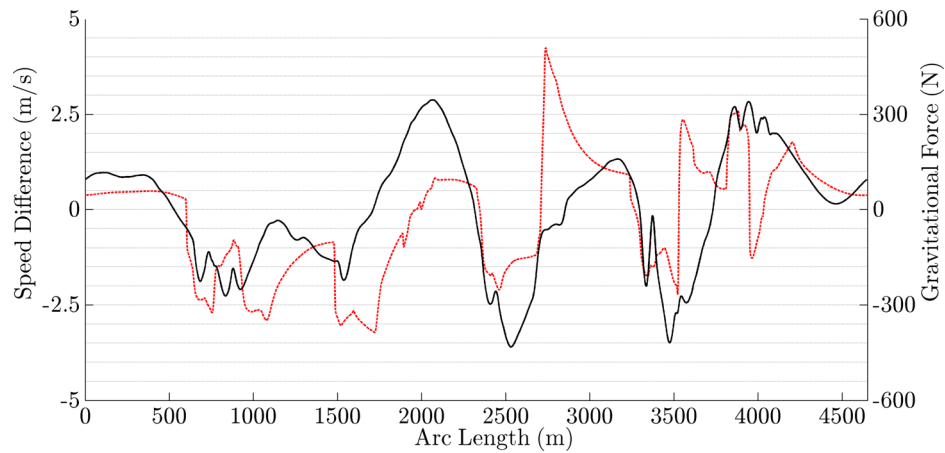

Fig. 15   Speed difference dashed (red) and the gravitational force component in the car's *x*-direction solid (black). The speed difference is positive when the car is faster on the 3D track. The car is traveling downhill when the gravitational force component is positive.
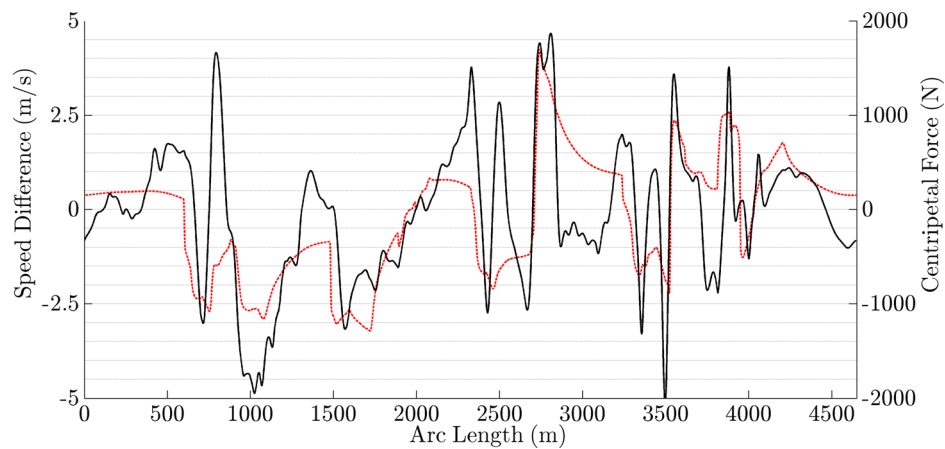

Fig. 16   Speed difference dashed (red) and the centripetal force in the car's *z*-direction solid (black). The speed difference is positive when the car is faster on the 3D track. The centripetal force provides additional down force when it is positive.

18 m over this section and so lower speeds can be expected. In addition to this, there are significant stretches of adverse camber as well as a significant down-force reduction due to centripetal forces (see Fig. 16). Track three dimensionality therefore causes the car to loose approximately 0.4 s on this section of track.

As is evident from Fig. 12, there are significant differences between the 3D and flat-road car speeds between turns ⑦ and ⑩. These speed differences are both positive and negative over this section. Before analysing the reasons for these variabilities, it is worth recalling that $\Delta v \propto \int \Delta F dt$, and so speed changes are a lagging indicator of driving force changes. It is clear that the car on the 3D track performs relatively poorly on the exit of turn ⑦. Figure 15 shows that the track has a positive gradient here, which is a first reason for the slower exit speed from turn ⑦. As can be seen in Fig. 10, the track also has an adverse camber over this section—a further reason for a reduced exit speed from turn ⑦. As is shown in Fig. 16, the car is also subject to a reduced down force on the exit of this corner due to centripetal force effects. This reduced exit speed is carried all the way to turn ⑨. Fortunes are reversed at the exit of turn ⑨, where gravitational effects appear to be neutral. Figure 16 shows that significantly improved centripetal down force occurs for approximately 100 m on the exit of the corner. This is also a section of sustained positive camber; see Fig. 10. This combination of advantageous circumstances leads to an increased exit speed from turn ⑨, which is then carried to turn ⑩, thereby reversing the earlier setback.

