Numer. Math. (2010) 114:653–695
DOI 10.1007/s00211-009-0262-2

Numerische
Mathematik

Optimal control of systems with discontinuous differential equations

David E. Stewart · Mihai Anitescu

Received: 5 June 2005 / Revised: 11 May 2009 / Published online: 25 September 2009

Abstract In this paper we discuss the problem of verifying and computing optimal controls of systems whose dynamics is governed by differential systems with a discontinuous right-hand side. In our work, we are motivated by optimal control of mechanical systems with Coulomb friction, which exhibit such a right-hand side. Notwithstanding the impressive development of nonsmooth and set-valued analysis, these systems have not been closely studied either computationally or analytically. We show that even when the solution crosses and does not stay on the discontinuity, differentiating the results of a simulation gives gradients that have errors of a size independent of the stepsize. This means that the strategy of "optimize the discretization" will usually fail for problems of this kind. We approximate the discontinuous right-hand side for the differential equations or inclusions by a smooth right-hand side. For these smoothed approximations, we show that the resulting gradients approach the true gradients provided that the start and end points of the trajectory do not lie on the discontinuity and that Euler's method is used where the step size is "sufficiently small" in comparison with the smoothing parameter. Numerical results are presented

The submitted manuscript has been created by the University of Chicago as Operator of Argonne National Laboratory ("Argonne") under Contract No. W-31-109-ENG-38 with the U.S. Department of Energy. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

D. E. Stewart (✉)
Department of Mathematics, University of Iowa,
Iowa City, IA 52242, USA
e-mail: dstewart@math.uiowa.edu

M. Anitescu
Mathematics and Computer Science Division, Argonne National Laboratory,
Argonne, IL, USA

for a crude model of car racing that involves Coulomb friction and slip showing that this approach is practical and can handle problems of moderate complexity.

Mathematics Subject Classification (2000) Primary: 49Q12; Secondary: 34A36 · 49J24

1 Introduction

Consider a block on a table subject to Coulomb friction on the contacting surface, pulled by a force $g(t)$ [11,28,33] (Fig. 1): The differential equation for this system is

$$m\frac{dv}{dt} \in -\mu N \operatorname{Sgn}(v) + g(t), \tag{1}$$

where Sgn is a *set-valued* function given by

$$\operatorname{Sgn}(z) = \begin{cases} \{+1\}, & z > 0, \\ [-1,+1], & z = 0, \\ \{-1\}, & z < 0. \end{cases} \tag{2}$$

The quantity N is the normal contact force ($= mg$ for a block of mass m) and μ the coefficient of Coulomb friction.

A differential inclusion

$$\frac{dx}{dt} \in F(x), \quad x(0) = x_0 \tag{3}$$

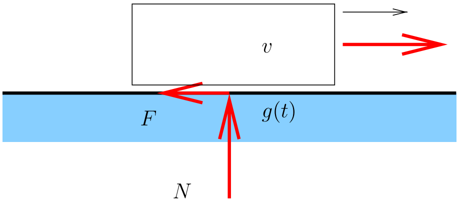
with F and x_0 given has unique solutions if F satisfies a one-sided Lipschitz condition: there is a constant $L \geq 0$ where

$$y_i \in F(x_i) \quad \text{for } i = 1, 2 \quad \text{implies} \quad (y_2 - y_1)^T (x_2 - x_1) \leq L\, \|x_2 - x_1\|^2. \tag{4}$$

Note that $-\operatorname{Sgn}$ satisfies this one-sided Lipschitz condition with $L = 0$. We write the solution as $x(t; x_0)$. The solution operator $x_0 \mapsto x(t; x_0)$ is Lipschitz with Lipschitz constant e^{Lt} [4].

Prior work has been done on theoretical aspects of nonsmooth optimal control problems, including [6,7,14–16]. However, none of this work deals with discontinuous dynamics. The work of Clarke [6,7] deals with nonsmooth but Lipschitz dynamics

Fig. 1 Block sliding on a table



and objective functions, while that of Frankowska [14–16] deals with set-valued but Lipschitz dynamics. Both approaches develop a *maximum principle* generalizing the well-known Pontryagin maximum principle [27] for optimal control.

Thus, this previous work of Clarke, Frankowska, and others cannot be directly applied to systems that have Coulomb friction. One approach which could be applied is that of Sussmann [31]. This approach requires computation of Jacobian matrix of the flow map $\Phi_{s,t}\colon x(s) \mapsto x(t)$. Since $\Phi_{s,t}$ is smooth *provided neither* $x(s)$ *nor* $x(t)$ *is on the discontinuity*, a modification of this approach could provide a suitable analog of the Pontryagin conditions. However, the Jacobian matrix of the flow map for problems with Coulomb friction is not easy to compute, and standard numerical approaches fail, as we show below. The examples below also show that the strategy of "optimizing the discretization" is unlikely to work for such problems.

Numerical work on optimizing systems with dynamics like (1) includes [9,20,36, 38]. Of these, Glowinski and Kearsley [20] used pattern search to carry out the optimization. Driessen and Sadegh [9] set up the entire dynamics as a mixed integer–linear program after using a standard time-discretization. The integer variables were used to represent the values of the "Sgn" function at each time-step. Ventura and Martinez [38] used a hybrid neural network/evolutionary computation approach to computing optimal controls. Van Willigenburg and Loop [36] used adjoint equations to compute gradients so that a conventional constrained optimization routine could be applied (BCPOL from IMSL in this paper). However, there is reason to believe that the adjoint functions computed by Van Willigenburg and Loop are not, in fact, correct, as the authors did not take into account that the discontinuity in the right-hand side (1) causes a discontinuity in the adjoint functions. This phenomenon of discontinuous adjoint functions has been noticed by Driessen and Sadegh [9] and is discussed in depth below.

We mention two examples of analytical investigation of optimal control problems with Coulomb friction. The first is the work of Lipp on the brachistochrone problem with Coulomb friction [24], although the slip is assumed to always be in a fixed direction so that the dynamics is continuous, although not smooth. The second is the work of Kim and Ha [23], who investigate a specific two-dimensional problem with Coulomb friction and find, for their simple system, that the adjoint variables have a jump; they compute the size of that jump.

There has been some success with optimizing *static* systems with Coulomb friction. In particular, Outrata et al. [26, Ch. 11] discuss using a bundle method of Lemarechal to optimize the friction coefficients for a contact problem.

As can be noticed in all of the above examples for optimal control of (1), gradient information either is not used or is probably incorrect.

Some authors have considered the problem of computing correct parametric sensitivities. The work of Barton et al. [18,35], for example, develops a "jump formula" for the sensitivities as the trajectory crosses a discontinuity. However, our work is different in the three important ways from that work.

1. The models considered by Barton et al. implicitly assume that the trajectory does not stay on the discontinuity for a positive length of time. This is commonly not

true for discontinuous systems arising Coulomb friction. We analyze such systems in depth in Sect. 6.

2. The same references contain the observation, which we also emphasize here, that in the case of numerical simulation, the derivatives are not computed correctly if the switching time is not accurately identified. However, we take this observation further in the context of optimal control, by showing that systems of the type described here whose derivative is computed by a fixed-step time-stepping procedure may exhibit local minima that accumulate to arbitrary points in a neighborhood of the actual minimum.
3. The models considered by Barton et al. also refer to differential algebraic equation that are index one on the smooth portions, whereas the differential algebraic equations that are equivalent to our model are index two.

Furthermore, in this paper we show that adjoints computed by smoothing the right-hand side of the differential equation will converge to the true adjoints, satisfying the relevant "jump conditions".

1.1 Organization of the paper

In Sect. 2 we look at the "optimize the discretization" strategy and show that it fails for problems of the same kind as (1) whether explicit, implicit, or partly implicit time-discretizations are used. In Sect. 3 we present the model differential inclusion and we discuss the implications of the one-sided Lipschitz Assumption. In Sect. 4 a smoothing approach is introduced, and some general properties of this approach are developed. This class of systems contains systems of type (1). As a result we develop a rule for computing the jumps in the adjoint functions for systems of this type. In Sects. 5 and 6 we show that provided the step-size goes to zero faster than the smoothing parameter, then the gradients and adjoints computed for Euler's method converge to the exact gradients for the discontinuous system. In Sect. 7 a crude model of a racing car is developed involving Coulomb friction, which is used as a test model. Numerical results are obtained via a smoothing approach that shows the practicality of the approach for a problem of moderate complexity.

1.2 Notation

Regarding the notation for gradients and Jacobians, most vectors are considered to be column vectors unless otherwise specified. For a function $f : \mathbb{R}^m \to \mathbb{R}^n$, $\nabla f(x)$ is an $n \times m$ matrix so $(\nabla f(x))_{ij} = \partial f_i / \partial x_j(x)$. This means that for scalar functions ($n = 1$), $\nabla f(x)$ is a *row* vector. However, if f is a function of one variable ($m = 1$), $\nabla f(x) = f'(x)$ is a column vector. This means that $f(x + \Delta x) = f(x) + \nabla f(x)\Delta x + o(\|\Delta x\|)$ for any differentiable f, regardless of whether f is scalar- or vector-valued. Where there are several inputs, we use a subscript on "∇" to indicate with respect to which variable the gradient is taken.

We use $\dot{x}$ to denote the derivative dx/dt, although sometimes the latter notation is used where it is clearer.

Note that we use C to denote a quantity that depends only on the data of the problem (that is, it does not depend on the other parameters introduced, such as the smoothing parameter σ, the step size h, the time t, or the step number k). These quantities can differ on each appearance. Since we use asymptotic notation, we remind the reader that $f(s) = O(g(s))$ (as $s \downarrow 0$) means that there are constants $C > 0$ and $s_0 > 0$ where for $0 < s < s_0$, $|f(s)| \leq C\, g(s)$. Also $f(s) = o(g(s))$ means that $\lim_{s\downarrow 0} f(s)/g(s) = 0$. Furthermore, $f(s) = \Omega(g(s))$ means that there are $C > 0$ and $s_0 > 0$ where for $0 < s < s_0$, $f(s) \geq C\,|g(s)|$, and $f(s) = \omega(g(s))$ means that $\lim_{s\downarrow 0} g(s)/f(s) = 0$.

Finally, we introduce the notation for the *tangent cone* of a set K at $u \in K$ to be

$$\mathcal{T}_K(u) = \left\{ \lim_{j\to\infty} \frac{u_j - u}{t_j} \mid u_j \in K, \quad t_j \downarrow 0 \text{ as } j \to \infty \right\}.$$

2 “Optimize the discretization” strategy

Consider first the simple differential inclusion

$$\frac{dx}{dt} \in -\mathrm{Sgn}(x), \quad x(0) = 1. \tag{5}$$

The exact solution is unique and is easily checked to be $x(t) = (1-t)_+$ where $z_+ = \max(z, 0)$ is the positive part of z. We can discretize this equation using the explicit Euler method or a partially explicit Euler method. If we set $t_k = t_0 + k\,h$ where $h > 0$ is the time step and x^k is our approximation to $x(t_k)$, the discrete-time trajectories for $dx/dt \in F(x)$ will satisfy

$$x^{k+1} \in x^k + h\, F(x^k + \chi(x^{k+1} - x^k)). \tag{6}$$

The parameter $\chi \in [0, 1]$ indicates how implicit the method is: $\chi = 0$ corresponds to the explicit Euler method, $\chi = \frac{1}{2}$ corresponds to the mid-point rule, and $\chi = 1$ corresponds to the fully implicit Euler method [1]. Solutions of the discretized problem are known to converge to solutions of the continuous time differential inclusion (5) (see [32–34]). However, we will shortly see that even though the numerical trajectories converge ($x_h(t; x_0) \to x(t; x_0)$ as $h \downarrow 0$ where x_h is the numerical trajectory), the gradients do not ($\nabla_{x_0} x_h(t; x_0) \not\to \nabla_{x_0} x(t; x_0)$) even where $x(t; \cdot)$ is smooth.

Note that if $L\,h < 1$, then there is only one solution to (6).

Consider the differential inclusion

$$\frac{dx}{dt} \in (1+\alpha) - \mathrm{Sgn}(x), \quad x(0) = -1, \tag{7}$$

with $\alpha > 0$. The exact solution is $x(t) = -1 + (2+\alpha)t$ for $0 \leq t \leq 1/(2+\alpha)$, and $x(t) = \alpha(t - 1/(2+\alpha))$ for $t \geq 1/(2+/\alpha)$. For $x(0) = x_0$ with $x_0 \approx -1$, the solution is nearly as simple: $x(t) = x_0 + (2+\alpha)t$ for $0 \leq -x_0/(2+\alpha)$, and $x(t) = \alpha(t + x_0/(2+\alpha))$ for $t \geq -x_0/(2+\alpha)$. This means that $\nabla_{x_0} x(2; x_0) = \alpha/(2+\alpha)$ at $x_0 = -1$.

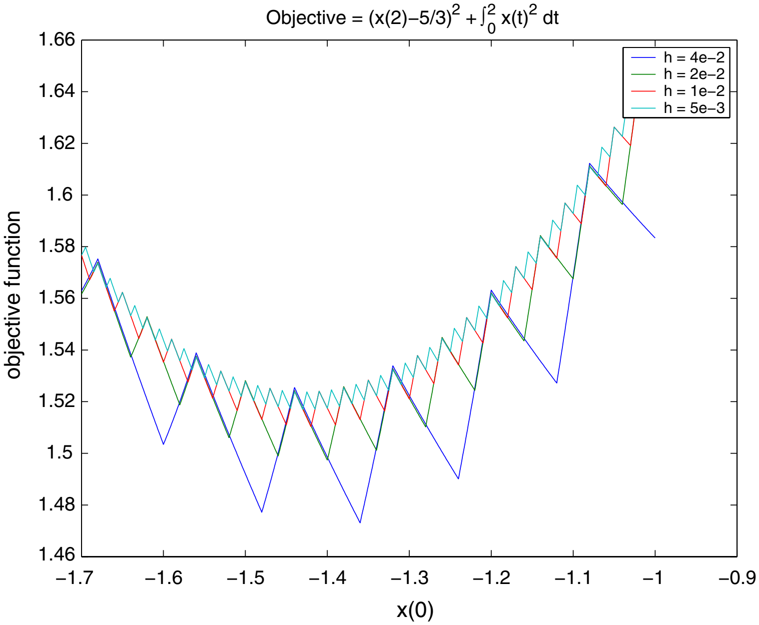


Fig. 2 Computed value of $(x(2) - 5/3)^2 + \int_0^2 x(t)^2\,dt$ against x_0 for various step-sizes

This differential inclusion should be easy to handle because it crosses the discontinuity, rather than staying on it as occurs in (5).

The discretization (6) for (7) is

$$x^{k+1} \in x^k + h\,(1+\alpha) - h\,\mathrm{Sgn}(x^k + \chi\,(x^{k+1} - x^k)). \tag{8}$$

If $x^k + \chi\,(x^{k+1} - x^k) < 0$, then $x^{k+1} = x^k + (2+\alpha)h$; if $x^k + \chi\,(x^{k+1} - x^k) > 0$, then $x^{k+1} = x^k + \alpha\,h$; if $x^k + \chi\,(x^{k+1} - x^k) = 0$, then $x^{k+1} = -((1-\chi)/\chi)x^k$. Inserting these formulas for x^{k+1} into the first two conditions gives the following: If $x^k + \chi(2+\alpha)h < 0$, then $x^{k+1} = x^k + (2+\alpha)h$; if $x^k + \chi\,\alpha\,h > 0$, then $x^{k+1} = x^k + \alpha\,h$. Neither of these occurs if $x^k \in -\chi\,h\,[\alpha,\ 2+\alpha]$, where $x^{k+1} = -((1-\chi)/\chi)x^k$. Note that $\nabla_{x^k} x^{k+1}$ is either $-(1-\chi)/\chi$ or one. If $\chi = 1$ (for fully implicit Euler), then $\nabla_{x^k} x^{k+1}$ is either zero or one. If $x_h(t)$ is the piecewise linear interpolant of $x_h(k\,h) = x^k$, then $\nabla_{x_0} x_h(2; x_0)$ computed from differentiating the numerical solutions of either zero or one for $h > 0$ sufficiently small.

Now consider $\frac{1}{2} < \chi < 1$. If $x^k \in -h\chi\,[\alpha,\ 2+\alpha]$, then $x^{k+1} = -(1-\chi)\,x^k/\chi > 0$, and so $x^{k+1} + \chi\,\alpha\,h > 0$ and $x^{k+2} = x^{k+1} + \alpha\,h \geq x^{k+1} > 0$, and so on. Thus there can be at most one k where $x^k \in -h\chi\,[\alpha,\ 2+\alpha]$. This means that the approximation to $\nabla_{x_0} x(2; x_0)$ obtained by differentiating the numerical solutions is either $-(1-\chi)/\chi$ or one. Both of these answers is clearly far from the correct answer of $\alpha/(2+\alpha)$.

