PITFALLS IN APPLYING OPTIMAL CONTROL TO DYNAMICAL SYSTEMS: AN OVERVIEW AND EDITORIAL PERSPECTIVE

URSZULA LEDZEWICZ
Institute of Mathematics
Lodz University of Technology, 90-924 Lodz, Poland
and Dept. of Mathematics and Statistics
Southern Illinois University Edwardsville
Edwardsville, Il, 62026-1653, USA

HEINZ SCHÄTTLER
Dept. of Electrical and Systems Engineering
Washington University
St. Louis, Mo, 63130, USA

(Communicated by Yuan Lou)

Abstract. In recent years, an increasing number of papers have been published (and many more submitted for publication) in which optimal control theory is superficially applied to specific problems, especially from the biological and health sciences, but also many other fields. A lack of understanding of what it actually means to solve an optimal control problem—complex infinite-dimensional optimization problems—often leads to heavily overblown claims about optimality of solutions. In this editorial, a critical assessment of these efforts is given.

1. Introduction. Optimal control problems are dynamic optimization problems in which the state of a system is to be guided in an “optimal” way from a prescribed initial condition into a set of allowed terminal states while satisfying applicable constraints. As a scientific field in mathematics and engineering, optimal control theory came to fruition in the 1960s with the development of space exploration. While its historical roots lie in the calculus of variations [26], it were the spectacular successes of putting satellites into orbit and of the space program in general that established the field as an independent discipline [9]. One omnipresent example of a solution to an optimal control problem is the so-called linear-quadratic regulator [14] which forms the basis for the regulation of dynamical systems around some reference solution. Its applications are far ranging from autopilots on commercial aircraft to chemical process control. Early applications to problems in economics such as optimal consumption-investment strategies [19] and the Black and Scholes formula [3] for the pricing of a European option helped earn those researchers Nobel prizes in economics.

2020 Mathematics Subject Classification. Primary: 49-01, 49K15; Secondary: 49L20, 93C10.
Key words and phrases. Optimal control, necessary conditions for optimality, multiplier rules, sufficient conditions for optimality, singularities, regular synthesis.
* Corresponding author: Urszula Ledzewicz.
Because of the great generality of the fundamental problem formulation—transfer the state of a dynamical system from a given initial condition into a set of desired terminal conditions subject to constraints while at the same time minimizing an objective associated with the motion and/or a penalty on the terminal state—applications of optimal control to practical areas are omnipresent. This is also attested to by a vast amount of current publications in the medical and health sciences. By now, for probably any scientific discipline papers exist where optimal control is applied to problems originating in that field.

Optimal control problems, however, are infinite-dimensional optimization problems which generally are difficult to solve. As it is to be expected, often, if not typically, they have a large number of stationary points and local minima. If the dimension is low and the dynamics is manageable, possibly because of underlying symmetries, as, for instance it was the case in the early applications in economics mentioned above, sufficient conditions for optimality in the form of a solution to the Hamilton-Jacobi-Bellman (HJB) equation can sometimes be given and thus a complete solution of the problem has been achieved. This, however, is the exception rather than the rule. In the vast majority of papers in which optimal control is applied to ‘solve’ some practical application, such complete solutions are rarely achieved. Instead, in many publications and submissions the mere computation of an extremal (which amounts to nothing more than what would be computing a stationary point for a finite-dimensional minimization problem) often is considered solving the problem and “victory is declared”. In various papers, a lack of understanding what it actually entails to solve an optimal control problem coupled with superficial arguments that supposedly ‘solve’ the problem leads to heavily overstated claims about optimality of solutions. In this short note, we want to clarify what really needs to be done to solve an optimal control problem while, at the same time, we highlight common pitfalls in applying both theory and numerical methods.

In Section 2, in order to have concrete equations at hand for our discussions, we formulate a standard, widely applicable model for an optimal control problem. While we do not discuss modeling aspects of the dynamics in this paper, we do comment on possible choices for the objective functional in Section 3. Section 4 then gives an overview of the steps that need to be undertaken in solving an optimal control problem with an emphasis on conceptual aspects. Leading from the necessary conditions of the Pontryagin maximum principle [22] we proceed to higher order necessary and sufficient conditions for local optimality (Jacobi conditions and/or perturbation feedback control) to the sufficient conditions for global optimality of field theory (regular synthesis [4, 5, 21]) and dynamic programming, both of which essentially construct solutions to the Hamilton-Jacobi-Bellman equation. We then use models for control-affine systems with linear ($L_1$-type) and quadratic ($L_2$-type) objectives in the control in Section 5 to highlight the most significant and common pitfalls in solving optimal control problems. These arise from the simple fact that the necessary conditions for optimality generally allow for multiple solutions. We only briefly discuss aspects of the numerical solution of optimal control problems in Section 6.