5 Conclusion

Prior work on the minimum-lap-time optimal control of a Formula One car has been generalized to the case of 3D tracks. In Part 1, an optimal estimation process is presented that allows the geodesic and normal curvatures, and the relative torsion of a race track to be identified from noisy measurement data. This curvature information is assumed known for the purposes of this paper. The dynamics of the car are modeled on the assumption that the vehicle moves on a tangent plane that is orthogonal to the principal normal vector m associated with the Darboux frame. Since the tangent plane under the car can pitch, roll and heave, several 3D dynamic influences result that include: (i) a gravitational force vector that moves in the car's reference frame. This motion produces lateral and longitudinal gravitational force components that can either supplement, or oppose the engine power, and that can either augment, or undermine the tires' side-force capability. (ii) The rolling road produces new centripetal forces that act along all three of the car's body-fixed axes. The most important influence appears to be the large vertical component that can either augment, or oppose the aerodynamic down force. (iii) Road camber influences that can either increase or reduce the car's maximum cornering speed; and (iv) road-related gyroscopic moments that affect the lateral and longitudinal load balances. Under most circumstances these effects are thought to be small. As an illustrative example we study the minimum-lap-time performance of a Formula One car on the Barcelona circuit. The study shows that 3D effects can significantly alter the car's maximum speed. The effect of 3D track features will vary considerably between tracks. Given that 3D influences are relatively easy to include in lap-time simulation studies, this source of error can now be addressed in an elegant and straightforward manner.

Appendix A

Vehicle and Tire Data

This Appendix contains nominal values for the tire and vehicle parameters used in this study. The air density is assumed to be 1.2 kg/m^3 (Table 1).

Table 1 Vehicle parameters

Symbol	Description	Value
P_{max}	Peak engine power	560 kW
M	Vehicle mass	660 kg
I_x	Moment of inertia about the x-axis	112.5 kg/m^2
I_y	Moment of inertia about the y-axis	450 kg/m^2
I_z	Moment of inertia about the z-axis	450 kg/m^2
W	Wheelbase	3.4 m
A	Distance of the mass center from the front axle	1.8 m
B	Distance of the mass center from the rear axle	$w - a$
H	Center of mass height	0.3 m
A	Frontal area	1.5 m^2
D_{roll}	Roll moment distribution (fraction at the front axle)	0.5
w_f	Front wheel to car centerline distance	0.73 m
w_r	Rear wheel to car centerline distance	0.73 m
R	Wheel radius	0.33 m
k_d	Differential friction coefficient	10.47 Nm s/rad

Appendix B

Tire Friction

The tire frictional forces are modeled using empirical formulae that are responsive to the tire normal loads and the combined slip. The tire's longitudinal slip is described by a longitudinal slip coefficient κ, while the lateral slip is described by a slip angle α [25]. Following standard conventions, we use

$$\kappa = -\left(1 + \frac{R\omega_w}{u_w}\right) \tag{B1}$$

$$\tan\alpha = -\frac{v_w}{u_w} \tag{B2}$$

where R is the wheel radius and ω_w the wheel's spin angular velocity. The terms u_w and v_w are the absolute speed components of the wheel center in a wheel-fixed coordinate system. The following formulae give the peak values and locations of the lateral and longitudinal friction coefficients using linear interpolation [26]

$$\mu_{x\,max} = (F_z - F_{z1})\frac{\mu_{x\,max\,2} - \mu_{x\,max\,1}}{F_{z2} - F_{z1}} + \mu_{x\,max\,1} \tag{B3}$$

$$\mu_{y\,max} = (F_z - F_{z1})\frac{\mu_{y\,max\,2} - \mu_{y\,max\,1}}{F_{z2} - F_{z1}} + \mu_{y\,max\,1} \tag{B4}$$

$$\kappa_{max} = (F_z - F_{z1})\frac{\kappa_{max\,2} - \kappa_{max\,1}}{F_{z2} - F_{z1}} + \kappa_{max\,1} \tag{B5}$$