As a more explicit example, consider the following results, which involve the above differential inclusion with $\alpha = 1$ and numerical solutions computed by using $\chi = 1$ (i.e., the fully implicit Euler method). In Fig. 2 the objective function is $(x(2) - 5/3)^2 +$

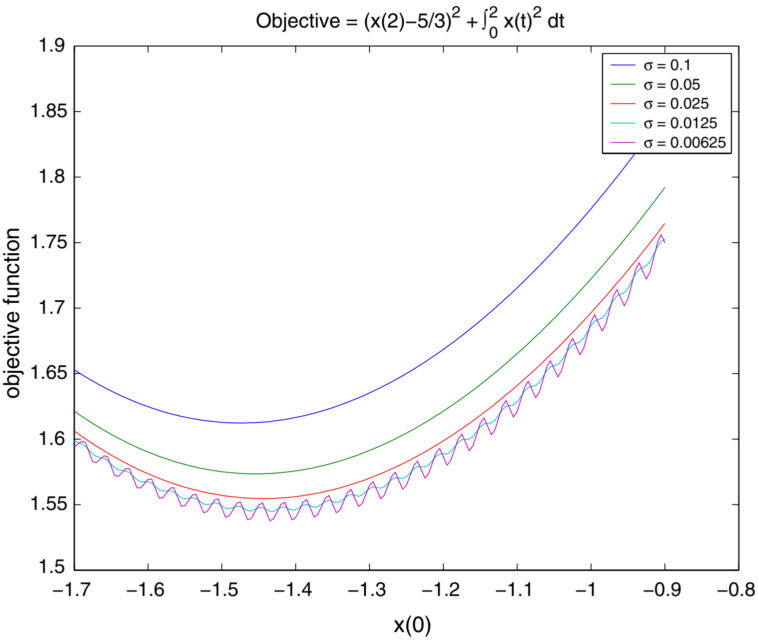


Fig. 3 Computed value of $(x(2) - 5/3)^2 + \int_0^2 x(t)^2\,dt$ against x_0 for various smoothing parameters ($h = 10^{-2}$)

$\int_0^2 x(t)^2\,dt$ is plotted against $x(0) = x_0$ with the integral computed by using the trapezoidal rule. The trapezoidal method should contribute only $O(h^2)$ error compared with the $O(h)$ bound for the errors from the implicit Euler method. As can be clearly seen in Fig. 2, the values of the computed objective function converge, but the gradients do not. Smoothing the right-hand side of the differential equation (in this case replacing $\mathrm{Sgn}(x)$ with $\mathrm{Sgn}_\sigma(x) := \tanh(x/\sigma)$) improves things greatly regarding the value of gradients, as can be seen in Fig. 3.

The results in Fig. 3 also point out the following fact about systems whose dynamics are not necessarily nonsmooth but are stiff enough to behave like an almost nonsmooth system, with $\sigma \neq 0$ but $\sigma \approx 0$. Since such systems are stiff, it is likely that the favorite way of simulating them is to use an implicit time-stepping scheme with a relatively large time step. If the derivative is computed on the same grid, the resulting optimization problem will have many local minima, not necessarily close to the target minimum. These disappear only when the time step is $o(1/L)$, where the right-hand side has a local Lipschitz constant of L.

Readers with a background in numerical analysis might wonder if higher-order schemes can reduce the requirement that $h = o(\sigma)$ for obtaining accurate gradients. Unfortunately, this does not appear to be the case as larger step-sizes result in $O(h)$ errors in the trajectory and $O(1)$ errors in the derivatives. This is related to the issue of finding numerical methods to solve a differential inclusion to an accuracy of $o(h)$. Methods of this kind can be found (see, e.g., [21,22]), but these will only give $o(h)$ accuracy if the trajectory does not leave or join the manifold on which the discontinuity lies. In either of these cases, using higher-order schemes does not obviate the condition that $h = o(\sigma)$.

3 The model differential inclusion

Consider the differential inclusion

$$\frac{dx}{dt} \in \begin{cases} \{f_1(x)\}, & \psi(x) < 0, \\ \{f_2(x)\}, & \psi(x) > 0, \\ \mathrm{co}\,\{f_1(x),\ f_2(x)\}, & \psi(x) = 0, \end{cases} \tag{9}$$

where “co (X)” is the convex hull of a set X.

We assume that this right-hand side satisfies a one-sided Lipschitz condition (4). We assume that $f_1,\ f_2\colon \mathbb{R}^n \to \mathbb{R}^n$ and $\psi\colon \mathbb{R}^n \to \mathbb{R}$ are C^2, and that $\nabla\psi(x) \neq 0$ whenever $\psi(x) = 0$.

We note that (9) can be in one of the following nondegenerate switching cases.

1. We have $\nabla\psi(x) \cdot f_1(x)$, $\nabla\psi(x) \cdot f_2(x) > 0$, whenever $\psi(x) = 0$. In this case $d\psi(x(t))/dt$ is strictly positive before and after the switching time. So the dynamical system described by (9) will switch from the set $\psi(x) < 0$ to the set $\psi(x) > 0$.
2. We have that $\nabla\psi(x) \cdot f_1(x) > 0$, $\nabla\psi(x) \cdot f_2(x) < 0$, whenever $\psi(x) = 0$. In this case once the dynamical system reaches the manifold $\psi(x) = 0$, it is trapped there.

Either of these cases will result in jumps in the sensitivities and the adjoint variables, as we show in the following sections. In addition we have the situation where the system “exits” a singularity, which may occur starting from the second case when $\nabla\psi(x) f_2(x)$ changes sign. In that case, however, there is no discontinuous transition, and the case does not need to be studied separately.

We address the properties of both cases below.

3.1 Consequences of the one-sided Lipschitz condition

In this subsection we show that the one-sided Lipschitz condition, which is satisfied by Coulomb-friction force laws, enforces certain constraints on the functions f_1 and f_2, which will be used later.

Lemma 1 *Suppose that the right-hand side of* (9) *satisfies the one-sided Lipschitz condition* (4)*. Suppose also that ψ is differentiable and $\nabla\psi(x) \neq 0$ for any x where $\psi(x) = 0$. Then on the discontinuity $\Sigma = \{\, x \mid \psi(x) = 0 \,\}$ we must have $(f_2 - f_1) \parallel \nabla\psi^T$ and $\nabla\psi \cdot (f_2 - f_1) \leq 0$.*

Proof Pick $\eta > 0, \epsilon > 0$, and $x \in \Sigma$. Consider any $0 \neq \zeta \perp \nabla\psi(x)^T$. Now put $x_1 = x - \epsilon\nabla\psi(x)^T$ and $x_2 = x + \epsilon\nabla\psi(x)^T + \eta\zeta$. We want to choose $\epsilon,\ \eta > 0$ small enough so that $\psi(x_1) < 0$ and $\psi(x_2) > 0$. Now $\psi(x_1) = \psi(x) - \epsilon\|\nabla\psi(x)\|^2 + o(\epsilon)$, so for any $\epsilon > 0$ sufficiently small, $\psi(x_1) < 0$. Also $\psi(x_2) = \psi(x) + \epsilon\|\nabla\psi(x)\|^2 + o(\epsilon + \eta)$.

So for any $\theta > 0$ there is an $\eta_0 > 0$ so that $\epsilon + \eta < \eta_0$ implies that the remainder term $o(\epsilon + \eta)$ is less than $\theta(\epsilon + \eta)$. For such ϵ and η,

$$\psi(x_2) \geq \epsilon \|\nabla\psi(x)\|^2 - \theta(\epsilon + \eta).$$

Provided $0 < \eta < (\|\nabla\psi(x)\|^2 - \theta)\epsilon/\theta$ and ϵ and η are sufficiently small, we have $\psi(x_2) > 0$. If we set $\eta = \frac{1}{2}(\|\nabla\psi(x)\|^2 - \theta)\epsilon/\theta$, then for sufficiently small $\epsilon > 0$, $\psi(x_2) > 0$. Turning this around, if we set $\epsilon = 2\theta\eta/(\|\nabla\psi(x)\|^2 - \theta)$, then for sufficiently small $\eta > 0$ and $0 < \theta \leq 1$, $\psi(x_1) < 0$ and $\psi(x_2) > 0$. Choose $\theta > 0$ sufficiently small so that $\epsilon \leq \eta$.

Now by the one-sided Lipschitz condition, for sufficiently small $\epsilon > 0$ and $\eta > 0$ given as above,

$$\begin{aligned} L\|x_2 - x_1\|^2 &\geq (x_2 - x_1)^T (f_2(x_2) - f_1(x_1)) \\ &= (2\epsilon\nabla\psi(x) + \eta\zeta^T)[f_2(x) - f_1(x) + O(\epsilon + \eta)]. \end{aligned} \tag{10}$$

Since $\|x_2 - x_1\| = O(\eta)$, after dividing both sides of (10) by η and taking $\eta \downarrow 0$, we have

$$0 \geq (2(\epsilon/\eta)\nabla\psi(x) + \zeta^T)[f_2(x) - f_1(x)]. \tag{11}$$

But $\epsilon/\eta = 2\theta/(\|\nabla\psi(x)\|^2 - \theta)$. So

$$0 \geq (2\theta/(\|\nabla\psi(x)\|^2 - \theta)\nabla\psi(x) + \zeta^T)[f_2(x) - f_1(x)],$$

for all $\theta > 0$ sufficiently small. Taking $\theta \downarrow 0$ gives the result that for any $x \in \Sigma$, $\zeta^T[f_2(x) - f_1(x)] \leq 0$. Since $-\zeta$ is also perpendicular to $\nabla\psi(x)^T$, it follows that $f_2(x) - f_1(x)$ is perpendicular to ζ. Noting that ζ is an arbitrary vector perpendicular to $\nabla\psi(x)^T$, we see that $f_2(x) - f_1(x) \parallel \nabla\psi(x)^T$.

To prove the final assertion, set $\zeta = 0$. Then for sufficiently small $\epsilon > 0$, $2\epsilon\nabla\psi(x)$ $[f_2(x) - f_1(x) + O(\epsilon)] \leq 0$. Taking $\epsilon \downarrow 0$ gives $\nabla\psi(x)[f_2(x) - f_1(x)] \leq 0$. □

3.2 The equivalent equation in the trapped case

We now work under the assumption that $\nabla\psi(x) f_1(x) > 0$ and $\nabla\psi(x) f_2(x) < 0$, that is, when the problem is "trapped" in the manifold $\psi(x) = 0$, where we must follow the trajectory.

From the third branch in (9), the problem of following the trajectory in the discontinuity should be expressed as

$$\frac{dx}{dt} = f^*(x) = (1 - \theta(x))\, f_1(x) + \theta(x) f_2(x).$$

The unknown weighting function $\theta(x)$ is computable from the condition that the mapping $\psi(x)$ is an invariant of the dynamical system defined by $f^*(x)$, that is,

$\nabla\psi(x)f^*(x) = 0$. In turn, the last equation leads to

$$\nabla\psi(x)\left((1-\theta(x))\,f_1(x) + \theta(x)f_2(x)\right) = 0.$$

Solving for the unknown weighting function from the last equation, we obtain

$$\theta(x) = \frac{\nabla\psi(x)f_1(x)}{\nabla\psi(x)f_1(x) - \nabla\psi(x)f_2(x)}. \tag{12}$$

Note that, under the assumption that $\nabla\psi(x)f_1(x) > 0$ and $\nabla\psi(x)f_2(x) < 0$ we must have that

$$0 < \theta(x) < 1,$$

at least in a neighborhood of the switching point.

This also shows that the dynamical system will exit the discontinuity manifold only when the weighing function $\theta(x)$ switches to either 0 or 1. From the expression of $\theta(x)$, it follows that such a switch can happen only when either $\nabla\psi(x)f_1(x)$ or $\nabla\psi(x)f_2(x)$ will switch signs.

With this identification, the dynamical system can, in effect, be represented by the following piecewise differential equation:

$$\dot{x} = \begin{cases} f_1(x), & \psi(x) < 0, \\ f^*(x), & \psi(x) \geq 0, \end{cases} \quad x(0) = x_0.$$

3.3 Exact sensitivities with respect to parameters

We are interested in evaluating the sensitivities with respect parameters. However, by making the parameter an extra dependent variable with zero derivative, we can make the sensitivity with respect to a parameter into sensitivity with respect to initial conditions. Thus we restrict ourselves to sensitivities with respect to the initial conditions. Let $x(t; x_0)$ be the value of $x(t)$ where x is the solution of the discontinuous ODE with initial value x_0.

3.3.1 The trapped case

We first compute these sensitivities in the case where the trajectory is trapped in the manifold $\psi(x) = 0$, that is, $\nabla\psi(x)f_1(x) > 0$ and $\nabla\psi(x)f_2(x) < 0$.

An important component in computing this sensitivities is the switching time $t_s = t_s(x_0)$, which can be defined implicitly by the equations

$$x(0; x_0) = x_0, \quad \dot{x}(t; x_0) = f_1(x), \quad \psi(x(t_s; x_0)) = 0.$$

The sensitivity $s(t, x_0)$ satisfies the following equation, before switching.

$$\dot{s} = \nabla f_1(x(t; x_0))s, \quad s(0) = I.$$

Using the implicit function theorem, we obtain that

$$\nabla_{x_0} t_s(x_0) = -\frac{\nabla\psi(x(t_s, x_0))s(t_s, x_0)}{\nabla\psi(x(t_s, x_0)) f_1\,(x(t_s, x_0))}.$$

From our assumptions, it is immediate that $\nabla\psi(x) f_1(x) \neq 0$.

To determine the equation satisfied by the sensitivity after switching, we proceed in two steps. First we consider the equation satisfied by the system once it enters the discontinuity,

$$\dot{y}(t, y_0) = f^*(y(t, y_0)), \quad y(0, y_0) = y_0,$$

and we analyze its sensitivity $s_2(t, y_0) = \nabla_{y_0} y(t; y_0)$ with respect to the parameter y_0. Here we have used the identification that is valid when $t > t_s$,

$$y(t; y_0) = x(t + t_s; x_0).$$

We obtain the following linear differential equation:

$$\dot{s}_2 = \nabla f^*(y(t; y_0))s_2, \quad s_2(0) = I.$$

To compute $s_2(t, y_0) = \nabla_{x_0} x(t; x_0)$, we glue the solutions before and after reaching to discontinuity by using that

$$y_0 = x(t_s(x_0); x_0).$$

We see that

$$\begin{aligned}
\nabla_{x_0} y_0 &= \nabla_{x_0}[x(t_s(x_0), x_0)] = f_1\,(x(t_s; x_0))\,\nabla t_s(x_0) + s(t_s, x_0) \\
&= -f_1\,(x(t_s; x_0))\,\frac{\nabla\psi(x(t_s; x_0))s(t_s, x_0)}{\nabla\psi(x(t_s; x_0))^T f_1\,(x(t_s; x_0))} + s(t_s, x_0) \\
&= \left[I - \frac{f_1\,(x(t_s; x_0))\,\nabla\psi(x(t_s; x_0))}{\nabla\psi(x(t_s; x_0)) f_1\,(x(t_s, x_0))}\right] s(t_s; x_0).
\end{aligned}$$

The following computation also shows that $x(t; x_0)$ is a differentiable function of x_0 and that its derivative $s(t)$ obeys the following differential equation

$$\dot{s}(t) = \begin{cases} \nabla f_1(x(t; x_0))s(t), & t < t_s, \\ \nabla f^*(x(t; x_0))s(t), & t > t_s. \end{cases} \tag{13}$$

To figure out the jump rule at the switching, we use that $x(t; x_0) = y(t - t_s; y_0)$.

We obtain that, whenever $t > t_s$, the following holds.

$$\nabla_{x_0} x(t; x_0) = \nabla_{x_0} y(t - t_s; y_0) = -f^*(y(t - t_s; y_0))\nabla_{x_0} t_s + s_2(t - t_s, y_0)\nabla_{x_0} y_0.$$

As $t \downarrow t_s$, we have that $s_2(t - t_s, y_0)$ approaches the identity. Using our computation for $\nabla_{x_0} t_s$ and $\nabla_{x_0} y_0$, we obtain that, at the switching point, the sensitivity will jump according to the rule

$$s(t_s^+) = \left[I + \frac{\left(f^*\left(x(t_s; x_0)\right) - f_1\left(x(t_s; x_0)\right)\right) \nabla\psi(x(t_s; x_0))}{\nabla\psi(x(t_s; x_0)) f_1\left(x(t_s; x_0)\right)} \right] s(t_s^-).$$

If we replace the expression for f^* in the above equation, we obtain that

$$s(t_s^+) = \left[I + \frac{\left(f_2\left(x(t_s; x_0)\right) - f_1\left(x(t_s; x_0)\right)\right) \nabla\psi(x(t_s; x_0))}{\nabla\psi(x(t_s; x_0)) \left(f_1\left(x(t_s; x_0)\right) - f_2\left(x(t_s; x_0)\right)\right)} \right] s(t_s^-). \tag{14}$$

From Lemma 1 we have that $(f_2 - f_1) \parallel \nabla\psi^T$, which implies that the matrix in the above relation is an orthogonal projection.

