Use of Concrete Examples and Visualizations to Improve the Discussion of Pontryagin's Maximum Principle in Control Education

Lilija Naiwert
Applied Mathematics
Hochschule RheinMain
Wiesbaden, Germany
lilija.naiwert@student.hs-rm.de

Karlheinz Spindler
Applied Mathematics
Hochschule RheinMain
Wiesbaden, Germany
karlheinz.spindler@hs-rm.de

***Abstract*—This paper, which grew out of an ongoing project geared towards improvements in control education, provides two examples of autonomous two-dimensional control systems which are simple enough to be analyzed analytically, but rich enough to exhibit interesting features. The examples are used to elaborate and visualize concepts which are relevant for the proof of Pontryagin's Maximum Principle (needle variations, Boltyanskiĭ cones, reachable sets). The examples are classroom-tested; they are presented in a comprehensible and illustrative way suitable for classroom use.**

***Index Terms*—Optimal control, control education, Pontryagin's Maximum Principle, reachable sets, Boltyanskiĭ cones.**

I. Introduction

Pontryagin's Maximum Principle, one of the pillars of optimal control theory, is well understood, and polished textbook treatments of this principle are available; see for example [1]-[5] and [7]. Nevertheless, a mathematically rigorous derivation in a first course on optimal control theory is still a challenge, and novices are often bewildered by the variety of concepts which come into play: differential equations with measurable right-hand side, needle variations, variational equations, Boltyanskiĭ cones, separation properties of convex sets, Brouwer's fixed point theorem, analysis on manifolds. It is therefore helpful to identify good examples which are complicated enough to reveal typical aspects, but are simple enough to be handled analytically. This is currently pursued in an ongoing student-professor project (see acknowledgment) geared towards improvements in control education. In this paper, which extends [7], we discuss two such examples (and purposefully present them in an elementary way suitable for classroom use).

II. First example

We first study the system

$$\begin{aligned} \dot{x} &= u \\ \dot{y} &= (1/2)\cdot(x^2+u^2) \end{aligned} \tag{1}$$

with the initial condition $\big(x(0),y(0)\big)=(0,0)$ subject to the control constraint $|u|\le 1$. We seek to determine, for any given fixed time $T>0$, the reachable set R_T consisting of those points to which the system can be steered within the time interval $[0,T]$ by an admissible control. Let us start by identifying some basic properties which can be ascertained without a detailed analysis.

Special role of the null steering. The trivial control $u_\star\equiv 0$ is admissible and yields the system response $(x_\star,y_\star)\equiv(0,0)$. Hence we can remain at the initial state $(0,0)$ as long as we wish, and since every system trajectory $t\mapsto\big(x(t),y(t)\big)$ satisfies $\dot{y}\ge 0$ and hence $y\ge y(0)=0$, the point $(0,0)$ is a boundary point of R_T for all $T>0$. Moreover, we see that $t_1<t_2$ implies $R_{t_1}\subseteq R_{t_2}$; namely, if a state can be reached at some time t_1 due to a control u, then this state can also be reached at any later time $t_2>t_1$ by using the control

$$\widehat{u}(t) \;:=\; \begin{cases} 0, & \text{if } 0\le t<t_2-t_1, \\ u\big(t-(t_2-t_1)\big), & \text{if } t_2-t_1\le t\le t_2. \end{cases} \tag{2}$$

A priori estimate. Since an admissible control satisfies $-1\le u(\tau)\le 1$ for all τ, we find first that $x(t)=\int_0^t u(\tau)\mathrm{d}\tau$ satisfies the estimate $|x(t)|\le\int_0^t 1\,\mathrm{d}\tau=t$ and then that $y(t)=\int_0^t\big(\big(x(\tau)^2+u(\tau)^2\big)/2\big)\,\mathrm{d}\tau$ satisfies the estimate $0\le y(t)\le\int_0^t\big((\tau^2+1)/2\big)\,\mathrm{d}\tau=(t^3/6)+(t/2)$. Thus for all $T>0$ we have the inclusion

$$R_T \;\subseteq\; [-T,T]\times\left[0,\frac{T^3}{6}+\frac{T}{2}\right]. \tag{3}$$

Symmetry. Let $t\mapsto\big(x(t),y(t)\big)$ be the system response to an admissible control $t\mapsto u(t)$. Then $U(t):=-u(t)$ is again an admissible control, and it is readily checked that the system response to the control U is $t\mapsto\big(-x(t),y(t)\big)$. This implies that the reachable set R_T is symmetrical with respect to the y-axis.

The work described in this paper was partially supported by the Klaus Tschira Foundation. This support is gratefully acknowledged.

III. Application of the Maximum Principle

In this section we use Pontryagin's Maximum Principle to formulate necessary conditions for a control to be a boundary control in our specific example.

Hamiltonian equations. The Hamiltonian is

$$H \;=\; pu + (q/2)\cdot(x^2+u^2). \tag{4}$$

Since $\partial H/\partial x = qx$ and $\partial H/\partial y = 0$, the Hamiltonian equations are given by

$$\begin{aligned} \dot{x} &= u, \\ \dot{y} &= (1/2)(x^2+u^2), \\ \dot{p} &= -qx, \\ \dot{q} &= 0. \end{aligned} \tag{5}$$

The last equation implies that q is a constant. Consequently, only the first three equations will be considered in the sequel. We first observe that the symmetry described in the previous paragraph carries over from the system equations to the Hamiltonian equations. More precisely: If (x,y,p,q) is a solution of the Hamiltonian equations belonging to a control u, if we define a new control U by $U(t) := -u(t)$ and if we let $X(t) := -x(t)$, $Y(t) := y(t)$ and $P(t) := -p(t)$, then (X,Y,P,q) is a solution of the Hamiltonian equations belonging to the control U.

Now if u is a boundary control, then there is a solution of the Hamiltonian equations for which $t \mapsto \big(p(t),q\big)$ never vanishes and for which H becomes maximal as a function of u. Hence we ask for which value of u, given fixed values for x, p and q, the expression $H = pu+(q/2)(x^2+u^2)$ becomes maximal. If $q=0$, this expression reduces to $H = pu$. In this case we find that $\dot{p}=0$, so that p is a constant function (necessarily nonzero because (p,q) is nonzero). If $p>0$, then H becomes maximal for $u=1$; if $p<0$, then H becomes maximal for $u=-1$. This yields the constant functions $u_\star \equiv 1$ and $u_\star \equiv -1$ as possible boundary controls. If $q \neq 0$, then a distinction between different cases is necessary.

Evaluation of the Maximum Condition. For fixed values of x, p and $q \neq 0$, the expression $pu+(q/2)(x^2+u^2)$ becomes maximal if and only if the function $\Phi(u) := 2pu + qu^2$ becomes maximal. Now this function can take its maximum in the interval $-1 \le u \le 1$ only at one of the boundary points $u = \pm 1$ or at a root of $\Phi'(u) = 2p+2qu$, i.e., at $-p/q$, if this point lies in the interval $[-1,1]$. Thus the maximum is necessarily one of the three numbers

$$\begin{aligned} &\Phi(1) = 2p+q, \\ &\Phi(-1) = -2p+q, \\ &\Phi(-p/q) = -p^2/q. \end{aligned} \tag{6}$$

Assuming $p \neq 0$ for the moment and going through the various possible cases, we find that the maximum condition yields

$$u_\star(t) \;=\; \begin{cases} 1, & \text{if } p(t)>0 \text{ and } q>-p(t), \\ -1, & \text{if } p(t)<0 \text{ and } q>p(t), \\ -p(t)/q, & \text{if } q<-|p(t)|. \end{cases} \tag{7}$$

The way the values of a boundary control u depend on the values of the adjoint variables is depicted in Fig. 1.