2. Standard model of an optimal control problem. The formulation of a typical optimal control problem consists of the following elements: (i) the dynamics which describes the evolution of the underlying system in time, (ii) a class of functions (the admissible controls) which model actions that are allowed to
influence/control the system, (iii) additional constraints which need to be satisfied both during the time-evolution of the system and/or at the terminal point, and (iv) an objective functional which, through judicious choice of the control, is sought to be minimized or maximized.

For sake of specificity, and since this formulation covers a vast majority of applications, we consider the following framework: the state-space $M$ is an open subset of $\mathbb{R}^n$ (more generally, a manifold), the control set $U$ is a compact subset of $\mathbb{R}^m$, and the dynamics is described by an ordinary differential equation of the form

$$\dot{x} = f(x,u), \quad x(0) = x_0,$$  \hfill (1)

with $f : M \times U \to \mathbb{R}^n$, $(x,u) \mapsto f(x,u)$, sufficiently smooth, control-dependent vector fields $f(\cdot,u)$, $u \in U$, on $M$. Controls $u$ are Lebesgue-measurable functions $u : [0,T] \to U$, $t \mapsto u(t)$, defined over some finite interval $[0,T]$ where $T$ can be fixed or free, which take values in the control set $U$ (a.e.). It is implicitly assumed that for relevant initial conditions $x_0 \in M$ the solution to the corresponding initial-value problem is unique and exists on the full interval $[0,T]$. The pair $(x,u)$ is called a controlled trajectory. Among a variety of possible additional constraints, here we only consider constraints at the terminal point $x(T)$ in the form of a target set into which the controls need to steer the system. Specifically, we assume that

$$(T,x(T)) \in N = \{(t,x) \in [0,\infty) \times M : \Psi(t,x) = 0\}$$  \hfill (2)

where $\Psi : [0,\infty) \times M \to \mathbb{R}^{n+1-k}$ is such that the matrix $D\Psi$ of the partial derivatives with respect to $(t,x)$ is of full rank $n + 1 - k$ everywhere on $N$. This guarantees that $N$ is a $k$-dimensional embedded submanifold of the combined time-state-space.

The class $\mathcal{U}$ of admissible controls $u$ consists of all controls (Lebesgue measurable functions) $u : [0,T] \to U$ for which the solution $x$ of the corresponding dynamics exists on all of $[0,T]$ and satisfies $(T,x(T)) \in N$. Further constraints such as state-space or mixed control-state space constraints may arise in specific formulations, but adding them here would make the mathematical treatment more challenging and the model as it is formulated here is more than adequate for typical applications.

The last, but not the least ingredient in the formulation of an optimal control problem is the specification of the objective functional which we take as a weighted average of a penalty term $\varphi : N \to \mathbb{R}$ at the terminal point and the integral of a Lagrangian

$$L : M \times U \to \mathbb{R}, \quad (x,u) \mapsto L(x,u),$$

which is taken as a measure of the performance of the system during its evolution in time. This function could be identically zero if this behavior is of lesser importance. The objective to be minimized is then given in what is called a Bolza-formulation in optimal control as

$$J : \mathcal{U} \to \mathbb{R}, \quad u \mapsto J(u) = \int_0^T L(x(s),u(s))ds + \varphi(T,x(T)).$$  \hfill (3)

The optimal control problem $[OC]$ can then succinctly be formulated as to

minimize the objective $J = J(u)$ over all admissible controls $u \in \mathcal{U}$.

We emphasize that the variables of choice are functions defined on the interval $[0,T]$, i.e., this is a heavily infinite-dimensional optimization problem.
3. Pitfalls in modeling: Choice of the objective functional. As in any application of mathematical tools to a real problem, the question about the accuracy of the model is important. Approximations will be necessary, be it that the underlying dynamics is not fully understood or simply to limit the complexities of the problem. Clearly, to what extent the dynamics is a realistic representation of the underlying problem is an essential question, but it needs to be discussed on a case-by-case basis. Here we are interested in the optimization procedure and thus we assume the dynamics as given. The controls which realistically can be applied are generally clear and there rarely is an issue about what the control set $U$ should be. The requirement that controls be Lebesgue measurable, on the other hand, in its generality may appear unrealistic, but it is essential for certain compactness properties that are needed to prove the existence of an optimal control. Whether an optimal control thus represents a practically realistic function becomes an after thought once the problem has been solved.