3.3.2 The case where $\nabla\psi(x) f_1(x) > 0$ and $\nabla\psi(x) f_2(x) > 0$

Using the arguments of the previous section, we note that (13) should be replaced by

$$\dot{s}(t) = \begin{cases} \nabla f_1(x(t; x_0))s(t), & t < t_s, \\ \nabla f_2(x(t; x_0))s(t), & t > t_s. \end{cases} \tag{15}$$

Repeating the preceding analysis for (15) we obtain

$$s(t_s^+) = \left[I + \frac{\left(f_2\left(x(t_s; x_0)\right) - f_1\left(x(t_s; x_0)\right)\right) \nabla\psi(x(t_s; x_0))}{\nabla\psi(x(t_s; x_0)) f_1\left(x(t_s; x_0)\right)} \right] s(t_s^-). \tag{16}$$

4 A smoothing approach

We now investigate the approximation of (9) by the smoothed system

$$\frac{dx_\sigma}{dt} = \varphi_\sigma(\psi(x_\sigma))\, f_2(x_\sigma) + (1 - \varphi_\sigma(\psi(x_\sigma)))\, f_1(x_\sigma). \tag{17}$$

Here

$$\varphi_\sigma(w) = \int_{-\infty}^{w} \rho_\sigma(r')\, dr', \tag{18}$$

where $\rho_\sigma(r) = (1/\sigma)\rho(r/\sigma)$ and $\rho \geq 0$, $\operatorname{supp} \rho = [-1, +1]$, and $\int_{-\infty}^{+\infty} \rho(r)\, dr = 1$. We could, for example, take $\rho(r) = c\, \exp(-1/(1+r) - 1/(1-r))$ for $-1 < r < +1$ and $\rho(r) = 0$ otherwise, with c chosen to give $\int_{-1}^{+1} \rho(r)\, dr = 1$. The smoothing function is then given by $\varphi_\sigma(w) = (1/\sigma)\int_{-\infty}^{w} \rho(r/\sigma)\, dr$. Note that there is no finite closed-form representation of φ_σ. The resulting smoothing function $\varphi_\sigma(w)$ is rather Thus, $\varphi_\sigma(w) = 0$ for $w \leq -\sigma$ and $\varphi_\sigma(w) = 1$ for $w \geq +\sigma$.

In practice it is not necessary that $\varphi_\sigma(w)$ is exactly zero for $w \leq -\sigma$ or that $\varphi_\sigma(w)$ is exactly one for $w \geq +\sigma$. However, it is important that $\varphi_\sigma(w)$ approaches 1 rapidly and smoothly as w/σ increases past one, and approaches 0 rapidly and smoothly as w/σ decreases below zero. Thus, we could use $\varphi_\sigma(w) = (1/2)[1 + \tanh(w/\sigma)]$ in practice. The assumption that $\varphi_\sigma(w) = 0$ for $w \leq -\sigma$ and $\varphi_\sigma(w) = 1$ for $w \geq +\sigma$ simplifies the analysis.

Readers may note that the results of the previous section can be understood as saying that gradients of the objective function with respect to parameters or changes in the initial conditions can be computed for this problem by identifying “break points” where the solution meets and leaves the discontinuity. This could be done using methods such as those in [29]. There are two reasons for dealing with the smoothing system. One is to show that using the smoothed system numerically can give accurate gradient information (provided at least that $h = o(\sigma)$). The other is so that the solution method for the (smoothed) differential equations can be incorporated into the constraints for using general purpose optimization software.

4.1 One-sided Lipschitz condition for the smoothed system

Lemma 1 is useful for showing that the smoothed right-hand side f_σ in (17) also satisfies a one-sided Lipschitz condition, although the Lipschitz constant might not be the same as for (9). To show this, we do need to assume that f_1 and f_2 satisfy an ordinary (“two-sided”) Lipschitz condition with constant L_f and that $\nabla\psi$ is also Lipschitz with constant $L_{\nabla\psi}$. As usual, L is the one-sided Lipschitz constant for (9). That means that both f_1 and f_2 satisfy the one-sided Lipschitz condition (4) with constant L. Since $\Sigma = \{\, x \mid \psi(x) = 0 \,\}$ is a C^1 manifold, there is a map $\pi(x) :=$ the nearest point in Σ to x, well-defined and continuous in a neighborhood of Σ. We can choose a $\sigma_0 > 0$ so that if $0 < \sigma < \sigma_0$, this map is well defined on the transition region.

We will show that for any $w \in \mathbb{R}^n$, $w^T \nabla f_\sigma(x) w \leq L\|w\|^2$. Note that $w^T \nabla f_1(x) w$, $w^T \nabla f_2(x) w \leq L\|w\|^2$ for all w. Outside the transition region we have $\nabla f_\sigma(x) = \nabla f_1(x)$ or $\nabla f_\sigma(x) = \nabla f_2(x)$, and the desired property of ∇f_σ follows immediately. Inside the transition region, we have

$$\nabla f_\sigma = (1 - \varphi_\sigma)\nabla f_1 + \varphi_\sigma \nabla f_2 + (f_2 - f_1)\varphi_\sigma'(\psi)\,\nabla\psi.$$

Thus

$$\begin{aligned} w^T \nabla f_\sigma(x) w &= (1 - \varphi_\sigma(\psi(x)) w^T \nabla f_1(x) w + \varphi_\sigma(\psi(x)) w^T \nabla f_2(x) w \\ &\quad + \varphi_\sigma'(\psi(x)) w^T (f_2(x) - f_1(x)) \nabla\psi(x) w \\ &\leq L\|w\|^2 + \varphi_\sigma'(\psi(x)) w^T (f_2(\pi(x)) - f_1(\pi(x))) \nabla\psi(\pi(x)) w \\ &\quad + O(1/\sigma)\, O(\|\pi(x) - x\|)\, \|w\|^2. \end{aligned}$$

Now from Lemma 1, $(f_2(\pi(x)) - f_1(\pi(x)))\nabla\psi(\pi(x))$ is a negative semi-definite matrix, so $w^T(f_2(\pi(x)) - f_1(\pi(x)))\nabla\psi(\pi(x))w \leq 0$. Also, the transition region is only $O(\sigma)$ wide, so $\|\pi(x) - x\| = O(\sigma)$. Thus, using Lemma 1, and the fact that $\varphi'_\sigma(\cdot) \geq 0$, we obtain that

$$w^T\nabla f_\sigma(x)w \leq O(1)\,\|w\|^2.$$

Thus f_σ satisfies a one-sided Lipschitz condition, although its one-sided Lipschitz constant may be considerably larger than for f.

5 Convergence of the gradients for the case $\nabla\psi(x)f_1(x) > 0$ and $\nabla\psi(x)f_2(x) > 0$

We now analyze the asymptotic properties, as $\sigma \to 0$ for the case where $\nabla\psi(x)$ $f_1(x) > 0$ and $\nabla\psi(x)f_2(x) > 0$, that is, the case where the trajectory switches from $\psi(x) < 0$ to $\psi(x) > 0$.

5.1 The variational equation of the smoothed differential equation

The variational equation for the smoothed system can be easily written down:

$$\frac{ds_\sigma}{dt} = \left\{(1-\varphi_\sigma)\nabla f_1 + \varphi_\sigma\nabla f_2 + (f_2 - f_1)\varphi'_\sigma(\psi)\,\nabla\psi\right\} s_\sigma. \tag{19}$$

For smooth f_1, f_2, the first two terms $\varphi_\sigma\,\nabla f_2$ and $(1-\varphi_\sigma)\,\nabla f_1$ in the braces of Eq. (19) are bounded, but the last term $(f_1 - f_2)\varphi'_\sigma(\psi)\,\nabla\psi$ might not be bounded. Thus the limiting equation as $\sigma \downarrow 0$ for $\psi(x(t)) \neq 0$ becomes

$$\frac{ds}{dt} = \left\{\begin{matrix}\nabla f_1, & \psi(x) < 0\\ \nabla f_2, & \psi(x) > 0\end{matrix}\right\} s, \tag{20}$$

which is uniform on compact sets bounded away from the discontinuity $\Sigma = \{x \mid \psi(x) = 0\}$. However, this ignores what happens near $\psi(x(t^*)) = 0$. From (16) we have determined what happens for the original system (9); but to complete a proof of convergence of s_σ to s, we must also prove that the jumps match.

Suppose that the limiting solution (which is unique by the one-sided Lipschitz assumption) reaches the surface $\psi(x(t)) = 0$ at time $t = t^*$. By our assumptions that $\psi(x) = 0$ implies $\nabla\psi(x)\cdot f_1(x) > 0$ and $\nabla\psi(x)\cdot f_2(x) > 0$, there can be only one time $t = t^*$ where $\psi(x(t)) = 0$. Put $x^* = x(t^*)$, $f_1^* = f_1(x^*)$, $f_2^* = f_2(x^*)$, $\nabla\psi^* = \nabla\psi(x^*)$. Note that if $t \approx t^*$ and $\sigma \approx 0$, then $x_\sigma(t) \approx x^*$.

Of particular interest to us is the fact that the term $(f_1 - f_2)\varphi'_\sigma(\psi)\nabla\psi$ is unbounded. Although we expect that $\varphi'_\sigma(\psi) \neq 0$ only for a time interval of length $O(\sigma)$, $\varphi'_\sigma(\psi)$ has a magnitude of $O(1/\sigma)$. In the limit as $\sigma \downarrow 0$, this could correspond to a Dirac-δ function. This can be interpreted in the sense of [25]. Since the matrix $(f_2(x_\sigma(t)) - f_1(x_\sigma(t)))\,\nabla\psi(x_\sigma(t))) \to (f_2^* - f_1^*)\,\nabla\psi^*$ as $\sigma \downarrow 0$ in the relevant time interval(s) ($\varphi'_\sigma(x_\sigma(t)) \neq 0$), in the limit the effect of this term is to include a factor of the form

$$\exp\left(\alpha\,(f_2^* - f_1^*)\,\nabla\psi^*\right), \tag{21}$$

where α is the limit of $\int \varphi'_\sigma(\psi(x_\sigma(t)))\,dt$. We will show that this limit exists and will give a simple formula for it and the matrix exponential (21).

Now for $-\sigma \leq \psi(x_\sigma(t)) \leq +\sigma$ we have $\|x_\sigma(t) - x^*\| = O(\sigma)$. Then we can write

$$\begin{aligned}\frac{d}{dt}\psi(x_\sigma(t)) &= \nabla\psi(x_\sigma(t)) \cdot \dot{x}_\sigma(t)\\ &= \varphi_\sigma(\psi(x_\sigma(t)))\nabla\psi(x_\sigma(t)) \cdot f_2(x_\sigma(t))\\ &\quad +(1 - \varphi_\sigma(\psi(x_\sigma(t))))\nabla\psi(x_\sigma(t)) \cdot f_1(x_\sigma(t))\\ &= \varphi_\sigma(\psi(x_\sigma(t)))\nabla\psi^* \cdot f_2^* + (1 - \varphi_\sigma(\psi(x_\sigma(t)))\nabla\psi^* \cdot f_1^* + O(\sigma).\end{aligned}$$

Put $\gamma_i = \nabla\psi^* \cdot f_i^*, i = 1, 2$. Then we can write

$$\frac{d\psi}{dt} = \varphi_\sigma(x_\sigma(t))\gamma_2 + (1 - \varphi_\sigma(x_\sigma(t)))\gamma_1 + O(\sigma).$$

For sufficiently small $\sigma > 0$, we have $d\psi/dt > 0$, so we can use a change of variables.

Returning to the value of α in (21), we consider the integrals

$$\begin{aligned}\int_{-\infty}^{+\infty} \varphi'_\sigma(\psi(x_\sigma(t)))\,dt &= \int_{-\sigma}^{+\sigma} \frac{\varphi'_\sigma(\psi)}{\gamma_1 + \varphi_\sigma(\psi)(\gamma_2 - \gamma_1) + O(\sigma)}\,d\psi\\ &= \frac{1}{\gamma_2 - \gamma_1}\int_{\gamma_1}^{\gamma_2} \frac{dw}{w} + O(\sigma) \quad (\text{using } w = \gamma_1 + \varphi_\sigma(\psi)(\gamma_2 - \gamma_1))\\ &= \frac{\ln(\gamma_2/\gamma_1)}{\gamma_2 - \gamma_1} + O(\sigma).\end{aligned}$$

Thus we obtain the value for the limit of $\alpha = \ln(\gamma_2/\gamma_1)/(\gamma_2 - \gamma_1)$. To compute the matrix exponential (21), we resort to the series definition of the matrix exponential. To simplify notation, put $u = f_2^* - f_1^*$ and $v^T = \nabla\psi^*$. Then

$$\begin{aligned}\exp(\alpha u v^T) &= I + \sum_{k=1}^{\infty} \frac{1}{k!}\alpha^k (uv^T)^k = I + \sum_{k=1}^{\infty} \frac{1}{k!}\alpha^k (v^T u)^{k-1} uv^T\\ &= I + \frac{1}{v^T u}\sum_{k=1}^{\infty} \frac{1}{k!}(\alpha\, v^T u)^k\, uv^T = I + \frac{uv^T}{v^T u}\left[e^{\alpha\, v^T u} - 1\right].\end{aligned}$$

Substituting for u and v, we see that the limiting matrix exponential is

$$\begin{aligned}
&\exp(\alpha(f_2^* - f_1^*)\nabla\psi^*) \\
&\quad = I + \frac{(f_2^* - f_1^*)\nabla\psi^*}{\gamma_2 - \gamma_1}\left[\exp((\gamma_2 - \gamma_1)\ln(\gamma_2/\gamma_1)/(\gamma_2 - \gamma_1)) - 1\right] \\
&\quad = I + \frac{(f_2^* - f_1^*)\nabla\psi^*}{\gamma_2 - \gamma_1}\left[\frac{\gamma_2}{\gamma_1} - 1\right] = I + \frac{(f_2^* - f_1^*)\nabla\psi^*}{\gamma_2 - \gamma_1}\frac{\gamma_2 - \gamma_1}{\gamma_1} \\
&\quad = I + \frac{(f_2^* - f_1^*)\nabla\psi^*}{\gamma_1}.
\end{aligned}$$

Thus taking $\sigma \downarrow 0$ we obtain the limiting result:

$$s(t^{*+}) = \left[I + \frac{(f_2^* - f_1^*)\nabla\psi^*}{\gamma_1}\right]s(t^{*-}). \qquad (22)$$

We have thus proved the following result.

Theorem 1 *For the case where $\psi(x) = 0$ implies that $\nabla\psi(x) f_1(x) > 0$, and $\nabla\psi(x) f_2(x) > 0$, the sensitivity of the solution of the smoothed problem, s_σ, approaches the sensitivity of the solution of the original problem, s, as $\sigma \to 0$.*

Proof Follows by comparing the right-hand side of the last displayed equality with (16), as well as our conclusion that the right-hand side of (19) converges to (20) away from the switching time t^*. □

5.2 Lagrange multipliers and the jump rule

There is another way to obtain this result via the adjoint equation from the Pontryagin conditions. Since this is a problem without control functions, we consider the problem of minimizing some objective function $g(x(T))$ by varying the initial value x_0. Again we consider using a smoothed right-hand side f_σ (17).

From the conventional Pontryagin conditions [5,19,27] we have the adjoint equations

$$\frac{d\lambda_\sigma}{dt} = -\nabla f_\sigma(x_\sigma(t))^T\lambda_\sigma, \quad \lambda_\sigma(T) = \nabla g(x_\sigma(T)). \qquad (23)$$

As above, we note that

$$\begin{aligned}
\nabla f_\sigma(x) = {}& \frac{\varphi'(\psi(x)/\sigma)}{\sigma}[f_2(x) - f_1(x)]\nabla\psi(x) \\
& + \{\varphi(\psi(x)/\sigma)\,\nabla f_1(x) + (1 - \varphi(\psi(x)/\sigma))\,\nabla f_2(x)\}
\end{aligned}$$

and that the terms enclosed in $\{\cdots\}$ are bounded as $\sigma \downarrow 0$. Integrating (23) backwards in time and using the matrix exponential, we get the approximation around $t = t^*$, where $\psi(x(t^*)) = 0$:

$$\lambda_\sigma(t^* - \epsilon) = \left[I + \frac{e^{\alpha(\gamma_2 - \gamma_1)} - 1}{\gamma_2 - \gamma_1} (f_2^* - f_1^*) \nabla \psi^* \right] \lambda_\sigma(t^* + \epsilon) + O(\epsilon), \quad (24)$$

where $\sigma = o(\epsilon)$ as $\epsilon \downarrow 0$, and α is some nonnegative quantity. Setting $\beta := (e^{\alpha(\gamma_2 - \gamma_1)} - 1)/(\gamma_2 - \gamma_1)$, we get

$$\lambda_\sigma(t^* - \epsilon) = \left[I + \beta (f_2^* - f_1^*) \nabla \psi^* \right] \lambda_\sigma(t^* + \epsilon) + O(\epsilon). \quad (25)$$

The only quantity that is not determined by this approach is β. But it can be computed from the property that the Hamiltonian $H_\sigma(x_\sigma(t), \lambda_\sigma(t)) := \lambda_\sigma(t)^T f_\sigma(x_\sigma(t))$ is a constant function of t. Applying this property at times $t^* \pm \epsilon$ we find that

$$\beta = \frac{1}{\nabla \psi^* f_1^*} + O(\epsilon). \quad (26)$$

Taking $\epsilon \downarrow 0$ and $\sigma = o(\epsilon)$ gives a simple "jump rule" for the adjoint variables:

$$\lambda(t^{*-}) = \left[I + \frac{(f_2^* - f_1^*) \nabla \psi^*}{\nabla \psi^* f_1^*} \right] \lambda(t^{*+}). \quad (27)$$

That the adjoint functions have discontinuities in problems with discontinuous right-hand sides was noted by, for example, Driessen and Sadegh [10] and Kim and Ha [23].