$$\alpha_{max} = (F_z - F_{z1})\frac{\alpha_{max\,2} - \alpha_{max\,1}}{F_{z2} - F_{z1}} + \alpha_{max\,1} \tag{B6}$$

where the quantities containing a "1" or a "2" in the subscript are measured tire parameters (Table 2). Next, the tire slip is normalized with respect to the peak slip

$$\kappa_n = \kappa/\kappa_{max} \tag{B7}$$

$$\alpha_n = \alpha/\alpha_{max} \tag{B8}$$

Table 2 Tire parameters

Symbol	Description	Value
F_{z1}	Reference load 1	2000 N
F_{z2}	Reference load 2	6000 N
μ_{x1}	Peak longitudinal friction coefficient at load 1	1.75
μ_{x2}	Peak longitudinal friction coefficient at load 2	1.40
κ_1	Slip coefficient for the friction peak at load 1	0.11
κ_2	Slip coefficient for the friction peak at load 2	0.10
μ_{y1}	Peak lateral friction coefficient at load 1	1.80
μ_{y2}	Peak lateral friction coefficient at load 2	1.45
α_1	Slip angle for the friction peak at load 1	9 deg
α_2	Slip angle for the friction peak at load 2	8 deg
Q_x	Longitudinal shape factor	1.9
Q_y	Lateral shape factor	1.9

Following normalization, the slip is characterized by a combined-slip coefficient

$$\rho = \sqrt{\alpha_n^2 + \kappa_n^2} \tag{B9}$$

The friction coefficients in the longitudinal and lateral directions are described by

$$\mu_x = \mu_{x\,\max} \sin\left(Q_x \arctan(S_x \rho)\right) \tag{B10}$$

$$\mu_y = \mu_{y\,\max} \sin\left(Q_y \arctan(S_y \rho)\right) \tag{B11}$$

with

$$S_x = \frac{\pi}{2 \arctan(Q_x)} \tag{B12}$$

$$S_y = \frac{\pi}{2 \arctan(Q_y)} \tag{B13}$$

Finally, the longitudinal and lateral components of the tire forces are given by

$$F_x = \mu_x F_z \frac{\kappa_n}{\rho} \tag{B14}$$

$$F_y = \mu_y F_z \frac{\alpha_n}{\rho} \tag{B15}$$

In the car model used here the normal loads are determined by solving the load balance equations given in Secs. 2.3 and 2.3.1. The four wheels slip angles are given by

$$
\begin{aligned}
\alpha_{rr} &= \arctan\left(\frac{v - \bar{\omega}_z b}{u - \bar{\omega}_z w_r}\right) \\
\alpha_{rl} &= \arctan\left(\frac{v - \bar{\omega}_z b}{u + \bar{\omega}_z w_r}\right) \\
\alpha_{fr} &= \arctan\left(\frac{\sin\delta(\bar{\omega}_z w_f - u) + \cos\delta(\bar{\omega}_z a + v)}{\cos\delta(u - \bar{\omega}_z w_f) + \sin\delta(\bar{\omega}_z a + v)}\right) \\
\alpha_{fl} &= \arctan\left(\frac{\cos\delta(\bar{\omega}_z a + v) - \sin\delta(\bar{\omega}_z w_f + u)}{\cos\delta(\bar{\omega}_z w_f + u) + \sin\delta(\bar{\omega}_z a + v)}\right)
\end{aligned}
\tag{B16}
$$

with the longitudinal slip coefficients given by

$$
\begin{aligned}
\kappa_{rr} &= -\left(1 + \frac{R\omega_{rr}}{u - \bar{\omega}_z w_r}\right) \\
\kappa_{rl} &= -\left(1 + \frac{R\omega_{rl}}{u + \bar{\omega}_z w_r}\right) \\
\kappa_{fr} &= -\left(1 + \frac{R\omega_{fr}}{\cos\delta(u - \bar{\omega}_z w_f) + \sin\delta(\bar{\omega}_z a + v)}\right) \\
\kappa_{fl} &= -\left(1 + \frac{R\omega_{fl}}{\cos\delta(u + \bar{\omega}_z w_f) + \sin\delta(\bar{\omega}_z a + v)}\right)
\end{aligned}
\tag{B17}
$$

References

[1] Sharp, R. S., and Peng, H., 2011, "Vehicle Dynamics Applications of Optimal Control Theory," Veh. Syst. Dyn., **49**(7), pp. 1073–1111.

[2] Perantoni, G., and Limebeer, D. J. N., 2014, "Optimal Control for a Formula One Car With Variable Parameters," Veh. Syst. Dyn., **52**(5), pp. 653–678.