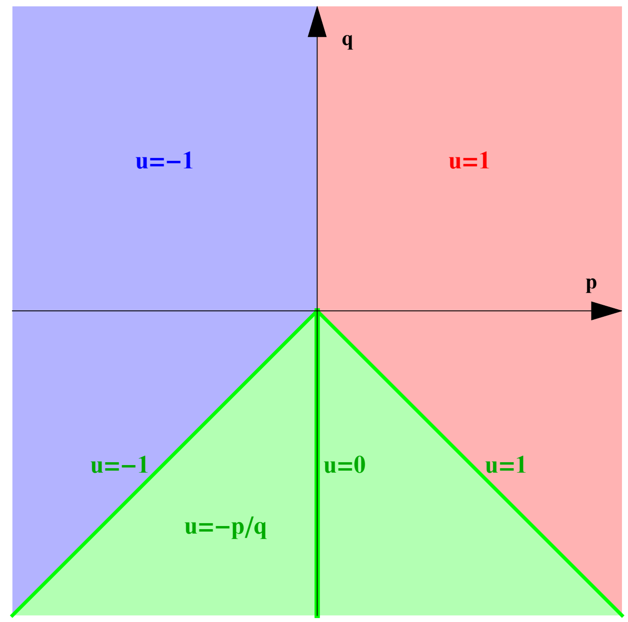

Fig. 1. Values of a boundary control as a function of the values of the adjoint variables.

The case $p=0$ has not been considered so far. This case is only of interest if the condition $p \equiv 0$ is satisfied on a whole time interval $I \subseteq [0,T]$. If this is the case then we have $q \neq 0$ due to the nontriviality condition and consequently $x \equiv 0$ on the interval I because $0 = \dot{p} = -qx$, which implies that $u = \dot{x} \equiv 0$ on I. Thus the case $p=0$ only yields the trivial control $u \equiv 0$ discussed before. We now turn to those solutions of the Hamiltonian equations for which $p(0) \neq 0$.

IV. Boundary Controls

The various possible boundary controls can be found by a case-by-case inspection.

Boundary controls, first case. We consider the case that $q<0$ and $p(0) \ge -q > 0$. In this case the control starts with the value $u=1$. Hence for some time we have $\dot{x} \equiv 1$, consequently $x(t)=t$ and therefore $\dot{p}(t) = -qx(t) = -qt > 0$, which yields $p(t) \ge p(0) > 0$ for all t. This implies that there is no switch in the control, and we have $u \equiv 1$ in this case.

Boundary controls, second case. Analogously, if $q<0$ and $p(0) \le q < 0$, then we start with the control $u=-1$, and the mapping $t \mapsto p(t)$ decreases, with the consequence that there cannot be a switch in the control. Thus we have $u \equiv -1$ in this case.

Boundary controls, third case. We consider the case that $q>0$ and $p(0)>0$. In this case the control starts with the value $u=1$.

First phase: Starting with the initial values $x(0) = 0$, $y(0) = 0$ and $p(0) =: p_0 > 0$, we integrate the Hamiltonian equations with the control $u \equiv 1$ and obtain

$$\begin{aligned} u(t) &= 1, \\ x(t) &= t, \\ y(t) &= (t^3/6) + (t/2), \\ p(t) &= -(q/2)t^2 + p_0. \end{aligned} \tag{8}$$

The control $u \equiv 1$ remains unchanged as long as $p(t) > 0$, i.e., as long as $t < \sqrt{2p_0/q} =: \tau$. Thus if $T \leq \tau$, we obtain the constant control $u \equiv 1$. If not, the function p changes sign at time τ from $+$ to $-$, because $p(\tau) = 0$ and $\dot{p}(\tau) = -q\tau < 0$, and the control u switches from the value $+1$ to the value -1. Then a second phase begins, starting with the values

$$\begin{aligned} x(\tau) &= \tau, \\ y(\tau) &= (\tau^3/6) + (\tau/2), \\ p(\tau) &= 0. \end{aligned} \tag{9}$$

Second phase: Starting with the initial values (9), we integrate the Hamiltonian equations with the control $u \equiv -1$ and obtain

$$\begin{aligned} u(t) &= -1, \\ x(t) &= -t + 2\tau, \\ y(t) &= (t-2\tau)^3/6 + (t/2) + (\tau^3/3), \\ p(t) &= (q/2)\big((t-2\tau)^2 - \tau^2\big) \\ &= (q/2)(t-\tau)(t-3\tau). \end{aligned} \tag{10}$$

The second phase ends (unless the given time T is reached before) as soon as the function p has its next zero, namely at $t = 3\tau$. At this time p changes sign from $-$ to $+$, because $p(3\tau) = 0$ and $\dot{p}(3\tau) = q\tau > 0$, and the control u switches from the value -1 to the value $+1$. Then a third phase begins, starting with the values taken at the end of the second phase, namely

$$\begin{aligned} x(3\tau) &= -\tau, \\ y(3\tau) &= (3\tau^3/6) + (3\tau/2), \\ p(3\tau) &= 0. \end{aligned} \tag{11}$$

Third phase: Starting with the initial values (11), we integrate the Hamiltonian equations with the control $u \equiv 1$ and obtain

$$\begin{aligned} u(t) &= 1, \\ x(t) &= t - 4\tau, \\ y(t) &= (t-4\tau)^3/6 + (t/2) + (2\tau^3/3), \\ p(t) &= -(q/2)\big((t-4\tau)^2 - \tau^2\big) \\ &= -(q/2)(t-3\tau)(t-5\tau). \end{aligned} \tag{12}$$

The third phase ends (unless the given time T is reached before) as soon as the function p hat its next zero, namely at $t = 5\tau$. At this time p changes sign from $+$ to $-$, because $p(5\tau) = 0$ and $\dot{p}(5\tau) = -q\tau < 0$, and the control u switches from the value 1 to the value -1. Then a fourth phase begins, starting with the values taken at the end of the third phase, namely

$$\begin{aligned} x(5\tau) &= \tau, \\ y(5\tau) &= (5\tau^3/6) + (5\tau/2), \\ p(5\tau) &= 0. \end{aligned} \tag{13}$$

Fourth phase: Starting with the initial values (13), we integrate the Hamiltonian equations with the control $u \equiv -1$ and obtain

$$\begin{aligned} u(t) &= -1, \\ x(t) &= -t + 6\tau, \\ y(t) &= (t-6\tau)^3/6 + (t/2) + (3\tau^3/3), \\ p(t) &= (q/2)\big((t-6\tau)^2 - \tau^2\big) \\ &= (q/2)(t-5\tau)(t-7\tau). \end{aligned} \tag{14}$$

The fourth phase ends (unless the given time T is reached before) as soon as the function p takes its next zero, namely at $t = 7\tau$. A this time p changes sign from $-$ to $+$, because $p(7\tau) = 0$ and $\dot{p}(7\tau) = q\tau > 0$, and the control u switches from the value -1 to the value $+1$. Then a fifth phase begins, starting with the values taken at the end of the fourth phase, namely

$$\begin{aligned} x(7\tau) &= -\tau, \\ y(7\tau) &= (7\tau^3/6) + (7\tau/2), \\ p(7\tau) &= 0. \end{aligned} \tag{15}$$