5.3 Convergence results

We now prove the main convergence result for this case.

Theorem 2 *Assume that* $\nabla \psi(x) f_1(x) > 0$, $\nabla \psi(x) f_2(x) > 0$, *whenever* $\psi(x) = 0$. *Assume that we integrate the smoothed model equation* (17) *and the corresponding sensitivity equation* (19) *for* $\nabla_{x_0} x(t; x_0)$ *using Euler's method* ($\chi = 0$) *with a time step* $h = o(\sigma)$. *Then, the numerical sensitivities and the numerical adjoints converge to the sensitivities of the original problem as* $\sigma \to 0$.

The assumption that $h = o(\sigma)$ is more restrictive than necessary away from the transition region, where larger values for h could be used. In this case, where the trajectory crosses the discontinuity without remaining on it, an adaptive ODE solver could be successfully used. However, if the trajectory is "trapped" in the transition region (as discussed in Sect. 6), this will not give much benefit in terms of efficiency.

We separate the proof in the following parts:

1. In Sect. 5.3.1 we prove that the sequence of state variables produced by Euler's method applied to the smoothed equation converge to the one of the model problem (9).

2. In Sect. 5.3.2 we prove that the sequence of adjoint variables is convergent. Proof of convergence of the sensitivities can be shown by essentially the same arguments as are used for the adjoint equations.

5.3.1 Errors in the computed trajectory

We first note that solutions to the smoothed ODE $dx_\sigma/dt = f_\sigma(x_\sigma)$ (17) converge to solution of the differential inclusion (9) as $\sigma \to 0$, since the graph of f_σ approaches the graph of the right-hand side of (9) [2]. We therefore concentrate on the errors involved in the numerical solution of (17).

Consider using the explicit Euler method for the numerical solution of $dx_\sigma/dt = f_\sigma(x_\sigma)$. Let $t_k = t_0 + k\,h$, where $h > 0$ is the step size:

$$\begin{aligned} x_\sigma^{k+1} &= x_\sigma^k + h\, f_\sigma(x_\sigma^k), \\ x_\sigma(t_{k+1}) &= x_\sigma(t_k) + h\, f_\sigma(x_\sigma(t_k)) + \xi_k, \end{aligned}$$

where $\|\xi_k\| \le \frac{1}{2}h^2 \max_{t_k \le t \le t_{k+1}} \|\ddot{x}_\sigma(t)\|$ by Taylor's theorem to second order. We suppose that $h\,L < 1$. Subtracting the equations for x_σ^{k+1} and $x_\sigma(t_{k+1})$ gives

$$x_\sigma(t_{k+1}) - x_\sigma^{k+1} = (x_\sigma(t_k) - x_\sigma^k) + h\,[f_\sigma(x_\sigma(t_k)) - f_\sigma(x_\sigma^k)] + \frac{1}{2}h^2\xi_k. \tag{28}$$

For all k, put $e_{\sigma,k} = x_\sigma(t_k) - x_\sigma^k$. Then

$$e_{\sigma,k+1} = e_{\sigma,k} + h\,[f_\sigma(x_\sigma^k + e_{\sigma,k}) - f_\sigma(x_\sigma^k)] + \frac{1}{2}h^2\xi_k. \tag{29}$$

Lemma 2 *Under our standing assumptions, the map $z \mapsto z + h[f_\sigma(x+z) - f_\sigma(x)]$ is Lipschitz with constant $1 + hL + Ch^2/\sigma^2$ for some constant C independent of h and σ.*

Proof Let $\Phi_{h,\sigma,x}(z) = z + h[f_\sigma(x+z) - f_\sigma(x)]$. Then for any z_1, z_2,

$$\begin{aligned} &\left\|\Phi_{h,\sigma,x}(z_1) - \Phi_{h,\sigma,x}(z_2)\right\|^2 \\ &\quad = \|z_1 - z_2 + h[f_\sigma(x+z_1) - f_\sigma(x+z_2)]\|^2 \\ &\quad = \|z_1 - z_2\|^2 + 2h(z_1 - z_2)^T[f_\sigma(x+z_1) - f_\sigma(x+z_2)] \\ &\qquad + h^2\,\|f_\sigma(x+z_1) - f_\sigma(x+z_2)\|^2 \\ &\quad \le \|z_1 - z_2\|^2 + 2hL\|z_1 - z_2\|^2 + h^2\,\|f_\sigma(x+z_1) - f_\sigma(x+z_2)\|^2 \quad \text{(using (4))} \\ &\quad \le (1 + 2hL + Ch^2/\sigma^2)\|z_1 - z_2\|^2, \end{aligned}$$

since f_σ is Lipschitz with constant $C^{1/2}/\sigma$ for some C independent of h or σ. Therefore,

$$\left\|\Phi_{h,\sigma,x}(z_1) - \Phi_{h,\sigma,x}(z_2)\right\| \le (1 + 2hL + Ch^2/\sigma^2)^{1/2}\|z_1 - z_2\|.$$

Note that for $w \geq 0$, $(1+w)^{1/2} \leq 1 + \frac{1}{2}w$, so

$$\left\| \Phi_{h,\sigma,x}(z_1) - \Phi_{h,\sigma,x}(z_2) \right\| \leq (1 + hL + Ch^2/(2\sigma^2)) \|z_1 - z_2\|,$$

as desired. □

Using Lemma 2, we see that

$$\|e_{k+1}\| \leq (1 + hL + Ch^2/\sigma^2)\|e_k\| + \frac{1}{2}h^2 \|\xi_k\|. \tag{30}$$

We can also bound $\ddot{x}_\sigma(t)$ because

$$\begin{aligned} \ddot{x}_\sigma(t) &= \frac{d}{dt} f_\sigma(x_\sigma(t)) \\ &= \nabla f_\sigma(x_\sigma(t))\, f_\sigma(x_\sigma(t)), \end{aligned}$$

which immediately gives the bounds

$$\|\ddot{x}_\sigma(t)\| \leq \|\nabla f_\sigma(x_\sigma(t))\|\, \|f_\sigma(x_\sigma(t))\|. \tag{31}$$

Using local boundedness of f, we can bound f_σ independently of σ over compact sets. Thus $\|\ddot{x}_\sigma(t)\| = O(\nabla f_\sigma(x_\sigma(t))) = O(1/\sigma)$.

Thus if $h = o(\sigma^2)$, we can combine these bounds to show that for $L' > L$ and sufficiently small $h > 0$

$$\|e_{k+1}\| \leq e^{L'h}\|e_k\| + C\frac{h^2}{\sigma}. \tag{32}$$

Using a discrete Gronwall lemma and $e_0 = 0$, we can see that

$$\|e_k\| \leq \frac{e^{L'kh} - 1}{L'h} C \frac{h^2}{\sigma} = (e^{L'(t_k - t_0)} - 1) O(h/\sigma). \tag{33}$$

This bound does not depend on how long the trajectory stays in the transition region, and will be used for the case where the trajectory stays in the discontinuity.

Bound (33) can be considerably improved if we know that $\nabla\psi^* f_1^*,\ \nabla\psi^* f_2^* > 0$, as then the time to cross the transition region is $O(\sigma)$. Outside the transition region, the error is $O(h)$, as is well known [1]. It will take $O(\sigma/h)$ time steps to cross the transition region under the assumptions that $\nabla\psi^* f_1^*,\ \nabla\psi^* f_2^* > 0$. Suppose we choose $k^* = k^*(\sigma, h)$ so that $x_\sigma(t_{k^*})$ is outside the transition region, but $x_\sigma(t_{k^*+1})$ is inside the transition region, and after $K = K(\sigma, h) = O(\sigma/h)$ steps we find that $x_\sigma(t_{k^*+K})$ is outside the transition region, and the discrete time trajectory remains outside the transition region for a positive time period (a period whose length does not go to zero as $h \to 0$).

For now we assume only that $h = o(\sigma)$. Then from (30) and (31) we find that

$$\begin{aligned}\|e_{k^*+K}\| &\le \exp((hL + Ch^2/\sigma^2)K(\sigma,h))\left[\|e_{k^*}\| + C\,\frac{h^2}{\sigma}K(\sigma,h)\right]\\ &\le \exp(C'(L\sigma + Ch/\sigma))\left[\|e_{k^*}\| + C\,C'\,h\right] = O(h),\end{aligned}$$

where $K \le C'\sigma/h$, for h, σ small enough. Once outside the transition region, standard bounds apply for $k \ge k^* + K$. Thus $e_k \to 0$ as $h,\ \sigma \to 0$ with $h = o(\sigma)$.

5.3.2 Errors in the adjoints

Recall that for the Euler method

$$x_\sigma^{k+1} = x_\sigma^k + h\, f_\sigma(x_\sigma^k).$$

Thus a small variation δx_σ^k in x_σ^k results in a variation

$$\delta x_\sigma^{k+1} = [I + h\,\nabla f_\sigma(x_\sigma^k)]\,\delta x_\sigma^k + o(\|\delta x_\sigma^k\|) \tag{34}$$

in x_σ^{k+1}. The discrete sensitivity equations are thus

$$s_\sigma^{k+1} = [I + h\,\nabla f_\sigma(x_\sigma^k)]\,s_\sigma^k,$$

with given $s_\sigma^0 = s^0$.

Alternatively, we can consider the discrete adjoint variables. Suppose we have a function $g\colon \mathbb{R}^n \to \mathbb{R}$ and we wish to determine the gradient of the $g(x^N)$, where $t_f = t_0 + Nh$ with respect to a change in the initial values x^0, where x^N is computed via Euler's method. Let $\Gamma_i(x^i) = g(x^N)$ where $x^{k+1} = x^k + h\, f_\sigma(x^k)$ for $k = i,\ i+1, \ldots, N-1$. Set $\lambda^i = \nabla\Gamma_i(x^i)^T$. If $J_k = I + h\,\nabla f_\sigma(x^k)$, for all k, then $J_k = \nabla_{x^k}(x^{k+1})$, so

$$\begin{aligned}\lambda^i &= \nabla\Gamma_i(x^i)^T = (\nabla\Gamma_{i+1}(x^{i+1})\,J_i)^T\\ &= [I + h\,\nabla f_\sigma(x^i)]^T\,\lambda^{i+1},\end{aligned}$$

and $\lambda^N = \nabla g(x^N)^T$, which are the discrete adjoint equations.

We can investigate the accuracy of either the direct sensitivity equations, or the discrete adjoint equations, to determine the accuracy of the computed gradients.

The adjoint equations for the differential equations are

$$\frac{d\lambda_\sigma}{dt} = -\nabla f_\sigma(x_\sigma(t))^T\lambda_\sigma, \quad \lambda_\sigma(T) = \nabla g(x_\sigma(T))^T.$$

Thus

$$\lambda_\sigma(t_k) = [I + h\,\nabla f_\sigma(x_\sigma(t_k))^T]\lambda_\sigma(t_{k+1}) + \eta_k, \tag{35}$$

$$\lambda_\sigma^k = [I + h\,\nabla f_\sigma(x_\sigma^k)^T]\lambda_\sigma^{k+1}, \tag{36}$$

where $\|\eta_k\| \leq \frac{1}{2}h^2 \max_{t_k \leq t \leq t_{k+1}} \|\ddot{\lambda}_\sigma(t)\|$. We can bound $\ddot{\lambda}_\sigma(t)$ by differentiating the adjoint equation:

$$\begin{aligned}\ddot{\lambda}_\sigma(t) &= -\frac{d}{dt}(\nabla f_\sigma(x_\sigma(t))^T \lambda_\sigma(t)) \\ &= -\nabla_x(\nabla f_\sigma(x)^T \lambda_\sigma(t))|_{x=x_\sigma(t)} \frac{dx_\sigma}{dt}(t) - \nabla f_\sigma(x_\sigma(t))^T \frac{d\lambda_\sigma}{dt}(t) \\ &= -\nabla_x(\nabla f_\sigma(x)^T \lambda_\sigma(t))|_{x=x_\sigma(t)} \frac{dx_\sigma}{dt}(t) + [\nabla f_\sigma(x_\sigma(t))^2]^T \lambda_\sigma(t).\end{aligned}$$

Now dx_σ/dt is bounded on finite intervals independently of σ as f_σ is bounded with a one-sided Lipschitz condition. Also λ_σ is bounded independently of σ. Noting that $\nabla f_\sigma(x) = O(1/\sigma)$ and $\nabla\nabla f_\sigma(x) = O(1/\sigma^2)$, we see that $\ddot{\lambda}_\sigma(t) = O(1/\sigma^2)$ in the transition region. Outside the transition region, $\ddot{\lambda}_\sigma(t) = O(1)$. Note that for t in a fixed finite interval, the constants implicit in the "O" expressions can be independent of t.

Thus η_k is $O(h^2/\sigma^2)$ with a constant independent of h, σ, and k, where $x_\sigma(t)$ is in the transition region for some $t_k \leq t \leq t_{k+1}$. Otherwise $\eta_k = O(h^2)$ with a constant independent of h, σ, and k.

To obtain bounds on the errors in the adjoints, we first subtract (36) from (35). This gives

$$\begin{aligned}\lambda_\sigma(t_k) - \lambda_\sigma^k &= [I + h\, f_\sigma(x_\sigma(t_k))]^T [\lambda_\sigma(t_{k+1}) - \lambda_\sigma^{k+1}] \\ &\quad + h[\nabla f_\sigma(x_\sigma(t_k)) - \nabla f_\sigma(x_\sigma^k)]\lambda_\sigma^{k+1} + \eta_k.\end{aligned}$$

Now $\|\nabla f_\sigma(x_\sigma(t_k)) - \nabla f_\sigma(x_\sigma^k)\| \leq (C/\sigma^2)\|x_\sigma(t_k) - x_\sigma^k\|$ as ∇f_σ is Lipschitz on bounded sets with a Lipschitz constant of $O(1/\sigma^2)$. Assuming that $\nabla\psi^* f_1^*$, $\nabla\psi^* f_2^* > 0$, so that we have $\|x_\sigma(t_k) - x_\sigma^k\| = O(h)$. Furthermore, $\|h[\nabla f_\sigma(x_\sigma(t_k)) - \nabla f_\sigma(x_\sigma^k)]\lambda_\sigma^{k+1} + \eta_k\| = O(h^2/\sigma^2)$ if $x_\sigma(t)$ is in the transition region for some $t_k \leq t \leq t_{k+1}$ or x_σ^k is in the transition region. Otherwise the more usual bounds $\|h[\nabla f_\sigma(x_\sigma(t_k)) - \nabla f_\sigma(x_\sigma^k)]\lambda_\sigma^{k+1} + \eta_k\| = O(h^2)$ hold. Again, assuming that the trajectories $x_\sigma(t)$ or x_σ^k are in the transition region for only $O(\sigma/h)$ many steps, we can apply a Gronwall lemma to obtain a bound

$$\|\lambda_\sigma(t_k) - \lambda_\sigma^k\| = O(h^2/\sigma^2)\, O(\sigma/h) + O(h^2)\, O(1/h) = O(h/\sigma) \tag{37}$$

with the implicit constants independent of k, h, and σ. Thus if $h = o(\sigma)$, and $h,\ \sigma \to 0$, the adjoint variables converge to the gradients of $g(x(t_f))$ with respect to the initial conditions for the discontinuous limit.

6 Convergence of gradients for time discretizations of the case where $\nabla\psi(x) f_1(x) > 0$ and $\nabla\psi(x) f_2(x) < 0$

In the case treated here, the solution of the smoothed equation (17) will stay in the transition region for $\omega(\sigma)$ time. In this case there is also a jump in the limit of the (smoothed) adjoint variables, as well as a "jump formula" for the limiting adjoint equations. Furthermore, the adjoint variables obtained in the limit are the correct adjoints for the discontinuous system.

Within the transition region, the adjoint variables change most rapidly in directions near to $\nabla\psi^*$. By Lemma 1, $f_2^* - f_1^* = -\rho^*\nabla\psi^*$ for some $\rho^* \geq 0$.