[3] Limebeer, D. J. N., Perantoni, G., and Rao, A. V., 2014, "Optimal Control of Formula One Car Energy Recovery Systems," Int. J. Control, **87**(10), pp. 2065–2080.

[4] Lanczos, C., 1986, *The Variational Principles of Mechanics*, 4th ed., Dover Publications, New York.

[5] Bertsekas, D. P., 1999, *Nonlinear Programming*, 2nd ed., Athena Scientific, Belmont, MA.

[6] Bliss, G. A., 1918, "The Problem of Mayer With Variable End Points," Trans. Am. Math. Soc., **19**, pp. 305–314.

[7] Bryson, A. E. J., and Ho, Y. C., 1975, *Applied Optimal Control: Optimization, Estimation, and Control*, CRC Press, Taylor & Francis Group, London, UK.

[8] Kirk, D. E., 2004, *Optimal Control Theory: An Introduction*, Dover Publication, New York.

[9] Sage, A. P., and White, I. C., 1977, *Optimum Systems Control*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

[10] Patterson, M. A., and Rao, A. V., 2013, GPOPS-II Version 1.0: *A General-Purpose MATLAB Toolbox for Solving Optimal Control Problems Using the Radau Pseudospectral Method*, University of Florida, Gainesville, FL.

[11] Davis, P. J., and Rabinowitz, P., 1984, *Methods of Numerical Integration*, Academic Press, Orlando, FL.

[12] Garg, D., Patterson, M. A., Hager, W. W., Rao, A. V., Benson, A. V., and Huntington, G. T., 2010, "A Unified Framework for the Numerical Solution of Optimal Control Problems Using Pseudospectral Methods," Automatica, **46**(11), pp. 1843–1851.

[13] Kameswaran, S., and Biegler, L. T., 2008, "Convergence Rates for the Direct Transcription of Optimal Control Problems Using Collocation and Radau Points," Comput. Optim. Appl., **41**(1), pp. 81–126.

[14] Hou, H., 2013, "Convergence Analysis of Orthogonal Collocation Methods for Unconstrained Optimal Control," Ph.D. dissertation, University of Florida, Gainesville, FL.

[15] Ruths, J., Zlotnik, A., and Li, J. S., 2011, "Convergence of a Pseudospectral Method for Optimal Control of Complex Dynamical Systems," 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, Dec. 12–15, pp. 5553–5558.

[16] Kelley, H. J., 1964, "A Second Variation Test for Singular Extremals," AIAA J., **2**(8), pp. 1380–1382.

[17] Goh, B. S., 1966, "Necessary Conditions for Singular Extremals Involving Multiple Control Variables," SIAM J. Control, **4**(4), pp. 716–731.

[18] Robbins, H. M., 1967, "A Generalized Legendre–Clebsch Condition for the Singular Cases of Optimal Control," IBM J. Res. Dev., **11**(4), pp. 361–372.

[19] Krenner, A. J., 1977, "The High Order Maximal Principle and Its Application to Singular Extremals," SIAM J. Control, **15**(2), pp. 256–293.

[20] Squire, W., and Trapp, G., 1998, "Using Complex Variables to Estimate Derivatives of Real Functions," SIAM Rev., **40**(1), pp. 110–112.

[21] Martins, J. R. R., Sturdza, P., and Alonso, J. J., 2003, "The Complex-Step Derivative Approximation," ACM Trans. Math. Software, **29**(3), pp. 245–262.

[22] Lantoine, G., Russell, R. P., and Dargent, T., 2012, "Using Multicomplex Variables for Automatic Computation of High-Order Derivatives," ACM Trans. Math. Software, **38**(3), pp. 16:1–16:21.

[23] Abramowitz, M., and Stegun, I. A., 1965, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York.

[24] Darby, C. L., Hager, W. W., and Rao, A. V., 2011, "An *hp*-Adaptive Pseudospectral Method for Solving Optimal Control Problems," Optim. Control Appl. Methods, **32**(4), pp. 476–502.

[25] Pacejka, H. B., 2008, *Tyre and Vehicle Dynamics*, 2nd ed., Butterworth-Heinemann, Oxford, UK.

[26] Kelly, D. P., 2008, "Lap Time Simulation With Transient Vehicle and Tyre Dynamics," Ph.D. thesis, Cranfield University School of Engineering, Bedford, UK.