The one item, however, which is equally important in the formulation of the optimal control problem, is the objective $J$ to be minimized and here there exists great freedom in choosing the functional forms for $\varphi$ and, especially, the Lagrangian $L$. The role of weights and scaling of the variables and controls is one important aspect. For example, even if the main interest is on the final outcome (e.g., minimization of the size of a tumor), care has to be taken that the solution does not lead to unacceptable behavior in between (e.g., a temporary spike in the tumor size) \cite{23}. Avoiding such features may require a judicious choice of weights represented by the functions $\varphi$ and $L$ and, at times, the mathematically optimal solution simply does not make sense because of ill-chosen weights. Appropriate scalings of the variables (e.g., tumor volumes) and the controls (which, using symmetries, often are normalized to $u_i \in [0,1]$) can sometimes avoid non-sensical or, what is more common, obvious and thus rather uninteresting outcomes.

A more important aspect is that the formulation of the objective $J$ should mostly be guided by the demands of the underlying problem and only to a lesser extent by the mathematical simplicity its formulation seemingly affords. All too often, and without any justification, the latter seems to determine the choice of the Lagrangian $L$. For example, if the control $u$ represents the dose-rate or concentration of some drug, then there is a clear pharmacological meaning to the integral $\int_0^T u(t)dt$ while the integral $\int_0^T u^2(t)dt$ is a rather meaningless quantity without any direct interpretation. Furthermore, if the controls are normalized to lie in $[0,1]$, taking the square distorts the amounts of agents used favoring lower doses. Hence such a conclusion is not because of optimality, but this was a bias a priori built into the modeling. Similarly, if $u$ is connected with some linear cost, then so is the integral $\int_0^T u(t)dt$, but a term of the form $\int_0^T u^2(t)dt$ has no viable interpretation: $u^2$ is not a meaningful quantity. And no matter how often the argument that some costs are nonlinear in the control is repeated, while true, this certainly does not justify to choose the dependence as a quadratic term. Generally, this is merely done because seemingly it offers mathematical simplicity in the solution. This need not even be the case and we shall discuss the mathematical implications of such a choice further below once we have formulated the necessary conditions for optimality of the Pontryagin maximum principle \cite{22}.

4. Optimal control: An overview. In this section, we discuss, with an emphasis on conceptual aspects, the steps that need to be undertaken in solving an optimal
control problem proceeding from necessary to sufficient conditions for optimality. In order to discuss the mathematical structures it becomes necessary to recall the fundamental necessary conditions for optimality.

4.1. The first-order necessary conditions of the Pontryagin minimum principle. An important role in the solution of an optimal control problem is played by the (control) Hamiltonian function $H$ which adjoins the dynamics to the Lagrangian $L$ with a multiplier $\lambda \in (\mathbb{R}^n)^*$ which we write as a row vector. Ignoring abnormal extremals for the sake of this editorial, the Hamiltonian is defined as

$$H = H(\lambda; x, u) = L(x, u) + \langle \lambda, f(x, u) \rangle.$$  \hspace{1cm} (4)

The Pontryagin Maximum principle [22] (for more recent references, see, e.g., [6, 8, 18, 23]) gives the fundamental first-order necessary conditions for optimality. Suppose $u_*$ is an optimal control for the problem [OC] defined over the interval $[0, T]$ with corresponding trajectory $x_*$. Then there exists an absolutely continuous co-vector $\lambda : [0, T] \rightarrow (\mathbb{R}^n)^*$ which is a solution of the adjoint equation

$$\dot{\lambda} = -\frac{\partial H}{\partial x}(\lambda; x_*(t), u_*(t)) = -\frac{\partial L}{\partial x}(x_*(t), u_*(t)) - \lambda \frac{\partial f}{\partial x}(x_*(t), u_*(t)),$$ \hspace{1cm} (5)

such that\footnote{The nontriviality of the multiplier is guaranteed as we only consider normal extremals.} the covector $(H, -\lambda + \partial \phi/\partial x)$ is orthogonal to the manifold $N$ at the terminal point and the optimal control $u_*$ minimizes the Hamiltonian $H$ pointwise along $\lambda$ and the controlled trajectory $(x_*, u_*)$ over the control set $U$, i.e.,

$$H(\lambda(t), x_*(t), u_*(t)) = \min_{v \in U} H(\lambda(t), x_*(t), v).$$ \hspace{1cm} (6)