If the trajectory stays on the discontinuity for any open interval, then while the trajectory is on the interval, an equivalent right-hand side can be used for the motion on the discontinuity [29].

6.1 The convergence result

In this section, we prove the following result concerning the convergence of the sensitivity of the Eq. (17) to the one of (9).

Theorem 3 *Assume that* $\nabla\psi(x)f_1(x) > 0$, $\nabla\psi(x)f_2(x) < 0$, *whenever* $\psi(x) = 0$. *Assume that we integrate the smoothed model equation* (17) *and the corresponding sensitivity equation* (19) *for* $\partial x/\partial x_0$ *using Euler's method* ($\chi = 0$) *with a time step* $h = o(\sigma^2)$. *Then, the numerical sensitivities and the numerical adjoints converge to the sensitivities of the original problem as* $\sigma \to 0$.

Note that we will need to take $h = o(\sigma^2)$ in order to resolve the trajectory as it passes through the transition region so that the gradient information will be accurate. It will turn out that this is sufficient for the Euler method to compute approximate gradients that converge to the true gradient as computed in the previous section.

Our initial investigations suggest that the result is true even for the case where $h = o(\sigma)$. Nonetheless, this would require additional complexity to an already very technical proof so we will use the stronger assumption. However, wherever possible in the course of the proof, we will invoke only the weaker assumption $h = o(\sigma)$.

The proof of this result is split into a number of items that are discussed in the following subsections. We note that the proof of the convergence of the state vectors is identical to the one for the preceding case, from Sect. 5.3.1. In addition, we will use results from Sects. 3 and 4 that apply to this case.

6.1.1 Asymptotic behavior of $\varphi_\sigma(\psi(x_\sigma(t)))$

Consider the differential equation

$$\begin{aligned}
&\frac{d}{dt}\varphi_\sigma(\psi(x_\sigma))) \\
&\quad = \varphi_\sigma'(\psi(x_\sigma))\nabla\psi(x_\sigma)[(1-\varphi_\sigma(\psi(x_\sigma)))f_1(x_\sigma) + \varphi_\sigma(\psi(x_\sigma))f_2(x_\sigma)] \\
&\quad = \varphi_\sigma'(\psi(x_\sigma))[(1-\varphi_\sigma(\psi(x_\sigma)))\gamma_1(x_\sigma(t)) + \varphi_\sigma(\psi(x_\sigma))\gamma_2(x_\sigma)],
\end{aligned}$$

where $\gamma_i(x) = \nabla\psi(x)^T f_i(x)$, $i = 1, 2$.

We consider this differential equation for a time interval $[t^*_{\sigma,-}, t^*_{\sigma,-} + \epsilon]$ where $t^*_{\sigma,-}$ is the first time when we reach the transition zone. Let t^* be the time when the limit

$x(\cdot)$ of $x_\sigma(\cdot)$ reaches Σ, and $x^* = x(t^*)$. Let $\gamma_i^* = \gamma_i(x^*)$. In an interval of this size, $\gamma_i(x_\sigma(t)) = \gamma_i^* + O(\epsilon + \sigma)$. We will consider $\epsilon = \omega(\sigma)$, so $\gamma_i(x_\sigma(t)) = \gamma_i^* + O(\epsilon)$. Setting $w(t) = \gamma_1^* - (\gamma_1^* - \gamma_2^*)\varphi_\sigma(\psi(x_\sigma(t)))$, we see that

$$\frac{dw}{dt} = -\mu(t)[w(t) + g(t)], \tag{38}$$

where $\mu(t) = (\gamma_1^* - \gamma_2^*)\varphi_\sigma'(\psi(x_\sigma(t)))$ and $g(t) = O(\epsilon/\sigma)$. Note that in a time interval of size $O(\sigma)$ the trajectory will reach a point where $\varphi_\sigma(\psi(x_\sigma(t)))$ is halfway between zero and $\gamma_1^*/(\gamma_1^* - \gamma_2^*)$. Call this time $t_{\sigma,-1/2}^*$. We can bound $\varphi'(\psi(x_\sigma(t)))$ away from zero, at least for a time interval of length bounded away from zero.

On our time interval of length ϵ, then, $\mu(t) = \Theta(1/\sigma)$ and $\gamma_i(x_\sigma(t)) = \gamma_i^* + O(\epsilon)$. We can solve the differential equation for w starting from $t_{\sigma,-1/2}^*$:

$$\begin{aligned} w(t_{\sigma,-1/2}^* + t) &= \exp\left(-\int_0^t \mu(t_{\sigma,-1/2}^* + \tau)\, d\tau\right) w(t_{\sigma,-1/2}^*) \\ &\quad + \int_0^t \exp\left(-\int_\tau^t \mu(t_{\sigma,-1/2}^* + s)\, ds\right) g(t_{\sigma,-1/2}^* + \tau)\, d\tau. \end{aligned}$$

Thus for $0 \le t \le \epsilon$ we get

$$w(t_{\sigma,-1/2}^* + t) = \int_0^t e^{-C(t-\tau)/\sigma}\, O(\epsilon/\sigma)\, d\tau = O(e^{-Ct/\sigma}) + O(\epsilon). \tag{39}$$

So in a time $\ge \text{const}\, \sigma \log(1/\epsilon)$ the difference between $\varphi_\sigma(\psi(x_\sigma(t)))$ and $\gamma_1^*/(\gamma_1^* - \gamma_2^*)$ is $O(\epsilon)$.

Set $\theta(x) = \gamma_1(x)/(\gamma_1(x) - \gamma_2(x))$. More careful analysis shows that for t large compared with $\sigma \log(1/\sigma)$, the difference between $\varphi_\sigma(\psi(x_\sigma(t)))$ and $\theta(x_\sigma(t))$ is $O(\sigma)$. To see this, construct the solution to the differential equation

$$\frac{d}{dt}(w - g) = -\mu(t)(w - g) - g'(t) \tag{40}$$

as

$$w(t) - g(t) = \exp\left(-\int_0^t \mu(\tau)\, d\tau\right)(w(0) - g(0)) - \int_0^t \exp\left(-\int_\tau^t \mu(s)\, ds\right) g'(\tau)\, d\tau$$

and substitute $w(t) = \varphi_\sigma(\psi(x_\sigma(t_{\sigma,-}^* + t)))$ and $g(t) = \theta(x_\sigma(t_{\sigma,-}^* + t))$.

6.1.2 Asymptotic behavior of $\varphi_\sigma(\psi(x_\sigma^k))$

Note that if $h = O(\epsilon\sigma^2)$, then we can use the global error bound in (33). Since $x \mapsto \varphi_\sigma(\psi(x))$ is Lipschitz with constant of $O(1/\sigma)$, for $t_k \geq t_{\sigma,-}^*$ we have $(\gamma_1 - \gamma_2)\varphi'_\sigma(\psi(x_\sigma^k)) = O(\epsilon)$. In addition, we have that $\varphi_\sigma(\psi(x_\sigma^k)) \to \theta(x(t))$ for t above the switching point.

This can be improved to merely requiring $h/\sigma = O(\epsilon)$. However, a detailed rigorous demonstration would require improved error bounds that take into account the exponential damping of order $O\ (1/\sigma)$ in the direction perpendicular to the manifold $\Sigma = \{x \mid \psi(x) = 0\}$. Since this would result in a substantial additional complexity to what is already very technical proof, we will not follow it in the context of this paper.

6.1.3 Sensitivity equations in the transition region, and their discretization

If we apply Euler's method (with $h = o(\sigma)$) to the sensitivity equations in the transition region, we get

$$\begin{aligned} s_\sigma^{k+1} &= [I + h\,\nabla f_\sigma(x_\sigma^k)]s_\sigma^k \\ &= \Big[I + h\varphi'_\sigma(\psi(x_\sigma^k))(f_2(x_\sigma^k)) - f_1(x_\sigma^k))\nabla\psi(x_\sigma^k) \\ &\qquad + h\ (1-\varphi_\sigma^k)\nabla f_1(x_\sigma^k) + h\varphi_\sigma^k\nabla f_2(x_\sigma^k)\Big]\, s_\sigma^k, \end{aligned}$$

where $\varphi_\sigma^k = \varphi(s_\sigma^k)$. Note that away from the boundaries of the transition zone, $\varphi'_\sigma(x) = \Theta(1/\sigma)$. Let $u_\sigma^k = f_1(x_\sigma^k) - f_2(x_\sigma^k)$, $v_\sigma^k = \nabla\psi(x_\sigma^k)^T$, and $F_\sigma^k = (1-\varphi_\sigma^k)\nabla f_1(x_\sigma^k) + \varphi_\sigma^k\nabla f_2(x_\sigma^k)$. Then we can write the discrete sensitivity equation as

$$s_\sigma^{k+1} = [I - h\varphi'_\sigma(\psi(x_\sigma^k))u_\sigma^k(v_\sigma^k)^T + hF_\sigma^k]s_\sigma^k. \tag{41}$$

If x_σ^k was on $\Sigma = \{x \mid \psi(x) = 0\}$, then since f satisfies a one-sided Lipschitz condition, $u_\sigma^k \parallel v_\sigma^k$ and $(v_\sigma^k)^T u_\sigma^k \geq 0$. However, while in the transition zone, the distance of x_σ^k from Σ is $O(\sigma)$. Thus the angle between u_σ^k and v_σ^k is $O(\sigma)$ by Lipschitz continuity of f_1, f_2 and $\nabla\psi$, and the assumption that $f_2 - f_1 \neq 0$ and $\nabla\psi \neq 0$ anywhere on Σ.

Let $\alpha_\sigma^k = \varphi'_\sigma(\psi(x_\sigma^k))(v_\sigma^k)^T u_\sigma^k$.

For the remainder of this subsection we will drop the σ subscripts.

Then we can write

$$s^{k+1} = \left[I - h\alpha_k\frac{u^k(v^k)^T}{(v^k)^T u^k} + hF^k\right]s^k. \tag{42}$$

Note that $\alpha_k = \Theta(1/\sigma)$ and $F^k = O(1)$. Since the angle between u^k and v^k is $O(\sigma)$ and $\alpha_k = O(1/\sigma)$,

$$\alpha_k \frac{u^k (v^k)^T}{(v^k)^T u^k} = \alpha_k \frac{u^k (u^k)^T}{(u^k)^T u^k} + O(1). \tag{43}$$

Thus

$$s^{k+1} = \left[I - h\alpha_k \frac{u^k (u^k)^T}{(u^k)^T u^k} + h\widehat{F}^k \right] s^k, \tag{44}$$

where $\widehat{F}^k = O(1)$.

Let $\widehat{u}^k = u^k / \|u^k\|_2$. Choose a family of orthogonal matrices Q_k where $Q_k \widehat{u}^k = \widehat{u}^{k+1}$. Since $\|u^{k+1} - u^k\| = O(h)$, and $\|u^k\|$ is bounded away from zero, we can choose Q_k so that $\|Q_k - I\|_2 = O(h)$. Put $R_k = Q_{k-1} Q_{k-2} \cdots Q_1 Q_0$, with $Q_j = I$ if x_σ^j is not in the transition zone. With this in mind, the discrete sensitivity equations can be rewritten as

$$\begin{aligned} s^{k+1} &= \left[I - h\alpha_k R_k \widehat{u}^0 (\widehat{u}^0)^T R_k^T + h\widehat{F}^k \right] s^k, \\ &= R_k [I - h\alpha_k \widehat{u}^0 (\widehat{u}^0)^T + h R_k^T \widehat{F}^k R_k] R_k^T s^k. \end{aligned}$$

For large k, we can write

$$\begin{aligned} s^k &= \left[\prod_{j=0}^{k-1} (R_j [I - h\alpha_j \widehat{u}^0 (\widehat{u}^0)^T + h R_j^T \widehat{F}^j R_j] R_j^T) \right] s^0 \\ &= R_k \left[\prod_{j=0}^{k-1} (R_{j+1}^T R_j [I - h\alpha_j \widehat{u}^0 (\widehat{u}^0)^T + h R_j^T \widehat{F}^j R_j]) \right] R_0^T\, s^0 \\ &= R_k \left[\prod_{j=0}^{k-1} (Q_j [I - h\alpha_j \widehat{u}^0 (\widehat{u}^0)^T + h R_j^T \widehat{F}^j R_j]) \right] R_0^T\, s^0, \end{aligned} \tag{45}$$

where $\prod_{j=0}^{k-1} A_j = A_{k-1} A_{k-2} \cdots A_1 A_0$. Noting that $Q_j = I + O(h)$, we can absorb the difference $Q_j - I$ in the product into the $O(1)$ term of the above product. This gives

$$s^k = R_k \left[\prod_{j=0}^{k-1} (I - h\alpha_j \widehat{u}^0 (\widehat{u}^0)^T + h G_j) \right] R_0^T\, s^0 \tag{46}$$

with $G_j = O(1)$. Provided $0 \le h\alpha_j \le 1$ for all j, one can easily show that this product is uniformly bounded as $h \downarrow 0$ with kh bounded.

We want to go further and show that this converges to a matrix of the form $R(I - \widehat{u}^0(\widehat{u}^0)^T)G(I - \widehat{u}^0(\widehat{u}^0)^T)$ with R orthogonal.

In addition to supposing that $h = o(\sigma)$, we choose p, an integer (depending on σ), so that $\sigma = o(ph)$, and $ph \to \epsilon$ as $\sigma \downarrow 0$.

6.1.4 Lower bounds for $\varphi_\sigma(\psi(x_\sigma(t)))$

In the following, we make certain assumptions about the properties of the function $\rho(r)$ that defines the smoothing function $\varphi_\sigma(s)$. Recall that we defined

$$\varphi_\sigma(s) = \int_{-\infty}^{s} \rho\left(\frac{r}{\sigma}\right)\frac{dr}{\sigma} = \int_{-\infty}^{x/\sigma} \rho\left(r'\right)dr' = \int_{-1}^{s/\sigma} \rho\left(r'\right)dr'.$$

Namely, we assume that there exists a positive parameter c such that

$$\int_{-1}^{s} \rho(r')dr' \leq c\rho(s), \quad s \in [-1, 0];$$

$$\int_{s}^{1} \rho(r')dr' \leq c\rho(s), \quad s \in [0, 1].$$

An immediate consequence of this assumption is that

$$\min(\varphi_\sigma(s), (1 - \varphi_\sigma(s)) \leq c\rho\left(\frac{s}{\sigma}\right) = c\rho_\sigma\left(\frac{s}{\sigma}\right), \quad s \in [-\sigma, \sigma].$$

Since we aimed to prove certain properties of the solution while in the transition region and while the solution follows the discontinuity manifold, it is important to define the transition region for the situation where $\sigma \neq 0$. In our case, we simply define it as the point $x_\sigma(t)$ that satisfies $\rho_\sigma(\psi(x_\sigma(t))) \geq k_1$, where $k_1 > 0$ is sufficiently small. In addition, since we have shown that $\varphi_\sigma(\psi(x_\sigma(t))) \overset{\sigma\to 0}{\longrightarrow} \rho(x(t))$ ($\theta(x)$ given in (12)), and since our assumption that $\nabla\psi(x)^T f_1(x) > 0$ and $\nabla\psi(x)^T f_2(x) < 0$ implies that $\theta(x(t))$ must be bounded away from 1, it follows that in the transition region and in the region that follows the discontinuity manifold, we have the following inequality:

$$k_1 \leq \varphi_\sigma(\psi(x_\sigma(t))) \leq 1 - k_2.$$

In turn, this implies that whenever $x_\sigma(t)$ is in the transition region or follows the discontinuity, we will have that

$$\min(k_1, k_2) \leq \min(\varphi_\sigma(\psi(x_\sigma(t))), (1 - \varphi_\sigma(\psi(x_\sigma(t)))) \leq c\rho\left(\frac{\psi(x_\sigma(t))}{\sigma}\right),$$

Therefore, in the same regime we have that

$$\varphi_\sigma'(\psi(x_\sigma(t)) = \frac{1}{\sigma}\rho\left(\frac{\psi(x_\sigma(t)}{\sigma}\right) \geq \frac{\min(k_1, k_2)}{c\sigma}. \tag{47}$$

As a result, we have for $I_\sigma = \{\, t \mid |\psi(x_\sigma(t))| \leq \sigma \,\}$,

$$h \sum_{t_k \in I_\sigma} \varphi_\sigma'(\psi(x_\sigma(t_k)) \to \infty$$

as soon as $m(I_\sigma) \geq \sigma^p,\ p \in (0, 1)$.