The general pattern now becomes obvious: The duration between two subsequent switching times is always the same, namely 2τ, und the control and the system response during the k-th phase are given by

$$\begin{aligned} u(t) &= (-1)^{k+1}, \\ x(t) &= (-1)^{k+1} \cdot \big(t - 2(k-1)\tau\big), \\ y(t) &= \big(t - 2(k-1)\big)^3/6 + (t/2) + (k-1)\tau^3/3. \end{aligned} \tag{16}$$

In principle, the first switching time τ can be any value between 0 and T, and this value sets the switching pattern, as follows:

- if $\tau < T \leq 3\tau$ (i.e., $T/3 \leq \tau < T$), there is exactly one switching;
- if $3\tau < T \leq 5\tau$ (i.e., $T/5 \leq \tau < T/3$), there are exactly two switchings;
- if $5\tau < T \leq 7\tau$ (i.e., $T/7 \leq \tau < T/5$), there are exactly three switchings,

and so on. Fig. 2 shows the system responses for different values of τ (where the value $T := 2$ was chosen).

Using the facts that, on the one hand, the set R_T is symmetric with respect to the y-axis and that, on the other hand, we have $R_t \subseteq R_T$ for all $t < T$, we see that as lower boundary points of R_T in the first quadrant only those points (x, y) are possible for which $0 \leq x \leq T$ and $y = (x^3/6) + (x/2)$ (which represent the system response to the control $u \equiv 1$), and that as upper boundary points in the first quadrant only the end points of such trajectories are possible which result from a control with exactly one switching and a switching

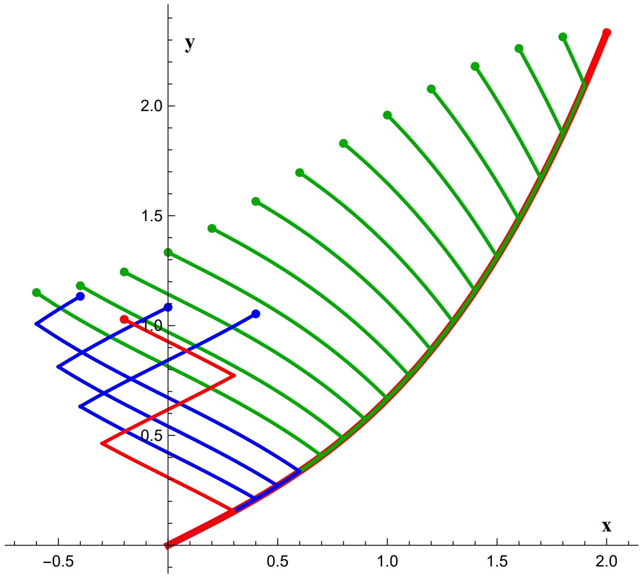

Fig. 2. Trajectories resulting from bang-bang controls with different first switching times (in the case $T = 2$).

time $\tau \geq T/2$. (Such points are marked green in Fig. 2.) The set of these points is

$$\left\{ \left(-T + 2\tau, \frac{(T - 2\tau)^3}{6} + \frac{T}{2} + \frac{\tau^3}{3} \right) \;\middle|\; \frac{T}{2} \leq \tau \leq T \right\} = \left\{ \left(x, \frac{-x^3}{6} + \frac{T}{2} + \frac{(x+T)^3}{24} \right) \;\middle|\; 0 \leq x \leq T \right\}. \tag{17}$$

Boundary controls, fourth case. We now consider the situation that $q > 0$ and $p(0) < 0$, in which case the control starts with the value $u = -1$. The resulting control is then exactly the negative of the control studied previously, and the resulting system trajectory is the mirror image with respect to the y-axis of the trajectory obtained previously, due to the symmetry properties discussed before.

Subsumption of previous cases. Following the previous considerations, the reachable set R_T contains the point set, shown in Fig. 3, whose lower and upper boundaries are the graphs of the functions

$$\begin{aligned} y_{\text{bottom}}(x) &:= \frac{|x|^3}{6} + \frac{|x|}{2} \quad \text{and} \\ y_{\text{top}}(x) &:= \frac{-|x|^3}{6} + \frac{T}{2} + \frac{(|x|+T)^3}{24}, \end{aligned} \tag{18}$$

each taken over the range $-T \leq x \leq T$. Only the upper boundary and the point $(0,0)$, marked black in Fig. 3, are possible boundary points of R_T; all other points in the gray area are necessarily inner points of R_T. This also applies to the points of the lower boundary of this set (marked by a dashed line), because a control u such that $u \equiv 0$ on some interval $[0, \theta]$ and then $u \equiv \pm 1$ on $[\theta, T]$ cannot be a boundary control, as we have seen.

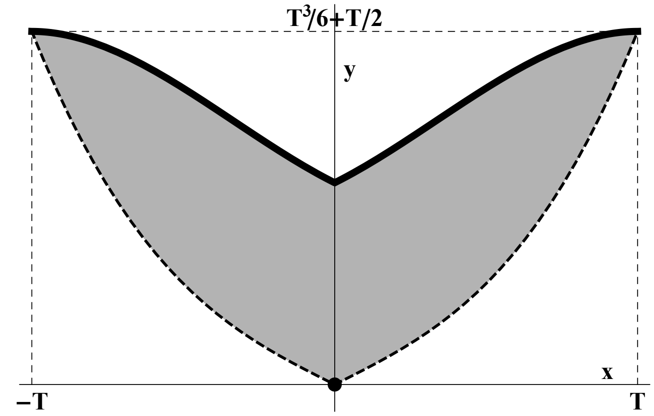

Fig. 3. Part of the reachable set R_T identified so far (here shown for $T = 2$).

Boundary controls, fifth case. We now consider the case that $q < 0$ and $|p(0)| \leq |q|$ so that $u(t) = -p(t)/q$. Plugging this control into the Hamiltonian equations results in

$$\begin{aligned} \dot{x} &= -p/q, \\ \dot{y} &= (1/2) \cdot (x^2 + p^2/q^2), \\ \dot{p} &= -qx. \end{aligned} \tag{19}$$

We find that $\ddot{p} = -q\dot{x} = p$; hence there are constants A and B such that $p(t) = Ae^t + Be^{-t}$. This yields $x(t) = -\dot{p}(t)/q = -(1/q) \cdot (Ae^t - Be^{-t})$; the initial condition $x(0) = 0$ then implies that $B = A$ and $p(t) = A(e^t + e^{-t})$. Plugging in $t = 0$ shows that $p_0 = p(0) = 2A$ so that

$$p(t) \;=\; p_0 \cosh(t). \tag{20}$$

If $p_0 = 0$ this yields the special case $p \equiv 0$ discussed before, which leads to the trivial control $u \equiv 0$. Assume that $p_0 \neq 0$. From $p(t) = p_0 \cosh(t)$ we find that

$$\begin{aligned} u(t) &= -\frac{p(t)}{q} = -\frac{p_0}{q} \cosh(t) \quad \text{and} \\ x(t) &= -\frac{\dot{p}(t)}{q} = -\frac{p_0}{q} \sinh(t). \end{aligned} \tag{21}$$