It is especially this last condition in which the power of this result lies as it reduces the minimization over $U$, an infinite-dimensional function space, to pointwise finite-dimensional minimizations over the control set $U \subset \mathbb{R}^m$. In particular, the result asserts the existence of this minimum. Furthermore, the orthogonality condition at the terminal point is equivalent to the existence of a multiplier $\nu \in (\mathbb{R}^{n+1-k})^*$ such that

$$H(\lambda(T), x_*(T), u_*(T)) + \nu \frac{\partial \Psi}{\partial t}(T, x_*(T)) = 0$$ \hspace{1cm} (7)

and

$$\lambda(T) = \frac{\partial \phi}{\partial x}(T, x_*(T)) + \nu \frac{\partial \Psi}{\partial x}(T, x_*(T)).$$ \hspace{1cm} (8)

Equation (7) restricts the terminal time (if free) while equation (8) gives conditions on the terminal value of the multiplier. The necessary conditions for optimality thus form a 2-point boundary value problem for the state and multiplier with the initial condition on the state given and restrictions on the time, state and multiplier at the terminal point.

A controlled trajectory $(x_*, u_*)$ for which there exists a multiplier $\lambda$ such that all these conditions are satisfied is called a (normal) extremal (controlled trajectory). On the level of minimizing a function $F$ in $\mathbb{R}^n$, extremals simply correspond to stationary points, but not more. The necessary conditions of the maximum principle are a multiplier rule which, through the solution of 2-point boundary value problems, identify all possible extremals in the joint state$\times$multiplier space (the cotangent bundle of the manifold that is the state-space). There are two steps in the analysis of these necessary conditions which can lead to multiplicities, the first one obvious, the second one more subtle:
1. the minimizer of the Hamiltonian $H$ over the control set $U$ need not be unique;
2. even if it is, then the solution to the associated 2-point boundary problem can
   have multiple, even infinitely many solutions.

In order to solve an optimal control problem, it is necessary to compute and compare
the values which all extremals generate for the objective. This requires to project
the extremals into the state-space (i.e., eliminate or find the optimal multiplier).
This procedure carries with it all the mathematical difficulties caused by multiple
solutions and singularities. Clearly, this is a challenging task.

4.2. Higher-order necessary conditions for optimality. Like for the mini-
mization of functions, higher-order necessary conditions for optimality help to elim-
inate extremals from consideration. For an optimal control problem these are the
analogues of the Jacobi-condition from the calculus of variations. The analogue of
the Legendre-condition is already superseded by the much stronger minimization
condition (6). For particular problems (see Section 5.1 below), however, there are
the so-called generalized Legendre-Clebsch conditions [9] for the optimality of sin-
gular controls which give independent additional high-order necessary conditions
for optimality. The Jacobi conditions, both as necessary and sufficient conditions
for local optimality, are formulated in terms of conjugate points (i.e., nontrivial sol-
solutions to the Jacobi equation). In a nondegenerate setting these computations can
be related to finite explosion times for solutions of matrix Riccati-differential equa-
tions in constructions known in the engineering literature [9] as sweep method or
perturbation feedback control. The only point we wish to convey here is that there
exists a well-established theory of second-order necessary conditions for optimality
which can be used to test further the optimality of a particular extremal.

4.3. Sufficient conditions for local and global optimality. Like in the cal-
culus of variations, the absence of conjugate points allows to embed a particular
extremal into a local field of extremals. This, and modulo some piecewise regu-
ularity conditions, is the construction of a parameterised family of extremals, i.e.,
state-multiplier pairs $(x, \lambda)$, which contains the reference extremal as a member
and is such that the projection $\pi$ from the cotangent bundle into the state space,
$\pi : (x, \lambda) \mapsto x$, is $1 - 1$. This is the analogue of the strengthened Jacobi condi-
tion from the calculus of variations and it allows to assert the local optimality of a
particular extremal. These arguments are developed, for instance, in our textbook
[23]. For particular problem formulations (e.g., when the matrix of the second par-
tial derivatives $\frac{\partial^2 H}{\partial u^2} (\lambda(t), x_*(t), u_*(t))$ is non-singular along the reference extremal)
there exist relatively simple algorithmic procedures to test this. The perturbation
feedback control algorithms of the engineering literature [9] relate the existence of
conjugate points to the existence of finite explosion times for a matrix Riccati differ-
etial equation whose coefficients are given by second derivatives of the hamiltonian
$H$ evaluated along the reference extremal. Equivalently, if this matrix Riccati differ-
etial equation has a solution on the full closed interval $[0, T]$ (and some other,
lesser regularity conditions are met), then there are no conjugate points along the
reference extremal and it is possible to embed the reference into a local field of
extremals. Thus this extremal is a local minimum. The formal calculations in [9]
have been put on a sound theoretical basis in [23, 2].
Constructing a field of extremals ultimately is related to dynamic programming techniques and solutions to the Hamilton-Jacobi-Bellman equation,