6.1.5 Results on products of nearby projections

Consider the product with $\|\widehat{u}_i\|_2 = 1$ for all i and $\widehat{u}_{i+1} - \widehat{u}_i = O(h)$:

$$P := \prod_{i=1}^{p} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T).$$

Let $Q_{i,i+1}\widehat{u}_i = \widehat{u}_{i+1}$ be an orthogonal matrix: $Q_{i,i+1} = I + (\widehat{u}_{i+1} - \widehat{u}_i)\widehat{u}_i^T - \widehat{u}_i(\widehat{u}_{i+1} - \widehat{u}_i)^T + O(\|\widehat{u}_{i+1} - \widehat{u}_i\|^2)$ Then put $Q_{r,s} = Q_{r,r+1}Q_{r+1,r+2} \cdots Q_{s-1,s}$ for $s > r$. Note that $Q_{r,s}\widehat{u}_r = \widehat{u}_s$. Also $Q_{r,s}^T\widehat{u}_s = \widehat{u}_r$ so we put $Q_{r,s}^T = Q_{s,r}$. Then

$$\begin{aligned}
P &= \prod_{i=1}^{p} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T) \\
&= \prod_{i=1}^{p} (I - h\alpha_i Q_{0,i} \widehat{u}_0 \widehat{u}_0^T Q_{0,i}^T) \\
&= \prod_{i=1}^{p} (Q_{0,i}(I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) Q_{0,i}^T) \\
&= Q_{0,j+1}\left[\prod_{i=1}^{p} (Q_{0,i+1}^T Q_{0,i}(I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T))\right] Q_{0,1}^T.
\end{aligned}$$

Note that $Q_{0,i+1} = Q_{i,i+1}Q_{0,i}$, so

$$\begin{aligned}
Q_{0,i+1}^T Q_{0,i} &= Q_{0,i}^T Q_{i,i+1}^T Q_{0,i} \\
&= Q_{0,i}^T \left[I + \widehat{u}_i(\widehat{u}_{i+1} - \widehat{u}_i)^T - (\widehat{u}_{i+1} - \widehat{u}_i)\widehat{u}_i^T + O(\|\widehat{u}_{i+1} - \widehat{u}_i\|^2)\right] Q_{0,i} \\
&= I + Q_{0,i}^T \widehat{u}_i(\widehat{u}_{i+1} - \widehat{u}_i)^T Q_{0,i} - Q_{0,i}^T(\widehat{u}_{i+1} - \widehat{u}_i)\widehat{u}_i^T Q_{0,i} \\
&\quad + O(\|\widehat{u}_{i+1} - \widehat{u}_i\|^2) \\
&= I + \widehat{u}_0 z_i^T - z_i \widehat{u}_0^T + O(\|\widehat{u}_{i+1} - \widehat{u}_i\|^2),
\end{aligned}$$

where $z_i = \widehat{u}_{i+1} - \widehat{u}_i$. Note that $z_i = O(h)$, and so $O(\|\widehat{u}_{i+1} - \widehat{u}_i\|^2) = O(h^2)$. If $\widetilde{G}_i := Q_{0,i+1}^T Q_{0,i} - I = \widehat{u}_0 z_i^T - z_i \widehat{u}_0^T + O(h^2)$, then

$$P = Q_{0,j+1} \left[\prod_{i=1}^{p} ((I + \widetilde{G}_i)(I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T)) \right] Q_{0,1}^T.$$

Expanding the factors with $I + \widetilde{G}_i$, we get

$$\begin{aligned} P = {} & Q_{0,j+1} \left[\prod_{i=1}^{p} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right] Q_{0,1}^T \\ & + Q_{0,j+1} \left[\sum_{i=1}^{p} \left\{ \prod_{j=i+1}^{p} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right\} \widetilde{G}_i \left\{ \prod_{j=1}^{i-1} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right\} \right] Q_{0,1}^T \\ & + O(h^2 p^2). \end{aligned}$$

Assume that all $\alpha_i = \Theta(1/\sigma)$ with $\sigma > 0$ small (bounds independent of i) and $h = o(\sigma)$. Then

$$\begin{aligned} \prod_{j=r}^{s} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) &= I - \left[1 - \prod_{j=r}^{s} (1 - h\alpha_j) \right] \widehat{u}_0 \widehat{u}_0^T \\ &= P_0 + O\left(\prod_{j=r}^{s} (1 - h\alpha_j) \right), \quad P_0 = I - \widehat{u}_0 \widehat{u}_0^T. \end{aligned}$$

Thus

$$\begin{aligned} & \sum_{i=1}^{p} \left\{ \prod_{j=i+1}^{p} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right\} \widetilde{G}_i \left\{ \prod_{j=1}^{i-1} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right\} \\ & = \sum_{i=1}^{p} P_0 \widetilde{G}_i P_0 + O(h\,(\sigma/h) \log(\sigma/h)). \end{aligned}$$

But $P_0 \widetilde{G}_i P_0 = (I - \widehat{u}_0 \widehat{u}_0^T)(\widehat{u}_0 z_i^T - z_i \widehat{u}_0^T + O(h^2))(I - \widehat{u}_0 \widehat{u}_0^T) = O(h^2)$, so

$$\sum_{i=1}^{p} \left\{ \prod_{j=i+1}^{p} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right\} \widetilde{G}_i \left\{ \prod_{j=1}^{i-1} (I - h\alpha_i \widehat{u}_0 \widehat{u}_0^T) \right\} = O(\sigma \log(\sigma) + ph^2).$$

Thus $P = Q_{0,j+1} P_0 Q_{0,1}^T + O(\sigma + ph^2 + p^2 h^2)$. In fact, noting that $Q_{0,1} = I + O(h)$, we get

$$P = Q_{0,j+1}(I - \widehat{u}_0 \widehat{u}_0^T) + O(\sigma + ph^2 + p^2 h^2).$$

Taking $ph \to \epsilon$, we get

$$P = I - \widehat{u}_0 \widehat{u}_0^T + O(\epsilon). \tag{48}$$

6.1.6 *Jump conditions for the sensitivity*

We wish to use the above result to show that if $G_i = O(1)$ for all i, then

$$\prod_{i=1}^{p} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T + hG_i) = I - \widehat{u}_0 \widehat{u}_0^T + O(\epsilon) \tag{49}$$

provided $ph = O(\epsilon)$.

Now

$$\begin{aligned}
\prod_{i=1}^{p} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T + hG_i) &= \prod_{i=1}^{p} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T) \\
&\quad + h \sum_{i=1}^{p} \left\{ \prod_{j=i+1}^{p} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T) \right\} G_i \left\{ \prod_{j=1}^{i-1} (I - h\alpha_i \widehat{u}_i \widehat{u}_i^T) \right\} \\
&= I - \widehat{u}_0 \widehat{u}_0^T + O(\epsilon) + h \sum_{i=1}^{p} (I - \widehat{u}_0 \widehat{u}_0^T + O(\epsilon)) G_i (I - \widehat{u}_0 \widehat{u}_0^T + O(\epsilon)) \\
&\qquad \text{(from (48))} \\
&= I - \widehat{u}_0 \widehat{u}_0^T + O(\epsilon).
\end{aligned}$$

Taking limits as $h, \sigma \to 0$ with $h = o(\sigma^2)$ gives $s(t^* + \epsilon) = (I - \widehat{u}_0 \widehat{u}_0^T) s(t^* - \epsilon) + O(\epsilon)$. That is, $s(t^{*+}) = (I - \widehat{u}_0 \widehat{u}_0^T) s(t^{*-})$, which is the required jump rule for the sensitivities (14).

6.1.7 *Convergence to the differential equation for the sensitivity on the discontinuity*

We assume that $x(t^*)$ (the exact solution) lies on Σ. Then $s(t^{*+}) \perp \nabla\psi(x(t^*))$ in Sect. 6.1.6.

We consider a time interval $[t^*, t^* + \epsilon]$ with $0 < \epsilon \ll 1$ and take $ph \to \epsilon$, $h = o(\sigma)$, $\sigma = O(\epsilon^2)$. Then the limit of the computed sensitivities can be computed from

$$s^{k^*+p} = \prod_{i=1}^{p} \left(I - h\alpha_i \frac{u^i (v^i)^T}{(v^i)^T u^i} + h\, F_i \right) s^{k^*},$$

where $F_i = (1 - \varphi_\sigma(\psi(x^{k^*+i}))\nabla f_1(x^{k^*+i}) + \varphi_\sigma(\psi(x^{k^*+i}))\nabla f_2(x^{k^*+i})$, and so forth. Note that $\varphi_\sigma(\psi(x^{k^*+i})) - \theta(x^{k^*+i}) = O(\sigma)$ with a constant independent of i, $1 \leq$

$i \le p$. From the above computations,

$$\begin{aligned} s^{k^*+p} &= Q_{0,p}(I - \widehat{u}^0(\widehat{u}^0)^T)\, s^{k^*} \\ &\quad + h \sum_{i=1}^{p} \left\{ \prod_{j=i+1}^{p} (I - h\alpha_{j+k^*}\widehat{u}^j(\widehat{u}^j)^T) \right\} F_i \left\{ \prod_{j=1}^{i} (I - h\alpha_{j+k^*}\widehat{u}^j(\widehat{u}^j)^T) \right\} s^{k^*} \\ &\quad + O(h^2 p^2 + \sigma). \end{aligned}$$

Note that $Q_{0,p}(I - \widehat{u}^0(\widehat{u}^0)^T) = (I - \widehat{u}^p(\widehat{u}^p)^T)Q_{0,p}$. Since the hidden constants in the "O" are independent of h, p or σ (provided hp bounded), and the F_i are bounded, it follows that the component of s^{k^*+p} orthogonal to $\widehat{u}^p$ is Lipschitz in hp with a Lipschitz constant independent of h, p or σ provided hp is bounded. The component in the direction of $\widehat{u}^p$ will be shown to go to zero, uniformly in σ, h and p.

Since $F_i - [(1 - \theta(x^{k^*+i})\nabla f_1(x^{k^*+i}) + \theta(x^{k^*+i})\nabla f_2(x^{k^*+i})] = O(\sigma)$, and the trajectories converge, we can take the limit as $h \to 0$, $ph \to \epsilon$. This gives the solution of the smoothed problem. We can also (simultaneously) take $\sigma \to 0$ with $h = o(\sigma)$ to get the limit of the computations with Euler's method. Without taking the limit as $\sigma \to 0$, we get

$$\begin{aligned} s^{k^*+p} &= Q_{0,p}(I - \widehat{u}^0(\widehat{u}^0)^T)\, s^{k^*} \\ &\quad + h\sum_{i=1}^{p} Q_{i+1,p}(I - \widehat{u}^{i+1}(\widehat{u}^{i+1})^T) F_i Q_{1,i-1}(I - \widehat{u}^1(\widehat{u}^1)^T)\, s^{k^*} \\ &\quad + O(h^2p^2 + \sigma) \\ &= Q_{0,p}(I - \widehat{u}^0(\widehat{u}^0)^T)\, s^{k^*} \\ &\quad + h\sum_{i=1}^{p} Q_{i+1,p}(I - \widehat{u}^{i+1}(\widehat{u}^{i+1})^T) F_i (I - \widehat{u}^{i-1}(\widehat{u}^{i-1})^T) Q_{1,i-1}\, s^{k^*} \\ &\quad + O(h^2p^2 + \sigma). \end{aligned}$$

Now

$$\begin{aligned} Q_{0,p} - I &= \prod_{i=1}^{p} (I + (\widehat{u}^{i+1} - \widehat{u}^i)(\widehat{u}^i)^T - \widehat{u}^i(\widehat{u}^{i+1} - \widehat{u}^i)^T) - I + O(ph^2) \\ &= \sum_{i=1}^{p} \left[(\widehat{u}^{i+1} - \widehat{u}^i)(\widehat{u}^i)^T - \widehat{u}^i(\widehat{u}^{i+1} - \widehat{u}^i)^T \right] + O(p^2h^2) \\ &= \sum_{i=1}^{p} \left[(\widehat{u}^{i+1} - \widehat{u}^i)(\widehat{u}^0)^T - \widehat{u}^0(\widehat{u}^{i+1} - \widehat{u}^i)^T \right] + O(p^2h^2) \\ &= (\widehat{u}^{p+1} - \widehat{u}^0)(\widehat{u}^0)^T - \widehat{u}^0(\widehat{u}^{p+1} - \widehat{u}^0)^T + O(p^2h^2). \end{aligned}$$

Assuming $(s^{k*})^T\nabla\psi(x(t^*)) = O(\sigma)$, we get

$$\begin{aligned}\frac{s^{k^*+p}-s^{k^*}}{ph} &= [(\widehat{u}^{p+1}-\widehat{u}^0)(\widehat{u}^0)^T - \widehat{u}^0(\widehat{u}^{p+1}-\widehat{u}^0)^T]s^{k^*}/(ph)\\
&\quad +\frac{1}{p}\sum_{i=1}^{p} Q_{i+1,p}(I-\widehat{u}^{i+1}(\widehat{u}^{i+1})^T)F_i(I-\widehat{u}^{i-1}(\widehat{u}^{i-1})^T)Q_{1,i-1}\,s^{k^*}\\
&\quad +O(hp+\sigma/(hp))\\
&= -\widehat{u}^0(\widehat{u}^{p+1}-\widehat{u}^0)^T s^{k^*}/(ph)\\
&\quad +\frac{1}{p}\sum_{i=1}^{p} Q_{i+1,p}(I-\widehat{u}^{i+1}(\widehat{u}^{i+1})^T)F_i(I-\widehat{u}^{i-1}(\widehat{u}^{i-1})^T)Q_{1,i-1}\,s^{k^*}\\
&\quad +O(hp+\sigma/(hp)).\end{aligned}$$

Now $\widehat{u}^{p+1} = u^{p+1}/\|u^{p+1}\|$ and $u^{p+1} = f_1(x^{k^*+p+1}) - f_2(x^{k^*+p+1})$. Since f_1 and f_2 are C^1 and $f_1 - f_2$ is nonzero on Σ, then $x \mapsto (f_1(x)-f_2(x))/\|f_1(x)-f_2(x)\|$ is a smooth map in a neighborhood of Σ. Thus $(\widehat{u}^{p+1}-\widehat{u}^0)/(ph) \to \nabla[(f_1-f_2)/\|f_1-f_2\|](x(t^*))\,f^*(x(t^*)) =: z$ as $ph \to 0$. Since $F_i = (1-\theta(x))\nabla f_1(x(t^*)) + \theta\,\nabla f_2(x(t^*)) + O(ph)$ and $Q_{i+1,p},\ Q_{1,i-1} = I + O(ph)$, it follows that if $P(t) = I - \widehat{u}(t)\widehat{u}(t)^T$, where $\widehat{u}(t) = u(t)/\|u(t)\|$ and $u(t) = f_1(x(t)) - f_2(x(t))$,

$$\begin{aligned}\frac{s^{k^*+p}-s^{k^*}}{ph} &= \Big[-\widehat{u}^0 z^T + P(t^*)\left\{(1-\theta(x))\nabla f_1(x(t^*)) + \theta\,\nabla f_2(x(t^*))\right\}P(t^*)\Big]s^{k^*}\\
&\quad +O(hp+\sigma/(hp)).\end{aligned}$$

Noting that $\sigma = o(\epsilon)$ and taking $ph = \epsilon \to 0$, we have

$$\begin{aligned}\frac{ds}{dt}(t^*) &= \left[-\frac{f_1(x(t^*)-f_2(x(t^*))}{\|f_1(x(t^*)-f_2(x(t^*))\|}z(t^*)^T\right]s(t^*)\\
&\quad +\big[P(t^*)\left\{(1-\theta(x(t^*)))\nabla f_1(x(t^*)) + \theta(x(t^*))\,\nabla f_2(x(t^*))\right\}P(t^*)\big]s(t^*),\end{aligned}$$

with $z(t) = \nabla[(f_1-f_2)/\|f_1-f_2\|](x(t))\,f^*(x(t))$. But $s(t^*) \perp u(t^*)$, so

$$\begin{aligned}\frac{ds}{dt}(t^*) &= \left[-\frac{f_1(x(t^*)-f_2(x(t^*))}{\|f_1(x(t^*)-f_2(x(t^*))\|}z(t^*)^T\right]s(t^*)\\
&\quad +\big[P(t^*)\left\{(1-\theta(x(t^*)))\nabla f_1(x(t^*)) + \theta(x(t^*))\,\nabla f_2(x(t^*))\right\}\big]s(t^*).\end{aligned}$$

Note that $\nabla f^*(x) = (1-\theta(x))\nabla f_1(x) + \theta(x)\nabla f_2(x) + (f_2(x) - f_1(x))\nabla\theta(x)$. Since the equation on the discontinuity is

$$\frac{dx}{dt} = f^*(x),$$

the associated variational equation

$$\frac{ds}{dt} = \nabla f^*(x(t))\, s$$

must keep the tangent plane of the discontinuity invariant. Note that

$$\begin{aligned} P(t^*)\frac{ds}{dt}(t^*) &= P(t^*)\left[(1-\theta(x))\nabla f_1(x(t^*)) + \theta\,\nabla f_2(x(t^*))\right] s(t^*) \\ &= P(t^*)\nabla f^*(x(t^*))\, s(t^*). \end{aligned}$$

In order to obtain the correct sensitivity in the limit, it suffices that the component of $ds/dt(t^*)$ in the direction of $u(t^*)$ is correct. But, $(\widehat{u}^{p+1})^T s^{k^*+p} = O(\sigma)$, so in the limit as h, $\sigma \to 0$ with $h = o(\sigma)$, $u(t) \perp s(t)$ for all $t \approx t^*$. Thus the component of $ds/dt(t^*)$ in the direction of $u(t)$ must also be correct, and so

$$\frac{ds}{dt}(t^*) = \nabla f^*(x(t^*))\, s(t^*).$$

Since this is true for all t^* in the interior of the set $\{\, \tau \mid \psi(x(\tau)) = 0 \,\}$, and since s is Lipschitz on this set (provided $u(t)^T s(t)$ for some t in any interval in $\{\, \tau \mid \psi(x(\tau)) = 0 \,\}$), the limit of the numerically computed sensitivities with $h = o(\sigma)$ satisfy the correct sensitivity equation on the discontinuity:

$$\frac{ds}{dt} = \nabla f^*(x(t))\, s.$$

6.1.8 The jump rule for λ

We have so far been able to obtain the jump rule for the sensitivities when the discontinuity manifold is reached:

$$s(t^{*+}) = \left(I - \frac{u(t^*)u(t^*)^T}{u(t^*)^T u(t^*)} \right) s(t^{*-}). \tag{50}$$

The corresponding jump rule for λ can be found from the following property of the adjoints and sensitivities, which is true for any $s(0)$:

$$\frac{d}{dt}(s_\sigma(t)^T \lambda_\sigma(t)) = 0.$$

Thus $s_\sigma(t)^T\lambda_\sigma(t)$ is independent of t. Now $s(t^{*+}) = P(t^*)\, s(t^{*-})$ where $P(t) = I - u(t)u(t)^T/(u(t)^T u(t))$. So taking $\sigma \to 0$ we obtain $s(t^{*+})^T\lambda(t^{*+}) = s(t^{*-})^T\lambda(t^{*-})$. Thus, $s(t^{*-})P(t^*)^T\lambda(t^{*+}) = s(t^{*-})^T\lambda(t^{*-})$. Since this is true for all values of $s(t^{*-})$, we get the jump rule for λ:

$$\lambda(t^{*-}) = P(t^*)^T \lambda(t^{*+}) = P(t^*)\lambda(t^{*+}), \tag{51}$$

since $P(t^*)$ is symmetric.