Plugging the last two equations into the equation $\dot{y} = (1/2) \cdot (x^2 + u^2)$, we find that

$$\begin{aligned} \dot{y}(t) &= \frac{p_0^2}{2q^2} \left(\sinh(t)^2 + \cosh(t)^2 \right) \\ &= \frac{p_0^2}{4q^2} (e^{2t} + e^{-2t}) = \frac{p_0^2}{2q^2} \cosh(2t) \end{aligned} \tag{22}$$

which, after integration, yields

$$y(t) \;=\; \frac{p_0^2}{4q^2} \sinh(2t) \;=\; \frac{p_0^2}{8q^2} (e^{2t} - e^{-2t}). \tag{23}$$

Introducing the constant $\sigma := -p_0/q$ (which necessarily satisfies the condition $|\sigma| \leq 1$), this results in

$$\begin{aligned} u(t) &= \sigma \cosh(t), \\ x(t) &= \sigma \sinh(t), \\ y(t) &= (\sigma^2/4) \sinh(2t), \\ p(t) &= -q\sigma \cosh(t). \end{aligned} \tag{24}$$

If $|\sigma| \leq 1/\cosh(T)$, then $|u(t)| \leq 1$ for all $t \in [0, T]$, and the control u is admissible. This yields the parabolic arc

$$\begin{aligned} &\left\{ \big(\sigma \sinh(T),\, (\sigma^2/4)\sinh(2T)\big) \;\middle|\; |\sigma| \leq \frac{1}{\cosh(T)} \right\} \\ &= \left\{ \left(x, \frac{x^2}{2\tanh(T)} \right) \;\middle|\; |x| \leq \tanh(T) \right\} \end{aligned} \tag{25}$$

as a potential part of the boundary ∂R_T of the reachable set R_T. If $1/\cosh(T) < |\sigma| \leq 1$ there is saturation at time $\tau := \operatorname{arcosh}(1/|\sigma|)$; i.e., at this time the control u reaches its extremal value -1 or 1, and from then on remains constant with this value. We only discuss the case $\sigma > 0$. (The case $\sigma < 0$ can be treated completely analogously.) We then have $\cosh(\tau) = 1/\sigma$, hence $\sigma = 1/\cosh(\tau)$, therefore $\sinh(\tau) = \sqrt{1-\sigma^2}/\sigma$ and consequently $u(\tau) = 1$ and

$$\begin{aligned} x(\tau) &= \sqrt{1-\sigma^2} = \tanh(\tau), \\ y(\tau) &= \sqrt{1-\sigma^2}/2 = \tanh(\tau)/2, \\ p(\tau) &= -q. \end{aligned} \tag{26}$$

Beginning with these initial conditions, we have to integrate the Hamiltonian equations

$$\begin{aligned} \dot{x} &= 1, \\ \dot{y} &= (x^2+1)/2, \\ \dot{p} &= -qx \end{aligned} \tag{27}$$

for $t \geq \tau$. We first find that $x(t) = t - \tau + \tanh(\tau)$ and then on the one hand $\dot{y}(t) = \big((t-\tau)^2 + 2\tanh(\tau)(t-\tau) + \tanh(\tau)^2 + 1\big)/2$ and hence

$$\begin{aligned} y(t) &= \frac{(t-\tau)^3}{6} + \frac{\tanh(\tau)}{2} \cdot (t-\tau)^2 \\ &\quad + \frac{\tanh(\tau)^2+1}{2} \cdot (t-\tau) + \frac{\tanh(\tau)}{2}, \end{aligned} \tag{28}$$

on the other hand $\dot{p}(t) = -q\big(t - \tau + \tanh(\tau)\big)$ and hence

$$p(t) = -\frac{q(t-\tau)^2}{2} - q\tanh(\tau)(t-\tau) - q. \tag{29}$$

This yields as potential boundary points of R_T the points $\big(x(T), y(T)\big)$ such that

$$\begin{aligned} x(T) &= T - \tau + \tanh(\tau), \\ y(T) &= \frac{(T-\tau)^3}{6} + \frac{\tanh(\tau)}{2} \cdot (T-\tau)^2 \\ &\quad + \frac{\tanh(\tau)^2+1}{2} \cdot (T-\tau) + \frac{\tanh(\tau)}{2} \end{aligned} \tag{30}$$

with $0 \leq \tau \leq T$, which are reached from $(0,0)$ by applying the control

$$u(t) = \begin{cases} \cosh(t)/\cosh(\tau), & 0 \leq t \leq \tau, \\ 1, & \tau \leq t \leq T. \end{cases} \tag{31}$$

(Since $\sigma\cosh(t) = 1$ the condition $1/\cosh(T) \leq \sigma \leq 1$ takes the form $\cosh(T) \geq \cosh(\tau) \geq \cosh(0)$, which means $0 \leq \tau \leq T$.) Additional points of the reachable set R_T (but not of the boundary of this set!) are obtained again by choosing, on some time interval $[0, \theta]$ with $0 < \theta < T$, the null control $u \equiv 0$ (which is tantamount to remaining in the initial state) and by then switching to a control of type (31) (with a time duration $T - \theta$ instead of T). Fig. 4 shows which states can be reached in this fashion. The clipping in Fig. 5 shows the situation more clearly. The states at those times at which saturation occurs lie on the curve $y = |x|/2$ (which is the dashed blue curve in Fig. 5). The total reachable set R_T, shown in Fig. 6, is the union of the sets depicted in Fig. 3 and Fig. 4.

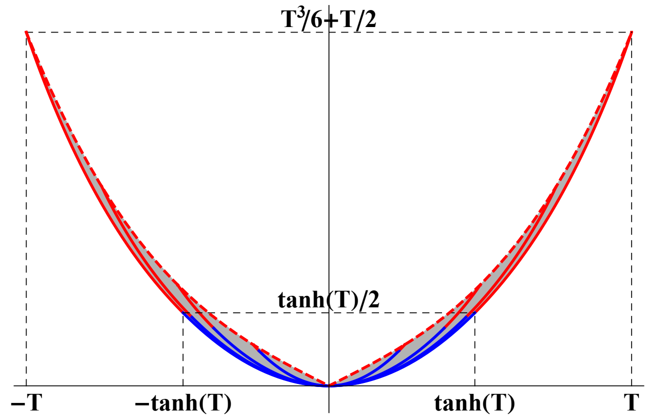

Fig. 4. Additional point set identified as a subset of R_T (here shown for $T = 2$).

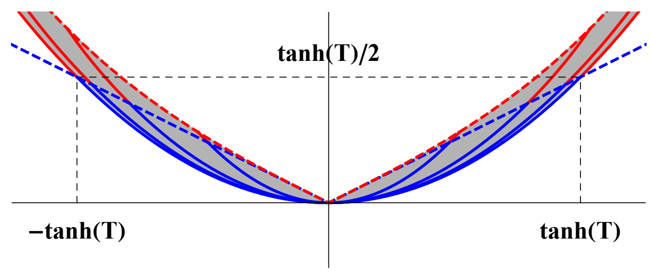

Fig. 5. Clipping from the previous diagram.