$$\frac{\partial V}{\partial t}(t, x) + \min_{u \in U} \left\{ \frac{\partial V}{\partial x}(t, x)f(x, u) + L(x, u) \right\} \equiv 0,$$

for the value function $V = V(t, x)$ of the optimal control problem, i.e., $V(t, x)$ is the infimum of all possible values that can be realized if the initial condition is given by $x$ at time $t$, $0 \leq t \leq T$: $V(t, x) = \inf_{u \in U(t, x)} J(u)$. This equation is the combination of a minimization problem over the control set $U \subset \mathbb{R}^m$ with what formally is a first-order partial differential equation (PDE). Once the minimization problem is solved, however, generally highly nonlinear PDEs arise which need to be solved numerically (and thus for specific parameter values for the dynamics and weights in the objective). Another alternative is to use the method of characteristics properly adjusted from first order PDEs to the optimal control problem (e.g., see [1, 23]). In this approach, the value function is computed precisely through a full analysis of extremals. This also is the contents of the construction of a regular synthesis of optimal controlled trajectories. This procedure was initially proposed and developed by Boltyansky [4, 5] with rather stringent regularity assumptions. Since then many of these have been relaxed [21] and we refer to Section 6.3 of our textbook [23] for a relatively simple exposition of the main result. Such a synthesis of extremal controlled trajectories or, equivalently, the construction of a global field of extremals, gives the global and hence true optimality of extremals. While its construction is a highly non-trivial endeavor, for low-dimensional systems this is not impossible [16, 23].

5. Pitfalls in applying the theory. We shall be more specific about the problems of solving optimal control problems focusing on multiplicities related to the issues (1) and (2) raised in Section 4.1. In order to make our main points, we consider systems whose dynamics is control-affine, i.e., of the form

$$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i.$$  

This, indeed, is a common form for the dynamics in many applications where the control variables $u = (u_1, \ldots, u_m)^T$ represent independent outside modes of intervention. In a vast majority of applications of optimal control to practical problems the underlying dynamics has such a structure. The vector field $f$, called the drift, represents the dynamics of the uncontrolled system\(^2\) and the vector fields $g_i$, the control vector fields, model the effects which the $i$th control has on the system. Typically the control set $U$ is a compact interval, $U = [u_{i_{\text{min}}}, u_{i_{\text{max}}}] \times \ldots \times [u_{m_{\text{min}}}, u_{m_{\text{max}}}]$, so that there exists a bounded range of action for each specific control independent of the values of the other controls. For such systems, it makes a significant difference mathematically whether the Lagrangian $L$ is chosen as an affine $L_1$-type function of the controls or as a quadratic $L_2$-type function.

5.1. Problems with a Hamiltonian function $H$ which is affine in the controls $u$. The key to analyzing optimal controls is the minimization condition (6).
If the Lagrangian function is taken of the form
\[ L_1(x, u) = L_0(x) + \sum_{i=1}^{m} L_i(x)u_i, \] (11)
then the minimization of \( H \) over the control set \( U \) splits into \( m \) scalar minimization problems which, in principle, are easily solved as follows:
\[ u_i^*(t) = \begin{cases} u_{i_{\text{min}}} & \text{if } \Phi_i(t) > 0, \\ ? & \text{if } \Phi_i(t) = 0, \\ u_{i_{\text{max}}} & \text{if } \Phi_i(t) < 0, \end{cases} \] (12)
where
\[ \Phi_i(t) = L_i(x_*(t)) + \langle \lambda(t), g_i(x_*(t)) \rangle \] (13)
denotes the so-called switching function of the problem. If this function vanishes, the control a priori is not determined by the minimum condition and in principle any value of the control set could be optimal. This ambiguity highlights the issue (1) when the minimizer of the hamiltonian \( H \) is not unique. However, this does not imply that it does not matter what this value is and that this value can be chosen arbitrarily. For a problem of this type the possibility of what are called singular controls arises which are controls \( u_i \) defined over an open interval \( I \) where the switching function \( \Phi_i \) vanishes identically. Generally these are then computed by differentiating the switching function along the dynamics until the variable \( u_i \) arises for the first time. Then solving for \( u_i \) in fact gives specific formulas which a singular control must obey. For low-dimensional system, the corresponding trajectories typically lie on thin sets, i.e., manifolds of positive codimension, and thus are difficult to locate with numerical methods alone. Singular controls do not arise if the derivative of the switching function \( \dot{\Phi}_i \), \( \dot{\Phi}_i(t) \), does not vanish whenever \( \Phi_i(t) = 0 \).

In this case the control \( u_i \) switches between its extreme values at any zero of the switching function and such a control is called a bang-bang control.