Similarly, we find that the adjoint variables will satisfy the following adjoint equations on the manifold of discontinuity:

$$\dot{\lambda}(t) = -\begin{cases} \nabla f_1(x(t; x_0))^T \lambda(t), & t < t_s \\ \nabla f^*(x(t; x_0))^T \lambda(t), & t > t_s \end{cases}, \quad \lambda(T) = \nabla g(x(T)),$$

which satisfies the following jump rule at the discontinuity

$$\lambda(t_s^-) = \left[I + \frac{(f^*\,(x(t_s; x_0)) - f_1\,(x(t_s; x_0)))\, \nabla\psi(x(t_s; x_0))^T}{\nabla\psi(x(t_s; x_0))^T f_1\,(x(t_s; x_0))} \right] \lambda(t_s^+).$$

6.2 Convergence of the controls

An issue that has been side-stepped in the above results, is whether the discrete control functions $u_h(t) = u^k$ for $t_k \le t < t_{k+1}$ as computed by some optimization technique converge to the optimal control $u^*(t)$ as $h \downarrow 0$. In general, we do not expect that an optimization method will be able to find the global minimizer for a finite-dimensional optimization problem. However, we should expect that suitable first-order conditions should be satisfied for the time-discretized problem. Then, are the corresponding first-order optimality conditions satisfied by an appropriate limit of the discrete control functions $u_h(\cdot)$?

The answer we believe is “yes”, and we give a partial explanation as to why, at least where the dynamics are affine in the control:

$$\frac{dx}{dt} = f(x(t)) + B(x(t))\, u(t)$$

and $u(t) \in U$, $U \subset \mathbb{R}^m$ a closed, convex, and bounded set, and with f and B smooth. Since the discrete control functions u_h are uniformly bounded in $L^\infty(t_0, T)$, by Alaoglu’s theorem there is a weakly* convergent subsequence (also denoted u_h, etc.). Within this subsequence there is a further subsequence of x_h (x_h being the piecewise linear interpolation of $x_h(t_k) = x^k$) of discrete trajectories that converges *uniformly* on $[t_0, T]$ by the Arzela–Ascoli theorem. Furthermore, the limits $u_h \rightharpoonup^* \widetilde{u}$ and $x_h \to \widetilde{x}$ in the subsequence satisfies the above differential equation.

An optimization algorithm for the time-discretized problem should satisfy the first order conditions:

$$(\lambda^k)^T B(x^k)\, \delta u^k \ge 0, \quad \text{for all } \delta u^k \in T_U(u^k).$$

That is,

$$(\lambda^k)^T B(x^k)\, (z^k - u^k) \ge 0, \quad \text{for all } z^k \in U.$$

Putting $z_h(t) = z^k$ for $t_k \leq t < t_{k+1}$, we can choose the z^k so that $z_h \rightharpoonup^* z$ for any $z \in L^\infty(t_0, T)$ with $z(t) \in U$. Since the x_h are uniformly Lipschitz as $h \downarrow 0$, we can obtain a lower bound

$$\int_{t_0}^{T} \lambda_h(t)^T B(x_h(t))\,(z_h(t) - u_h(t))\,dt \geq -C\,h.$$

We now wish to take limits $h \downarrow 0$ in the subsequence. Since $B(x_h(\cdot)) \to B(\widetilde{x}(\cdot))$ uniformly, all we need is for $\lambda_h \to \widetilde{\lambda}$ in $L^1(t_0, T)$ as $h \downarrow 0$. Away from the jumps, λ_h are uniformly Lipschitz. Supposing that the λ_h are bounded in, say, the space of functions of bounded variation, we can see that we would indeed obtain convergence of a further subsequence in this space. From the arguments above, $\widetilde{\lambda}(\cdot)$ is the adjoint function for the optimal control problem with the discontinuity. So, assuming uniformly bounded variation for λ_h, we get

$$\int_{t_0}^{T} \widetilde{\lambda}(t)^T B(\widetilde{x}(t))\,(\widetilde{z}(t) - u^*(t))\,dt \geq 0, \quad \text{for all measurable } z\colon [t_0, T] \to U,$$

as desired. While this is not a rigorous proof, it indicates that we should expect that the first order conditions hold for the limiting control function u^*.

7 A model problem and numerical results

We now investigate numerically the benefits of our theoretical results. We use our smoothing approach to investigate an optimal control problem whose discontinuous dynamics originates in the Coulomb friction. While the proofs of our theorems do not include the case with controls, the benefits of our analysis can be extended to that case as well.

Indeed, consider the problem

$$\begin{aligned} \min\nolimits_{u,x} \quad & g(x(T)) \\ \text{subject to} \quad & \dot{x} = f(x,u), \\ & x(0) = x_0, \\ & u(t) \in K, \end{aligned}$$

where K is a given convex set.

We construct the Lagrangian $\mathcal{L}(x,u,\lambda) = g(x(T)) - \int_0^T \lambda^T\,(\dot{x} - f(x,u))$. We compute its first-order variations with respect to feasible $\delta x(t)$ and $\delta u(t) \in \mathcal{T}_K(u(t))$:

$$\delta\mathcal{L} = \nabla g(x(T))\delta x(T) - \int_0^T \lambda^T\left(\dot{\delta x} - \nabla_x f(x,u)\delta x - \nabla_u f(x,u)\delta u\right).$$

Choosing the adjoint variable $\lambda(t)$ to satisfy

$$\dot{\lambda}(t) = -(\nabla_x f(x, u))^T \lambda(t), \ \lambda(T) = \nabla_x g(x(T))$$

as well as using integration by parts and $\delta x(0) = 0$, we obtain

$$\delta\mathcal{L} = \int\limits_0^T \lambda(t)^T \nabla_u f(x(t), u(t))\delta u(t)$$

for feasible $\delta u(t) \in \mathcal{T}_K(u(t))$.

Therefore, $\lambda(t)^T \nabla_u f(x(t), u(t))$ is the reduced gradient with respect to u. Using Theorems 3 and 2, we obtain that, if the problem has discontinuities and we use a smoothing approach with an explicit Euler time-stepping scheme with $h = O(\sigma^2)$, the reduced gradients for the smoothed problem approach the ones of the original problem, if $\nabla_u f(x, u)$ is continuous. Therefore, since the gradients converge, the divergence phenomena described at the beginning of this paper will not occur. Nonetheless, when solving our example we do not have to use the reduced gradient computed in this fashion; we have used it only to argue the convergence of the relevant gradients.

We have used the smoothing approach to compute optimal solutions for a crude approximation of a racing car model. This is a version of the "Michael Schumacher" problem described on CPNET, the Complementarity Problem Network [30]. The differential equations and constraints are different here from those in [30] in that aerodynamic drag is ignored here and the track considered here is a more complex S-bend instead of an ellipse.

The state space consists of a vector $\mathbf{x} \in \mathbb{R}^2$ denoting the position of the center of the vehicle, its velocity $\mathbf{v} \in \mathbb{R}^2$, and the angle in which the vehicle is pointing $\theta \in \mathbb{R}$. The controls consist of the throttle $a(t)$, which accelerates or decelerates the vehicle, and the steering control $s(t)$, which changes the vehicle orientation. The auxiliary functions used are $\mathbf{t}(\theta) = (\cos(\theta),\ \sin(\theta))$, the unit vector the vehicle is pointing in, and $\mathbf{n}(\theta) = (-\sin(\theta),\ \cos(\theta))$, a unit normal vector to $\mathbf{t}(\theta)$. The following differential equations are used:

$$\dot{\mathbf{x}} = \mathbf{v}, \tag{52}$$

$$\dot{\mathbf{v}} = a(t)\,\mathbf{t}(\theta) + F\,\mathbf{n}(\theta), \tag{53}$$

$$\dot{\theta} = s(t)\,(\mathbf{t}(\theta)^T\mathbf{v}), \tag{54}$$

$$F \in -\mu N \mathrm{Sgn}(\mathbf{n}(\theta)^T\mathbf{v}). \tag{55}$$

As usual, μ is the coefficient of friction and N is the normal contact force (assumed constant). These equations are clearly not sufficient to realistically describe a Formula 1 racing car. For example, "spin-outs" and "fish-tailing", where large uncontrolled angular velocities occur, cannot happen in this model. However, (52)–(55) do provide an interesting control system where friction appears in an essential way. Furthermore, slip is an essential characteristic of the solutions found for certain optimal control problems.

The initial conditions used were $\mathbf{x}(0) = 0$, $\mathbf{v}(0) = 0$, and $\theta(0) = 0$. These are the conditions for a vehicle initially at rest at the origin, pointing horizontally to the right.

The vehicle is constrained not to leave the track. This constraint introduces state constraints of the form $\mathbf{x}(t) \in C$, where $C \subset \mathbb{R}^2$ is the track. For our particular model problem, we take C to be

$$\{ (x, y) \mid |y - y_{cl}(x)| \le w/2 \}, \tag{56}$$

where w is the "width" of the track, and $\{ (x, y_{cl}(x)) \mid x \in \mathbb{R} \}$ is the curve of the centerline of the track. For an interesting but easily implementable system, we set

$$y_{cl}(x) = \begin{cases} \sin(x), & x \le \pi, \\ \pi - x, & \pi \le x \le 2\pi, \\ -\pi - \sin(x), & 2\pi \le x. \end{cases} \tag{57}$$

This generates a C^1, but not C^2, curve for the centerline.

The controls are subject to simple bounds constraints:

$$|a(t)| \le a_{\max} \quad \text{for all } t, \tag{58}$$

$$|s(t)| \le s_{\max} \quad \text{for all } t. \tag{59}$$

The objective function chosen was a combination of a penalty term for missing a target $\mathbf{x}_{tgt}$, and the time taken to reach the endpoint:

$$g(T, x(T)) = \alpha \|\mathbf{x}(T) - \mathbf{x}_{tgt}\|^2 + T. \tag{60}$$

7.1 Specific parameter values

The following default values were used:

- The penalty parameter for the final target was $\alpha = 10$.
- The target point was $\mathbf{x}_{tgt} = (3\pi, -\pi)^T$, which is on the center line of the track.
- The maximum acceleration and steering controls were $a_{\max} = 2$ and $s_{\max} = 2$.
- The maximum friction force was $\mu N = 4$.

7.2 Numerical results

The discretized optimal control problem was set up with AMPL modeling language [12] and solved with LOQO [37] under Linux. Note that the AMPL representation of the problem also included the discretization of the differential equations (explicit Euler); this meant that adaptive ODE solvers could not be applied. On the other hand, AMPL is able to provide LOQO with all the necessary gradient information. The baseline problem used a time step of $h = T/N$, with T the final time and N fixed at

Table 1 Objective, final time and algorithm performance for minimum time problem

σ	N	Objective	T	No. of iterations	CPU time
0.1	250	5.544654	5.53393	303	14.1
	500	5.549708	5.53877	531	53.3
	1,000	5.552177	5.54111	349	53.3
	2,000	5.553451	5.54239	420	227.0
0.05	500	5.590849	5.58285	1,501	153.5
	1,000	5.409497	5.39842	977	242.2
	2,000	5.409183	5.39813	643	300.8
0.025	1,000	5.353886	5.34289	1,240	184.5
	2,000	5.354256	5.34321	1,759	913.2
	4,000	5.354451	5.34341	1,368	1,552.8

1,000; the smoothing parameter was $\sigma = 0.1$. A number of runs were carried out with different values of N and σ. The results are shown in Table 1.

For the baseline problem, the value of T computed for this "minimal time" problem was $T = 5.541106358$, making $h \approx 5.5 \times 10^{-3}$. The final objective function value was 5.552177023, obtained in 349 iterations and taking 53.3 s of CPU time on a Pentium 4 running Linux. LOQO reported a dual objective function value of 5.552176961. While the objective function and the feasible region are highly nonconvex, this result does indicate that the objective function value is likely within about 6×10^{-8} of the value of a local minimum. The objective function value indicates that $\alpha \|\mathbf{x}(T) - \mathbf{x}_{tgt}\|^2 \approx 0.01107$; since $\alpha = 10$, this indicates that $\|\mathbf{x}(T) - \mathbf{x}_{tgt}\| \approx 0.03327$; the target is approached to high accuracy. Similar results are apparent for the other values of N and σ (see Table 1).

We note that LOQO, like most software for mathematical programming, can guarantee only that the computed solution is close to a local minimum; guaranteeing a global minimum without convexity is a computationally challenging task and there is no reason to expect that these optimal control problems do not have many local minima. However, the objective function values reported are remarkably consistent, and the controls used to achieve these local minima are also remarkably similar. This gives us some confidence that the computed objective values and controls are close to a global minimum for the true unsmoothed continuous time problem (52)–(57). The computed optimal trajectory for the baseline problem is shown in Fig. 4.

The computed optimal control functions, along with the normal and tangential velocities, are shown in Fig. 5. A complex maneuver takes place near $t = 3$; this is shown in more detail in Fig. 6.

From Fig. 4 we can see that the computed solution is somewhat complex, but has features similar to those expected from state constrained optimal control problems: the trajectory touches the boundary of the feasible region at a few points. The control functions, shown in Fig. 5 give a further indication of the complexity of the controls. However, the optimal acceleration control $a(t)$ is a pure bang–bang type (with control values lying on the boundary of the set of admissible control values). On the other hand, the steering control $s(t)$ is a bang–singular–bang control [3] where the control values lie in the interior of the set of possible control values for an interval in time.