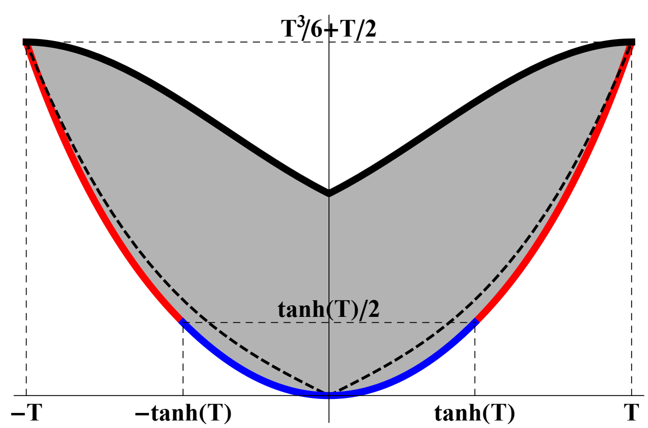

Fig. 6. Reachable set R_T (here shown for $T = 2$).

To understand the time evolution of the control system under consideration, it is helpful to see how the reachable set R_T changes as a function of T. This behavior is shown in Fig. 7, along with two typical solutions $t \mapsto \big(x(t), y(t), p(t), q\big)$ of the Hamiltonian equations for different boundary controls, where (x, y, p, q) is represented by the vector $(p, q)^T$ attached to the point (x, y).

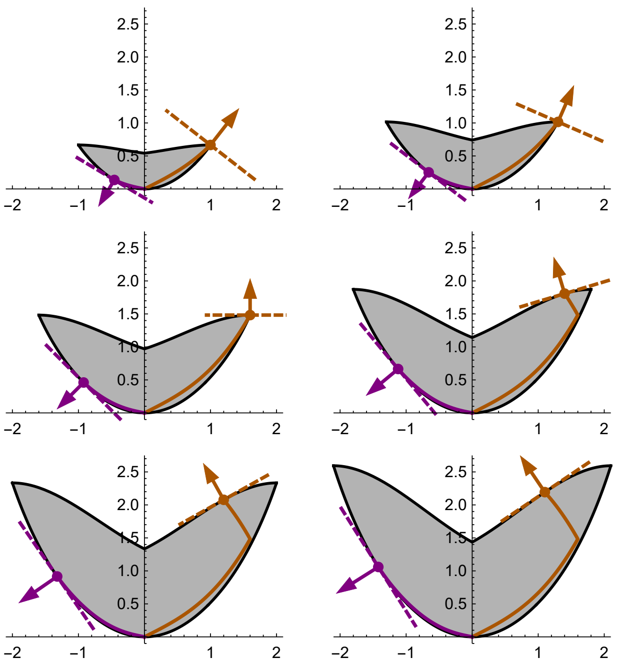

Fig. 7. Reachable set R_T for $T = 1.0$ (upper left), $T = 1.3$ (upper right), $T = 1.6$ (middle left), $T = 1.8$ (middle right), $T = 2.0$ (lower left) and $T = 2.1$ (lower right).

V. Boltyanskiĭ Cones

A key ingredient in the proof of Pontryagin's Maximum Principle is the investigation of the effect of needle variations of a boundary control on the resulting state trajectory. Applying such needle variations results in the formation of the Boltyanskiĭ cone associated with this boundary control, which serves as a local approximation of the reachable set at the point to which the boundary control steers the system. We will now determine the Boltyanskiĭ cones of several different reference controls (both boundary controls and other controls). Doing so is somewhat tedious (especially if done by hand), but turned out to be beneficial in classroom use.

First reference control. We first consider the null control $u_\star \equiv 0$ as a reference control and a simple needle variation

$$u_\varepsilon(t) := \begin{cases} 0, & \text{if } 0 \le t < \tau - \varepsilon\ell \\ w, & \text{if } \tau - \varepsilon\ell \le t < \tau \\ 0, & \text{if } \tau \le t \le T \end{cases} \tag{32}$$

with $w \in [-1, 1]$, $\tau \in [0, T]$ and $\ell > 0$. Without loss of generality, the scaling factor ℓ is chosen to be $\ell := 1$. We denote by $(x_\varepsilon, y_\varepsilon)$ the system response to the control u_ε; since no misunderstandings are possible we simply write (x, y) instead of $(x_\varepsilon, y_\varepsilon)$ for the moment. In the first phase $(0 \le t \le \tau - \varepsilon)$ we simply have $x \equiv 0$ and $y \equiv 0$. In the second phase $(\tau - \varepsilon \le t \le \tau)$ we obtain (x, y) as the solution of the initial value problem

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} w \\ (1/2)(x^2 + w^2) \end{bmatrix}, \quad \begin{bmatrix} x(\tau-\varepsilon) \\ y(\tau-\varepsilon) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{33}$$

which is given by

$$\big(x(t), y(t)\big) = \left(w(t-\tau+\varepsilon), \ \frac{w^2}{2}\left[\frac{(t-\tau+\varepsilon)^3}{3} + (t-\tau+\varepsilon) \right] \right). \tag{34}$$

In the third phase $(\tau \le t \le T)$ we then obtain (x, y) as the solution of the initial value problem

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 \\ x^2/2 \end{bmatrix}, \quad \begin{bmatrix} x(\tau) \\ y(\tau) \end{bmatrix} = \begin{bmatrix} w\varepsilon \\ (w^2/2)(\varepsilon^3/3 + \varepsilon) \end{bmatrix} \tag{35}$$

which is given by

$$\big(x(t), y(t)\big) = \left(w\varepsilon, \ \frac{w^2}{2} \cdot \left[\varepsilon^2 (t-\tau) + \frac{\varepsilon^3}{3} + \varepsilon \right] \right). \tag{36}$$

For clarity's sake we now write again $(x_\varepsilon, y_\varepsilon)$ instead of (x, y). The endpoint of this trajectory is the point

$$\big(x_\varepsilon(T), y_\varepsilon(T)\big) = \left(w\varepsilon, \ \frac{w^2}{2} \cdot \left[\varepsilon^2 (T-\tau) + \frac{\varepsilon^3}{3} + \varepsilon \right] \right). \tag{37}$$

By construction, the curve $\varepsilon \mapsto \big(x_\varepsilon(T), y_\varepsilon(T)\big)$ is a curve in R_T which approaches $\big(x_\star(T), y_\star(T)\big) = (0, 0)$ as $\varepsilon \to 0$. The direction at which this point is approached is

$$\frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \begin{bmatrix} x_\varepsilon(T) \\ y_\varepsilon(T) \end{bmatrix} = \begin{bmatrix} w \\ w^2/2 \end{bmatrix}. \tag{38}$$

The directions in which the vectors of this form point are given by the angles $\varphi \ge 0$ with $\tan\varphi = w/2$ and $w \in [-1, 1]$, i.e., by $\tan\varphi \in [-1/2, 1/2]$. These are exactly the vectors pointing into that part of the reachable set which lies below the set depicted in Fig. 3. The Boltyanskiĭ cone is then the convex cone generated by these vectors, which is the upper half-plane; see Fig. 8.

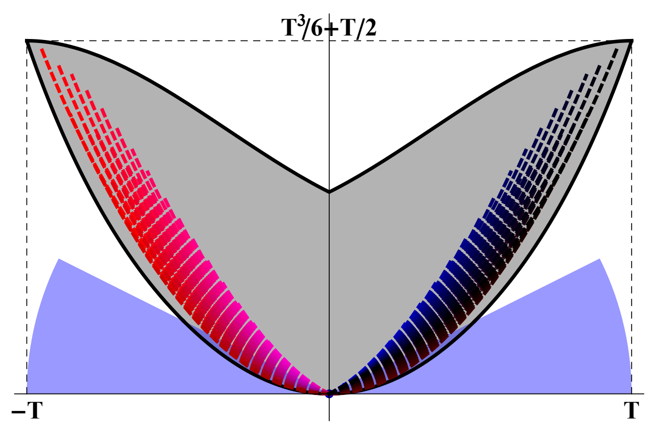

Fig. 8. Formation of the Boltyanskiĭ cone of the control $u_\star \equiv 0$.