For problems with a control-affine hamiltonian function \( H \) it is necessary to analyze all possible concatenations between bang and singular controls which the necessary conditions of the maximum principle allow. If singular controls can be excluded from optimality (typically through higher-order necessary conditions for optimality such as the generalized Legendre-Clebsch or Goh condition), then there exist simple algorithmic procedures to compute bang-bang extremals and to check their local optimality through second-order sufficient conditions. For example, this holds for cell-cycle specific models for cancer chemotherapy [27, 28, 29, 23]. Pure bang-bang controls, however, are rather the exception than the norm for a nonlinear dynamics. Singular controls are natural candidates for optimality and often form the essential part of the optimal solutions that determines the optimal synthesis [16].

In such a case, the corresponding singular arcs are limits of bang-bang extremal trajectories with an increasing number of switchings, i.e., there exists a huge number of (both locally optimal and non-optimal) extremals for such problems. This raises doubts about any claims of optimality of specific, numerically computed bang-bang extremals with a large number of switchings for such systems.

The analysis of possible optimal concatenation sequences of bang and singular arcs is generally quite involved (e.g., see [16, 17]) and for this reason all too often researchers shy away from \( L_1 \)-type formulations even if these would be the most appropriate ones for the problem under consideration. Rather, Lagrangian functions \( L \) are chosen which are quadratic in the controls as this resolves the multiplicities
that arise in the minimization of the Hamiltonian $H$ over the control set. But this does not resolve the second issue (2) and still the issue of multiple extremals needs to be addressed.

5.2. Problems with a Hamiltonian function $H$ which is quadratic in the controls $u$. If the Lagrangian function $L$ is chosen in the form

$$L_2(x, u) = L_0(x) + \sum_{i=1}^{m} L_i(x) u_i^2,$$

then, as for an $L_1$-type objective (11), minimizing $H$ is simple. For sake of argument (and since this is the most common case considered) suppose the functions $L_i$ are positive. In this case the Hamiltonian $H$ is strictly convex in $u_i$ and the minimizing control is unique. The stationary point is given by

$$\tilde{u}_i(t) = -\frac{\langle \lambda(t), g_i(x_*(t)) \rangle}{2L_i(x_*(t))}$$

and depending on whether this point lies in the control interval or not, the control which minimizes $H$ is given by

$$u_i^*(t) = \begin{cases} u_{\text{min}}^i & \text{if } \tilde{u}_i(t) \leq u_{\text{min}}^i, \\ \tilde{u}_i(t) & \text{if } u_{\text{min}}^i \leq \tilde{u}_i(t) \leq u_{\text{max}}^i, \\ u_{\text{max}}^i & \text{if } u_{\text{max}}^i \leq \tilde{u}_i(t). \end{cases}$$

(16)

This can be succinctly expressed in the commonly used form

$$u_i^*(t) = \max\{u_{\text{min}}^i, \min\{\tilde{u}_i(t), u_{\text{max}}^i\}\}$$

which also shows that the minimizing control is continuous.

The problem, however, is far from being solved by this formula. Having a Hamiltonian function $H$ which is quadratic in the controls does not make the optimal control problem a convex minimization problem which would have a unique solution. We refer, for example, to Lectures 15 and 16 in Girsanov’s notes [12] for a brief discussion of when an optimal control problem is a convex optimization problem.

An in depth analysis can be found in the textbook by Ioffe and Tikhomirov [13]. Contrary to the belief of many authors which use this approach, formula (17) does not define an optimal control: the stationary point $\tilde{u}_i(t)$ depends on the multiplier $\lambda(t)$ which, together with the state $x_*(t)$ is only given as the solution to a 2-point boundary value problem. There simply may exist multiple solutions and it is easy to give mathematical examples when this happens [10, 11]. Geometrically, the formula (17) determines the control $u_*$ only as a function in the cotangent bundle, $u_* = u_*(x, \lambda)$, while the optimal control $u_*$ lives in the state-space only, i.e., $u_*=u_*(x)$. Thus the projections of multiple solutions in the cotangent-bundle into the state-space need to be considered and various kinds of singularities (shock waves in the solution to the Hamilton-Jacobi-Bellman equation) can arise. A simple test of second-order conditions for local optimality, which is not too elaborate to carry out—because of the quadratic structure in the controls the hamiltonian function $H$ is strictly convex in the control and perturbation feedback control algorithms, the aforementioned Jacobi equations, easily apply and only require to integrate a small number of ODEs along the reference extremal [25]—would at least guarantee the

\[3\] Typically, the dynamics itself already forms a highly non-convex constraint. More or less, the function $f = f(x, u)$ which describes the dynamics should be convex in both variables.
local optimality of a particular, numerically computed extremal. Regretfully, this is rarely done.