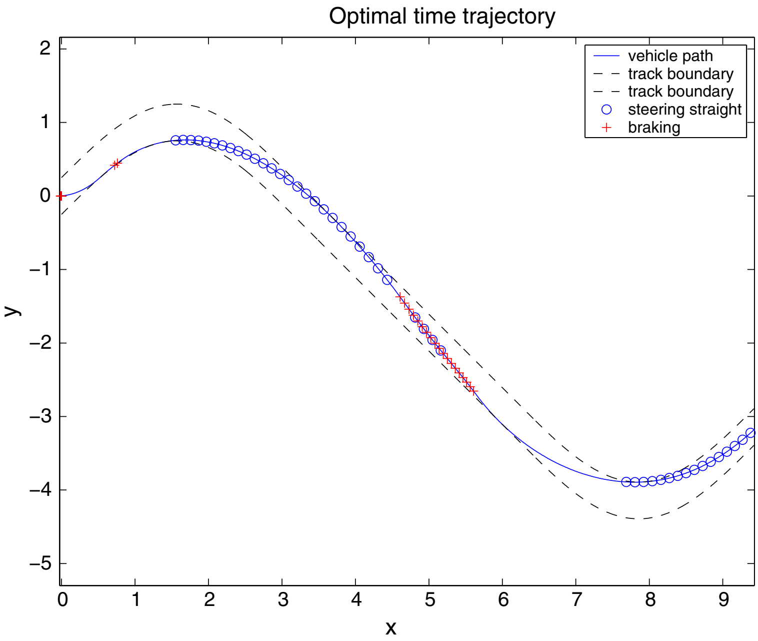


Fig. 4 Trajectory of "race-car"

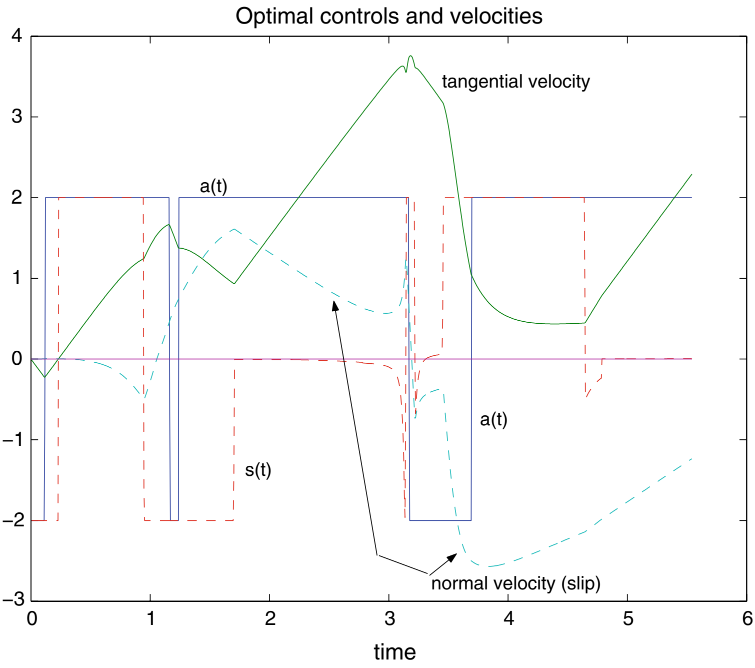


Fig. 5 Computed optimal control functions and velocities

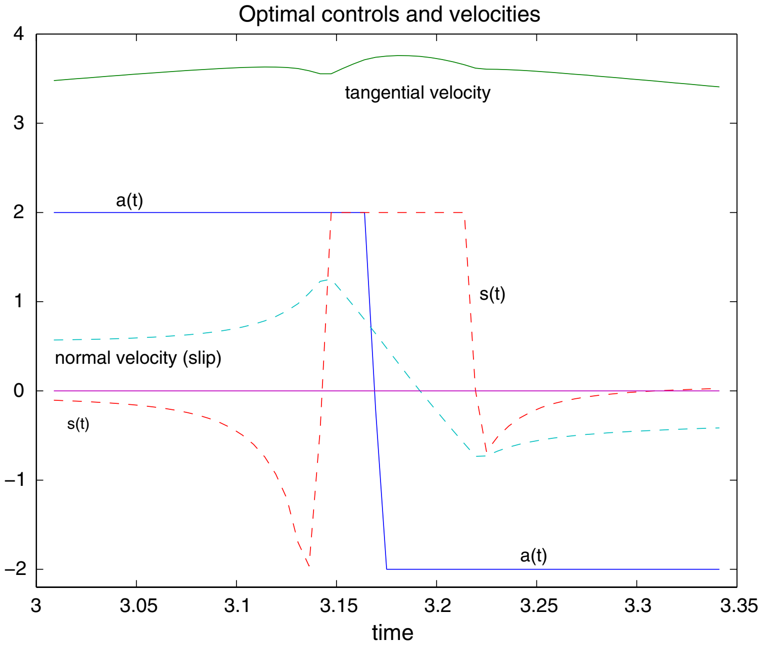


Fig. 6 Computed optimal control functions and velocities (zoom)

For the numerical problem, the "singular arc" for $s(t)$occurs for t roughly in the intervals (1.7, 3.5) and (4.7, 4.8).

It should also be noted that the discontinuity in the right-hand side, which occurs when the normal velocity or slip is zero, is met by the optimal trajectory in several places. Initially (from $t = 0$ to $t \approx 0.3$) the normal slip is zero, while the normal velocity passes through zero twice (once for $t \approx 1.1$ and again for $t \approx 3.2$). This means that the analysis for both the "crossing" and "trapped" cases are relevant for this problem.

7.3 Comparison with solutions for different N and σ

Different values of N and σ do result in some differences in the control functions and the velocities. These can be seen in Table 1, and also in Fig. 7. Most of the control and velocities are indistinguishable except for a few features.

For $\sigma = 0.1$ there is the initial reversing maneuver, which is not present for smaller values of σ. This is presumably because reducing the value of σ keeps the normal velocity closer to zero and enables the vehicle to make the first counter-clockwise turn without leaving the track. However, for $\sigma = 0.05$ there is a deceleration maneuver near $t = 0.35$ that appears to be of a bang-singular-bang type; for $\sigma = 0.025$ there is another deceleration maneuver at about the same time of a bang-bang type.

For $\sigma = 0.05$ and $N = 500$ the computed solution appears to be a local but not global minimum. Note that the objective function for this case is ≈ 5.59 compared

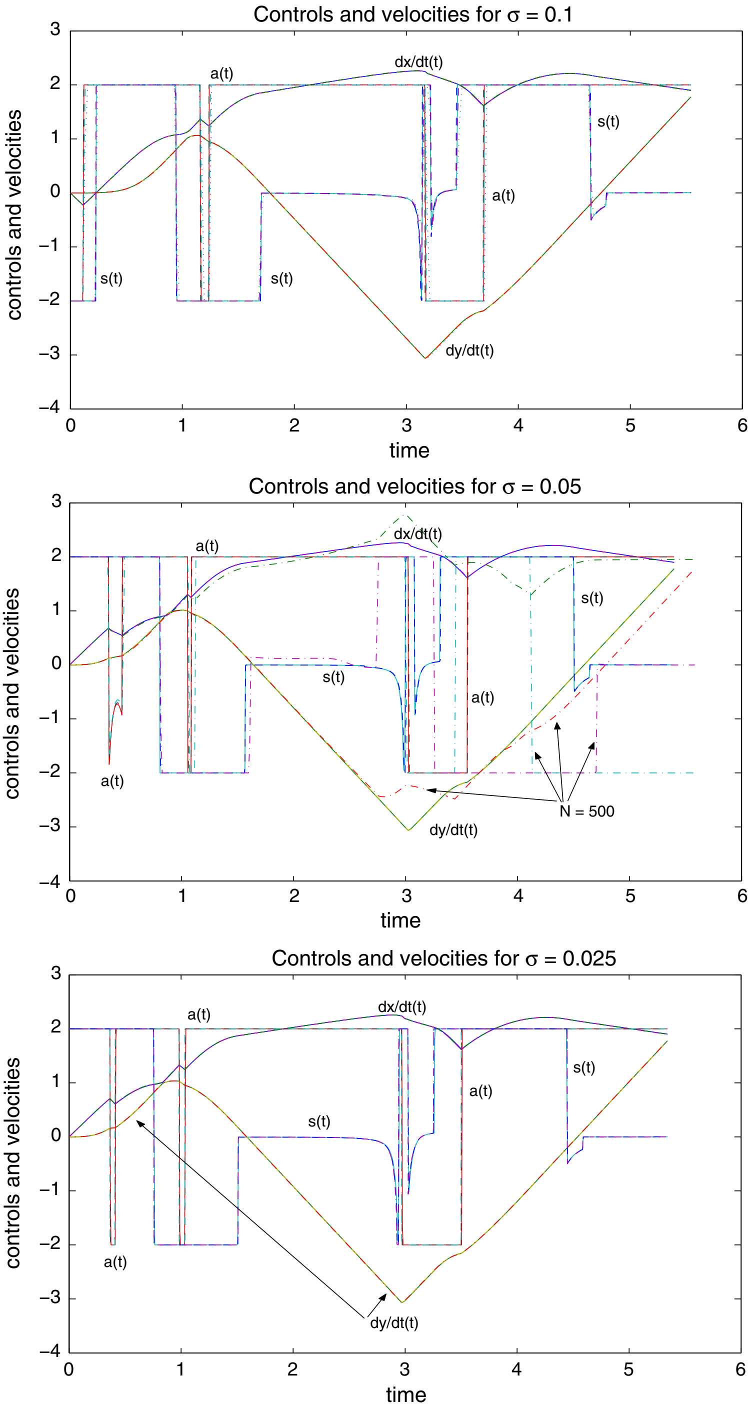


Fig. 7 Controls and velocities for varying N and σ ($\sigma = 0.1,\ 0.05,\ 0.025$)

with ≈ 5.41 for $N = 1{,}000$ and $N = 2{,}000$ with $\sigma = 0.05$, a difference of 3%. This solution (plotted with "dot-dash" lines) does not have the interesting steering maneuvers near $t = 3$ and near $t = 4.6$ and has quite different control functions near the finish line.

Otherwise the controls are qualitatively and quantitatively similar. Thus, it would seem that most of the maneuvers observed in the optimal controls are not artifacts but are truly part of the optimal strategy.

8 Conclusions

Optimal control problems where the dynamics includes discontinuous right-hand sides or differential inclusions can be handled both analytically and computationally by smoothing the right-hand side, at least when the right-hand side satisfies a one-sided Lipschitz condition. Furthermore, not smoothing the right-hand side, and computing the gradients of the discretized system directly, leads to incorrect gradient information with errors comparable to the size of the true gradients even with fully implicit discretization. For the computed gradients to converge to the true gradients, we need $h = o(\sigma^2)$, where h is the step size and σ is the smoothing parameter; this is sufficient under non degeneracy assumptions. Furthermore, we used the computed gradients with modern optimization software to compute optimal controls for moderately complex optimization problems with solutions that could not be easily predicted a priori. We expect that such results can be replicated for a large class of problems.

The usual adjoint equation or inclusion must be modified to allow for jumps in the adjoint variables if the trajectory crosses the discontinuity. This is quite different from the case of differential inclusions with Lipschitz right-hand sides [8,13,17], where an adjoint differential inclusion can be constructed and solved.

A "jump formula" for the adjoint variable in the case where the trajectory crosses a codimension-one discontinuity has been developed that uses only the right-hand side on opposite sides of the discontinuity and the normal vector of the discontinuity manifold. This jump formula opens up the possibility of accurately computing gradients by using a detect/locate/restart method of handling discontinuities in the forward ODE solver.

Acknowledgments This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-ENG-38. David Stewart was supported in part by NSF collaborative grant DMS-0139708. He would also like to thank the Argonne National Laboratory for their support and hospitality; part of this work was carried out while he was visiting Argonne National Laboratory.

References

1. Atkinson, K.E.: An Introduction to Numerical Analysis, 1st edn. Wiley, New York (1978)
2. Aubin, J.-P., Cellina, A.: Differential Inclusions: Set-Valued Maps and Viability Theory. Springer, Berlin (1984)

3. Bell, D.J., Jacobsen, D.H.: Singular Optimal Control Problems. Mathematics in Science and Engineering, vol. 117. Academic Press, New York (1975)
4. Brézis, H.: Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert. North-Holland Publishing, Amsterdam, North-Holland Mathematics Studies, No. 5. Notas de Matemática (50) (1973)
5. Clarke, F.H.: Optimal control and the true hamiltonian. SIAM Rev. **21**, 157–166 (1979)
6. Clarke, F.H.: Methods of Dynamic and Nonsmooth Optimization. CBMS–NSF Reg. Conf. Ser. #57. SIAM, Philadelphia (1989)
7. Clarke, F.H.: Nonsmooth Analysis and Optimization. SIAM, Philadelphia (1990). Originally published by the Canadian Mathematical Society (1983)
8. Clarke, F.H.: The maximum principle under minimal hypotheses. SIAM J. Control Optim. **14**(6), 1078–1091 (1976)
9. Driessen, B.J., Sadegh, N.: Minimum-time control of systems with Coulomb friction: near global optima via mixed integer linear programming. Optimal Control Appl. Methods **22**(2), 51–62 (2001)
10. Driessen, B.J., Sadegh, N.: On the discontinuity of the costates for optimal control problems with Coulomb friction. Optimal Control Appl. Methods **22**(4), 197–200 (2001)
11. Filippov, A.F.: Differential Equations with Discontinuous Right-Hand Side. Kluwer, Dordrecht (1988)
12. Fourer, R., Gay, D.M., Kernighan, B.W.: AMPL: a Modeling Language for Mathematical Programming, 2nd edn. Brooks/Cole, Thomson Learning, Pacific Grove (2003)
13. Frankowska, H.: Adjoint differential inclusions in necessary conditions for the minimal trajectories of differential inclusions. Ann. Inst. H. Poincaré Anal. Non Linéaire **2**(2), 75–99 (1985)
14. Frankowska, H.: The maximum principle for a differential inclusion problem. In: Analysis and Optimization of Systems, Part 1 (Nice, 1984), pp. 517–531. Springer, Berlin (1984)
15. Frankowska, H.: Le principe de maximum pour une inclusion différentielle avec des contraintes sur les états initiaux et finaux. C. R. Acad. Sci. Paris Sér. I Math. **302**(16), 599–602 (1986)
16. Frankowska, H.: The maximum principle for an optimal solution to a differential inclusion with end points constraints. SIAM J. Control Optim. **25**(1), 145–157 (1987)
17. Frankowska, H., Kaŝkosz, B.: A maximum principle for differential inclusion problems with state constraints. Syst. Control Lett. **11**(3), 189–194 (1988)
18. Galán, S., Feehery, W.F., Barton, P.I.: Parametric sensitivity functions for hybrid discrete/continuous systems. Appl. Numer. Math. **31**(1), 17–47 (1999)
19. Gamkrelidze, R.V.: Principles of Optimal Control Theory. Plenum Press, London (1978). Original in Russian (1975)
20. Glowinski, R., Kearsley, A.J.: On the simulation and control of some friction constrained motions. SIAM J. Optim. **5**(3), 681–694 (1995)
21. Kastner-Maresch, A.: Diskretisierungsverfahren zur Lösung von Differentialinklusionen. PhD thesis, Universität Bayreuth (1990)
22. Kastner-Maresch, A.: Implicit Runge–Kutta methods for differential inclusions. Numer. Funct. Anal. Optim. **11**, 937–958 (1990)
23. Kim, T.-H., Ha, I.-J.: Time-optimal control of a single-DOF mechanical system with friction. IEEE Trans. Autom. Control **46**(5), 751–755 (2001)
24. Lipp, S.C.: Brachistochrone with Coulomb friction. SIAM J. Control Optim. **35**(2), 562–584 (1997)
25. Maso, G.D., Rampazzo, F.: On systems of ordinary differential equations with measures as controls. Differ. Integral Equ. **4**(4), 739–765 (1991)
26. Outrata, J., Kočvara, M., Zowe, J.: Nonsmooth Approaches to Optimization Problems with Equilibrium Constraints. Nonconvex Optimization and Its Applications, vol. 28. Kluwer, Dordrecht (1998)
27. Pontryagin, L.S., Boltjanskij, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: The Mathematical Theory of Optimal Processes. Interscience, New York (1962). Original in Russian (1956)
28. Stewart, D.E.: Numerical methods for friction problems with multiple contacts. J. Aust. Math. Soc. Ser. B **37**(3), 288–308 (1996)
29. Stewart, D.: A high accuracy method for solving ODEs with discontinuous right-hand side. Numer. Math. **58**(3), 299–328 (1990)
30. Stewart, D.E.: The "Michael Schumacher" problem. http://www.cs.wisc.edu/cpnet/cpnetmeetings/iccp99/race-car/race-car.html. Accessed June 1999
31. Sussmann, H.J.: Optimal control of nonsmooth systems with classically differentiable flow maps. In: Proceedings of the Sixth IFAC Symposium on Nonlinear Control Systems (NOLCOS 2004), Stuttgart (2004)

32. Taubert, K.: Differenz Verfahren für gewöhnliche Anfangswertaufgaben mit unstetiger rechte Seite. In: Dold, A., Eckmann, B. (eds.) Numerische Behandlung nichtlinearer Integrodifferential- und Differentialgleichungen, pp. 137–148. Lecture Notes Series, vol. 395 (1974)
33. Taubert, K.: Differenzverfahren für Schwingungen mit trockener und zäher Reibung und für Regelungssysteme. Numer. Math. **26**, 379–395 (1976)
34. Taubert, K.: Converging multistep methods for initial value problems involving multivalued maps. Computing **27**, 123–136 (1981)
35. Tolsma, J.E., Barton, P.I.: Hidden discontinuities and parametric sensitivity calculations. SIAM J. Sci. Comput. **23**(6), 1861–1874 (electronic) (2002)
36. van Willigenburg, L.G., Loop, R.P.H.: Computation of time-optimal controls applied to rigid manipulators with friction. Int. J. Control **54**(5), 1097–1117 (1991)
37. Vanderbei, R.J.: LOQO User's Manual, Version 4.05. Princeton University, Operations Research and Financial Engineering Department, October 2000
38. Ventura, D., Martinez, T.: Optimal control using a neural/evolutionary hybrid system. In: Proceedings of the International Joint Conference on Neural Networks, pp. 1036–1041, May (1998)