Second reference control. We consider the control $u_\star \equiv 1$ as a reference control and a simple needle variation

$$u_\varepsilon(t) \;:=\; \begin{cases} 1, & \text{if } 0 \le t < \tau - \varepsilon\ell \\ w, & \text{if } \tau - \varepsilon\ell \le t < \tau \\ 1, & \text{if } \tau \le t \le T \end{cases} \tag{39}$$

with $w \in [-1,1]$, $\tau \in [0,T]$ and $\ell > 0$. We denote by $(x_\varepsilon, y_\varepsilon)$ the system response to the control u_ε; since no misunderstandings are possible we simply write (x,y) instead of $(x_\varepsilon, y_\varepsilon)$ for the moment. In the first phase ($0 \le t < \tau - \varepsilon\ell$) we have $x(t) = t$ and $y(t) = (t^3/6) + (t/2)$; this phase ends at the time $\tau - \varepsilon\ell$ with the state

$$\begin{aligned} x(\tau - \varepsilon\ell) &= \tau - \varepsilon\ell, \\ y(\tau - \varepsilon\ell) &= (\tau - \varepsilon\ell)^3/6 + (\tau - \varepsilon\ell)/2. \end{aligned} \tag{40}$$

In the second phase ($\tau - \varepsilon\ell \le t < \tau$) we obtain (x,y) as the solution of the system of differential equations $\dot{x} = w$ and $\dot{y} = (x^2 + w^2)/2$ satisfying the initial conditions (40). The solution is given by

$$\begin{aligned} x(t) &= w(t-\tau+\varepsilon\ell) + \tau - \varepsilon\ell, \\ y(t) &= \frac{w^2(t-\tau+\varepsilon\ell)^3}{6} + \frac{w(t-\tau+\varepsilon\ell)^2(\tau-\varepsilon\ell)}{2} \\ &\quad + \frac{(t-\tau+\varepsilon\ell)(\tau-\varepsilon\ell)^2}{2} + \frac{(\tau-\varepsilon\ell)^3}{6} \\ &\quad + \frac{w^2(t-\tau+\varepsilon\ell)}{2} + \frac{\tau-\varepsilon\ell}{2}. \end{aligned} \tag{41}$$

The second phase ends at the time τ with the state

$$\begin{aligned} x(\tau) &= w\varepsilon\ell + \tau - \varepsilon\ell, \\ y(\tau) &= \left(\frac{\tau^3}{6} + \frac{\tau}{2}\right) + \frac{w^2-1}{2} \cdot \varepsilon\ell \\ &\quad + \frac{\tau(w-1)}{2} \cdot \varepsilon^2\ell^2 + \frac{(w-1)(w-2)}{6} \cdot \varepsilon^3\ell^3. \end{aligned} \tag{42}$$

In the third phase ($\tau \le t \le T$) we then obtain (x,y) as the solution of the system of differential equations $\dot{x} = 1$ and $\dot{y} = (x^2+1)/2$ satisfying the initial conditions (42). This solution is given by

$$\begin{aligned} x(t) &= t + (w-1)\varepsilon\ell, \\ y(t) &= \left(\frac{t^3}{6} + \frac{t}{2}\right) + \frac{(w-1)\big(t^2 - \tau^2 + w + 1\big)}{2} \cdot \varepsilon\ell \\ &\quad + \frac{(w-1)(tw - \tau w - t + 2\tau)}{2} \cdot \varepsilon^2\ell^2 \\ &\quad + \frac{(w-1)(w-2)}{6} \cdot \varepsilon^3\ell^3. \end{aligned} \tag{43}$$

For clarity's sake we now write again $(x_\varepsilon, y_\varepsilon)$ instead of (x,y). By construction, the curve $\varepsilon \mapsto \big(x_\varepsilon(T), y_\varepsilon(T)\big)$ is a curve in R_T which approaches $\big(x_\star(T), y_\star(T)\big) = (T, T^3/6 + T/2)$ as $\varepsilon \to 0$. The direction at which this point is approached is easily computed from (43) to be

$$\frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} \begin{bmatrix} x_\varepsilon(T) \\ y_\varepsilon(T) \end{bmatrix} \;=\; \frac{\ell(w-1)}{2} \begin{bmatrix} 2 \\ T^2 - \tau^2 + w + 1 \end{bmatrix}. \tag{44}$$

The vectors of this form (where $|w| \le 1$, $0 \le \tau \le T$ and $\ell \ge 0$) are exactly the nonpositive multiples of vectors of the form $(2,a)^T$ where $0 \le a \le T^2 + 2$. Since these vectors already form a convex cone (without the need to apply multiple needle variations), they constitute the Boltyanskiĭ cone of the reference control $u_\star \equiv 1$. The vectors in this cone are exactly the vectors pointing from $\big(x_\star(T), y_\star(T)\big)$ into the reachable set. This is depicted in Fig. 9.

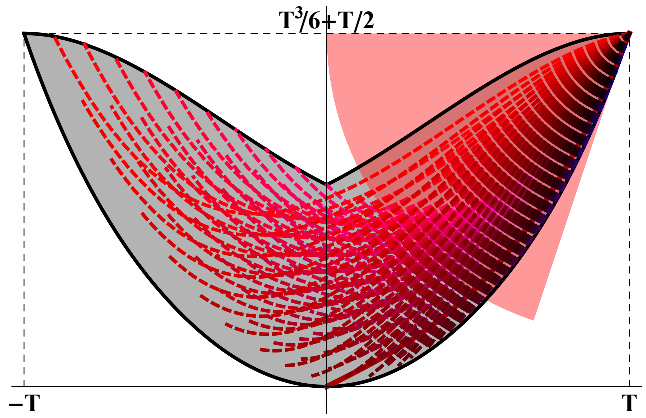

Fig. 9. Formation of the Boltyanskiĭ cone of the control $u_\star \equiv 1$.

Third reference control. We consider the reference control $u_\star$ given by $u_\star \equiv 1$ on $[0,\,T/2)$ and $u_\star \equiv -1$ on $[T/2,\,T]$. Rather than calculating the effect of needle variations by hand in this example, we determine the elements of the Boltyanskiĭ cone of $u_\star$ by integrating the variational equations of the associated reference trajectory, which is depicted as the green curve in Fig. 10 showing the formation of the Boltyanskiĭ cone. Note that this cone approximates only part of the reachable set in a vicinity of the endpoint of the reference trajectory.

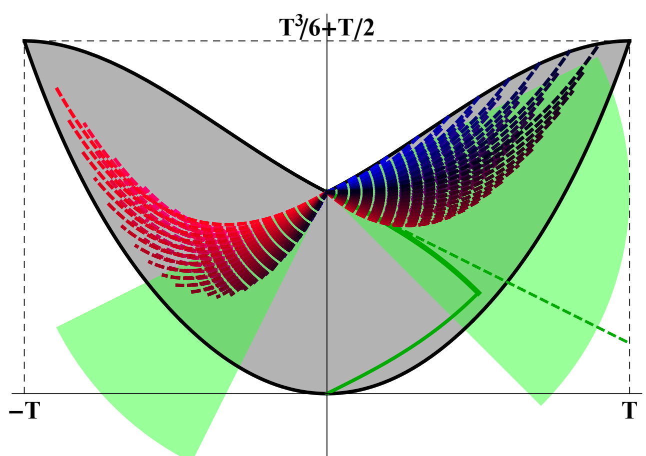

Fig. 10. Formation of the Boltyanskiĭ cone of the control $u_\star$ given by $u_\star \equiv 1$ on $[0,\,T/2)$ and $u_\star \equiv -1$ on $[T/2,\,T]$.