6. Pitfalls in using numerical methods.

6.1. Shooting methods and multiple extremals. Standard ways to compute extremals use shooting methods to solve the 2-point boundary value problem which is defined by the necessary conditions for optimality of the maximum principle, i.e., the differential equations (1) and (5) with the control satisfying the minimum condition (6), initial condition $x(0) = x_0$, and terminal conditions imposed (possibly) both on the state and the multiplier by the transversality conditions (7) and (8). In these procedures, an initial guess is taken for the value $\lambda(0)$ of the multiplier at the initial time and the solutions are iteratively adjusted in order to meet the terminal conditions. Aside from the fact that these procedures are notoriously sensitive to the initial guess for $\lambda(0)$—convergence may only occur in small neighborhoods of the correct value about which there exists no a priori knowledge—also slight changes in the guess for $\lambda(0)$ may lead to vastly different solutions. This, however, is merely a reflection of the fact that indeed there may exist multiple extremals. Even a convergent algorithm picks only one particular extremal (and does so arbitrarily through its initial guess for $\lambda(0)$) and at a minimum its local optimality should be asserted through the verification of second-order sufficient conditions for optimality. There simply are possibly many solutions for the extremals and the one to which any given numerical method converges may depend on the rather arbitrary choice that has been made for the initialization of the shooting procedure.

6.2. Termination rules for numerical procedures. Direct numerical methods, which instead use discretizations of the state and adjoint equations, generally rely on changes in the value of the objective to terminate the algorithm. The value function $V$, however, may have many regions (both around the optimal solution, but also in regions away from the optimum) where this function is rather flat. While some cases when a thus computed control violates the necessary conditions for optimality are rather obvious (e.g., if it is discontinuous for an $L_2$-type objective), in most situations this is not the case. For example, in the optimal solution for the antiangiogenic monotherapy problem considered in [16] there exist optimal solutions when a singular control saturates. In these cases, it is not optimal to follow the singular arc until saturation occurs, but optimal controlled trajectories exit from the singular arc before this happens. The numerical differences in the objective compared with the control which lets the singular control saturate are tiny and will not be picked up unless computations are done to high precision. Similarly, optimal chattering controls will not be found using ordinary numerical procedures. In each case, it may be argued that for all practical purpose the control which was computed is a close enough suboptimal solution. This indeed is correct, but it then should be stated thus. But generally it is difficult to make such a claim without knowing what the true optimal control is.

Another unrelated difficulty of numerical solutions obviously is that these are for specific parameter values. In many applications, especially medical ones or those to public health concerns, these parameters often are highly uncertain and can vary over broad ranges, often still changing in time. Numerical computations thus only provide a snapshot of one particular condition. Theoretical analysis, on the other hand, aims at a full understanding of the dynamical system and its optimal
controls with all the bifurcations or changes in the structures of optimal solutions that come along with this. Clearly, this is a difficult undertaking only possible for small systems or systems whose dynamics have special properties. Yet, numerical solutions should not be overstated, especially if all that was done was to compute one extremal.

6.3. **Unjustified claims about optimality.** For the reasons outlined above, and also for sometimes blind reliance on commercially available, not very sound software, many papers published on applications of optimal control to specific problems are fraught with heavily overstated and sometimes even incorrect claims about the optimality of solutions. Claims of the form “the optimal solutions are . . .” without any application of at least second-order optimality conditions are not justified. Categorical statements about the optimality of particular solutions without the analysis of all extremals are unjustified. In particular, the fact that the Hamiltonian is quadratic (or strictly convex) in the controls does not guarantee a unique solution as the optimal control problem is only a convex optimization problem if also all other constraints that define the problem are convex. Owing to a generally nonlinear dynamics, this is rarely the case for the set of all controlled trajectories.

7. **Conclusion.** Optimal control problems are infinite-dimensional optimization problems which exhibit all the intricacies and complexities one would expect from such a setting. Generally, there just is no simple or quick way to obtain the true solutions. Superficial schemes which, for example, follow the standard recipe of choosing a model formulation so that the Hamiltonian function $H$ is strictly convex in $u$ and then compute one solution of the associated 2-point boundary value problem for the state and multiplier numerically do not guarantee that a ‘solution’ to the problem has been found and should not be sold as such. Optimal control is far too often used as a superficially understood tool that was ‘learned’ overnight to add another section to a paper focused on dynamical systems. This naturally causes overstatements of the significance of numerical results which are unjustified.