Fourth reference control. We consider the reference control $u_\star$ given by $u_\star \equiv 1$ on $[0, T/4) \cup [3T/4,\,T]$ and $u_\star \equiv -1$

on $[T/4, 3T/4]$. As before, we integrated the variational equations of the reference trajectory rather than determining the Boltyanskiĭ cone by calculating by hand the effect of needle variations. Fig. 11, in which the reference trajectory is the purple curve, shows the formation of the Boltyanskiĭ cone. Note that this cone is not the full space even though $u_\star$ is not a boundary control (while satisfying the conditions of Pontryagin's Maximum Principle).

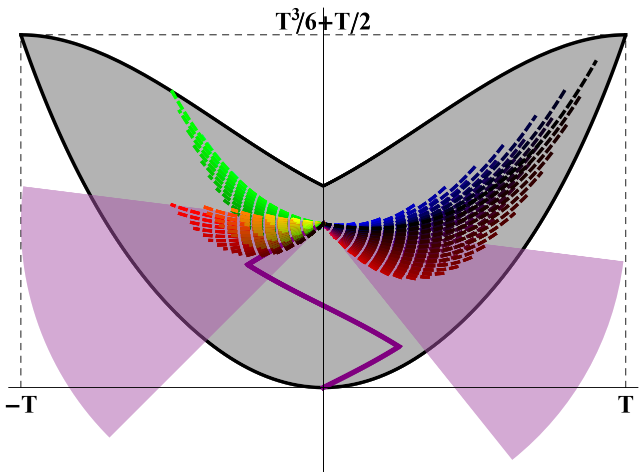

Fig. 11. Formation of the Boltyanskiĭ cone of the control $u_\star$ given by $u_\star \equiv 1$ on $[0, T/4) \cup [3T/4, T]$ and $u_\star \equiv -1$ on $[T/4, 3T/4]$.

Fig. 12 shows simultaneously the Boltyanskiĭ cones of the four reference controls considered so far.

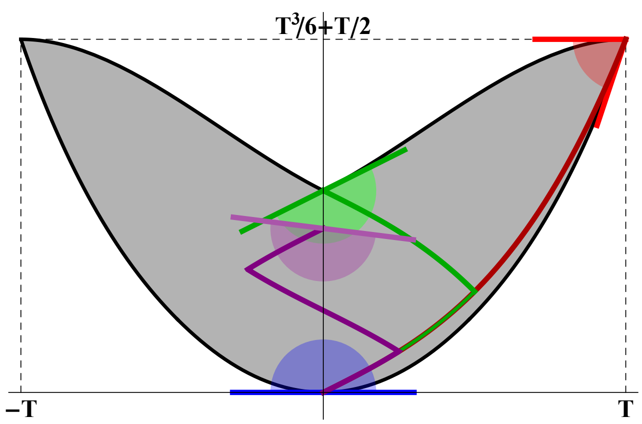

Fig. 12. Boltyanskiĭ cones of the reference controls considered so far.

Fifth reference control. An admissible control is given by $u(t) := \cos(t)$; one readily checks that the system response to this control is given by $\big(x(t), y(t)\big) = (\sin t, t/2)$. We treat u as a reference control and the resulting system response as a reference trajectory. The upper left part of Fig. 13 shows this reference trajectory for $T = 3$. The other parts of Fig. 13 show the system responses to different needle variations of u, namely for $\tau = 0.7$ and $w = -1$ (red), for $\tau = 2.5$ and $w = 1$ (blue) and for $\tau = 2$ and $w = -1$ (green), each with $\ell = 1$.

If all trajectories are included in a single diagram (see Fig. 14), one realizes that the directions determined by the

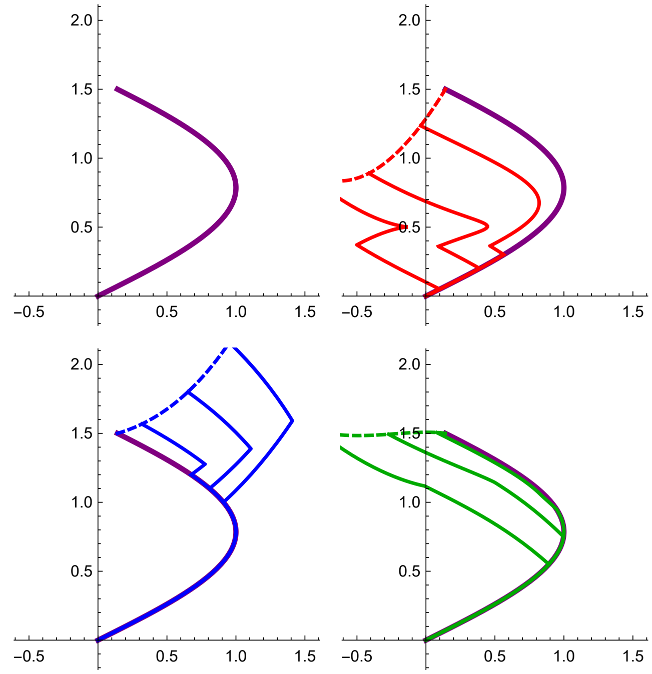

Fig. 13. Reference trajectory and system responses to different needle variations of the reference control.

three chosen needle variations of the reference control u are not contained in a half-space. Hence the Boltyanskiĭ cone of u is all of $\mathbb{R}^2$, which shows that the end point $\big(x(T), y(T)\big)$ of the reference trajectory must be an inner point of the reachable set R_T. (This was, of course, clear *a priori* because we identified all possible boundary controls beforehand and hence know that the chosen reference control u is not such a boundary control.)

VI. Second Example: Insecticide Problem

Assume that an insect population with a natural growth rate $a > 0$ is fought with an insecticide. If $u(t)$ is the killing rate due to the insecticide at time t, the population evolution $t \mapsto x(t)$ satisfies the system equation $\dot{x}(t) = ax(t) - u(t)$. Starting with a given initial population $x(0) = x_0$, we want all insects to be eliminated at a given time $T > 0$. Assume the environmental damage caused by the use of the insecticide over the time interval $[0, T]$ is given by $(1/2)\int_0^T u(t)^2 \mathrm{d}t$. How do we have to choose the control u in order to meet the requirement $x(T) = 0$ while minimizing the environmental damage? Introducing the function $c(t) := (1/2)\int_0^t u(\tau)^2 \mathrm{d}\tau$, which expresses the environmental damage done up to time t, and treating c as an addition al state variable, we consider the augmented system

$$\begin{aligned} \dot{x} &= ax - u \\ \dot{c} &= (1/2) \cdot u^2 \end{aligned} \tag{45}$$

subject to the initial conditions $x(0) = x_0$ and $c(0) = 0$. Let us first (unrealistically) allow all (measurable) functions u as admissible control functions. (A negative value for u can be