**REFERENCES**

[1] L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, 1974.
[2] S. Bhan and H. Schättler, A variational approach to perturbation feedback control for optimal control problems with terminal constraints and free terminal time, *Set-Valued Var. Anal.*, 27 (2019), 309–330.
[3] F. Black and M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.*, 81 (1973), 637–654.
[4] V. G. Boltyansky, Sufficient conditions for optimality and the justification of the dynamic programming method, *SIAM J. Control*, 4 (1966), 326–361.
[5] V. G. Boltyansky, *Mathematical Methods of Optimal Control*, Holt, Rinehart and Winston, Inc., 1971.
[6] B. Bonnard and M. Chyba, *Singular Trajectories and their Role in Control Theory*, Mathématiques & Applications, vol. 40, Springer Verlag, Paris, 2003.
[7] U. Boscain and B. Piccoli, *Optimal Syntheses for Control Systems on 2-D Manifolds*, Mathématiques & Applications, Vol. 43, Springer-Verlag, Berlin, 2004.
[8] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, American Institute of Mathematical Sciences, 2007.
[9] A. E. Bryson Jr. and Y. C. Ho, *Applied Optimal Control*, Revised Printing, Hemisphere Publishing Company, New York, 1975.
[10] C. Byrnes and H. Frankowska, Unicité des solutions optimales et absence de chocs pour les équations d’Hamilton–Jacobi–Bellman et de Riccati, *C. R. Acad. Sci. Paris*, 315 (1992), 427–431.
[11] C. I. Byrnes and A. Jhemi, Shock waves for Riccati partial differential equations arising in nonlinear optimal control, in: Systems, Models and Feedback: Theory and Applications, (A. Isidori and T. J. Tarn, eds.), Birkhäuser, (1992), 211–227.
[12] I. V. Girsanov, Lectures on Mathematical Theory of Extremum Problems, Lecture Notes in Economics and Mathematical Systems, Vol. 67, Springer-Verlag, Berlin-New York, 1972.
[13] A. D. Ioffe and V. M. Tikhomirov, Theory of Extremal Problems, North-Holland, Amsterdam, 1979.
[14] H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley–Interscience, 1972.
[15] U. Ledzewicz and H. Schättler, Optimal bang-bang controls for a 2-compartment model in cancer chemotherapy, J. Optim. Theory Appl., 114 (2002), 609–637.
[16] U. Ledzewicz and H. Schättler, Antiangiogenic therapy in cancer treatment as an optimal control problem, SIAM J. Control Optim., 46 (2007), 1052–1079.
[17] U. Ledzewicz and H. Schättler, Combination of antiangiogenic treatment with chemotherapy as a multi-input optimal control problem, Math. Methods in the Applied Sciences, publ. online.
[18] D. Liberzon, Calculus of Variations and Optimal Control, Princeton University Press, Princeton, 2012.
[19] R. C. Merton, Lifetime portfolio selection under uncertainty: The continuous-time case, The Review of Economics and Statistics, 51 (1969), 247–257.
[20] H. G. Moyer, Sufficient conditions for a strong minimum in singular control problems, SIAM J. Control, 11 (1973), 620–636.
[21] B. Piccoli and H. J. Sussmann, Regular synthesis and sufficient conditions for optimality, SIAM J. Control Optim., 39 (2000), 359–410.
[22] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, Macmillan, New York, 1964.
[23] H. Schättler and U. Ledzewicz, Geometric Optimal Control, Interdisciplinary Applied Mathematics, Vol. 38, Springer, New York, 2012.
[24] H. Schättler and U. Ledzewicz, Optimal Control for Mathematical Models of Cancer Therapies, Interdisciplinary Applied Mathematics, Vol. 42, Springer, New York, 2015.
[25] H. Schättler, U. Ledzewicz and H. Maurer, Sufficient conditions for strong local optimality in optimal control problems with $L^2$-type objectives and control constraints, Discrete Contin. Dyn. Syst. Ser. B, 19 (2014), 2657–2679.
[26] H. J. Sussmann and J. C. Willems, 300 years of optimal control: From the brachistochrone to the maximum principle, IEEE Control Systems, 17 (1997), 32–44.
[27] G. W. Swan, Applications of Optimal Control Theory in Medicine, Marcel Dekker, New York, 1984.
[28] G. W. Swan, Role of optimal control in cancer chemotherapy, Mathematical Biosciences, 101 (1990), 237–284.
[29] A. Swierniak, Cell cycle as an object of control, Journal of Biological Systems, 3 (1995), 41–54.

Received for publication September 2021; early access February 2022.

E-mail address: uledzew@siue.edu
E-mail address: hms@email.wustl.edu