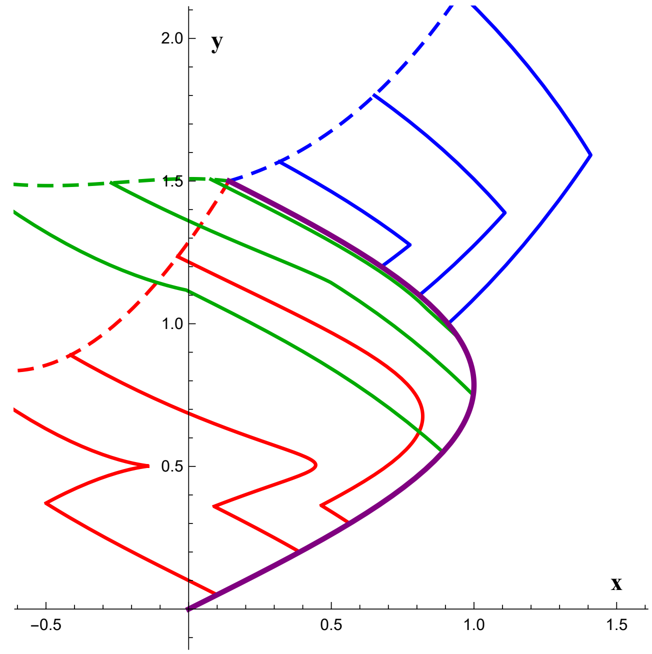

Fig. 14. Reference trajectory and system responses to different needle variations of the reference control, shown in a single diagram.

interpreted as “feeding” the insects rather than fighting them.) The Hamiltonian for the system (45) is

$$H = p(ax-u) + (q/2)\cdot u^2; \tag{46}$$

hence the adjoint equations are given by

$$\begin{aligned}\dot p &= -ap \\ \dot q &= 0\end{aligned} \tag{47}$$

and imply that q is a constant whereas $p(t) = e^{-at}p(0)$. If $u_\star$ is a boundary control for (45), then $u_\star$ maximimizes the Hamiltonian pointwise as a function of u. Now for a maximum of H with respect to u to exist, we must necessarily have $q < 0$; then the maximum is found by letting $0 = \partial H/\partial u = -p + qu$ so that

$$u_\star(t) = \frac{p(t)}{q} = \frac{p(0)}{q}\cdot e^{-at} =: Ce^{-at}. \tag{48}$$

Plugging this into the system equations yields

$$\begin{aligned}x_\star(t) &= e^{at}x_0 - \sigma(e^{at} - e^{-at}) \\ c_\star(t) &= a\sigma^2(1-e^{-2at})\end{aligned} \tag{49}$$

where $\sigma := C/(2a)$. Thus the boundary of the reachable set at time t is

$$\begin{aligned}\partial R_T &= \left\{ \begin{bmatrix} e^{aT}x_0 - \sigma(e^{aT}-e^{-aT}) \\ \sigma^2(1-e^{-2aT}) \end{bmatrix} \mid \sigma\in\mathbb{R} \right\} \\ &= \left\{ (x,c) \Big| \, c = \frac{a}{e^{2aT}-1}(x - e^{aT}x_0)^2 \right\}.\end{aligned} \tag{50}$$

This is an upward-pointing parabola with vertex at the point $(e^{aT}x_0, 0)$ (which corresponds to the control $u\equiv 0$ which leaves the insect population undisturbed). The reachable set R_T is the region over this parabola and includes the parabola itself. Each value of C (or, equivalently, of σ) determines a boundary control $u_\star$ which, in turn, determines a boundary point $\big(x_\star(T), c_\star(T)\big) \in \partial R_T$ and an adjoint vector $\big(p_\star(T), q_\star(T)\big) = \big(e^{-aT}p(0), q\big)^T = -q\cdot\big(-2a\sigma e^{-aT}, -1\big)^T$. This adjoint vector represents a normal vector to the curve ∂R_T pointing away from the reachable set R_T, as is shown in Fig. 15.

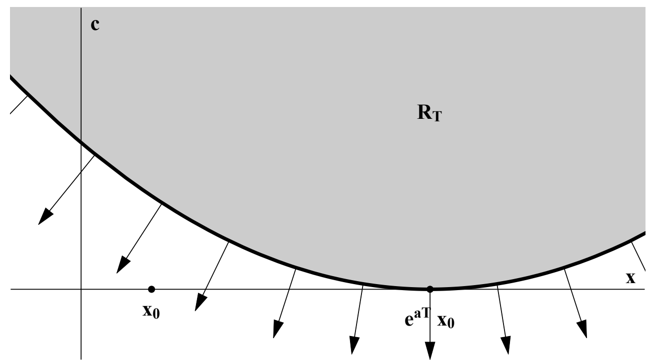

Fig. 15. Reachable set R_T and adjoint vectors attached to different boundary points of R_T.

As T grows larger, this parabola becomes flatter and moves further to the right, as is shown in Fig. 16.

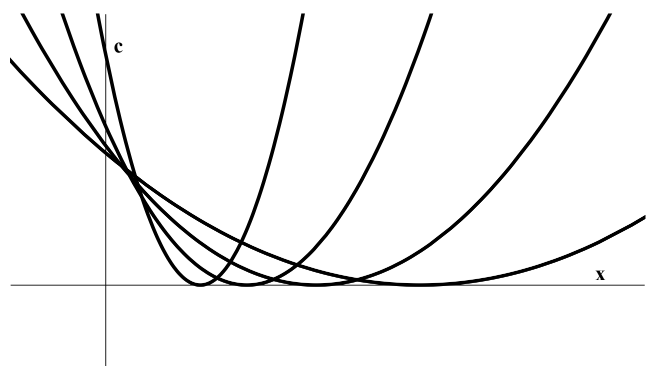

Fig. 16. Boundary of the reachable set R_T for different values of T.

What changes if we use $U = [0,\infty)$ rather than $U = \mathbb{R}$ as the control set, i.e., if only controls with nonnegative values are admissible? The adjoint equations remain unchanged; only the maximum condition is affected. For H to possess a maximum subject to the constraint $u \geq 0$ we must have either $q < 0$ or else ($q = 0$ and $p > 0$). If $q < 0$, the graph of H as a function of u is a downward-pointing parabola with a vertex at p/q. If $p < 0$, this vertex represents an admissible control value, and we find that $u_\star(t) = p(t)/q$ as before. If $p > 0$, this vertex represents a nonnegative and hence inadmissible control value; since only arguments $u \geq 0$ are allowed, H is now maximized for $u = 0$, which yields the boundary control $u_\star \equiv 0$. (In the case $p = 0$ the two controls coincide.) Finally, if $q = 0$ then necessarily $p > 0$, and again $u_\star \equiv 0$. Thus in (48), only controls $u_\star(t) = Ce^{-at}$ with $C \geq 0$ can occur, and in (50) not all $\sigma\in\mathbb{R}$ are allowed, but only values

$\sigma \geq 0$. The boundary of the reachable set consists of a half-parabola and an upward-pointing half-line whose points do not actually belong to the reachable set R_T. (Note that this set is not closed in this example.) The geometrical interpretation is as follows: The vector $\big(p(T), q(T)\big)$ is a normal vector of a support hyperplane of R_T at $\big(x_\star(T), c_\star(T)\big)$ pointing away from R_T. For all boundary points of R_T other than $(e^{aT}x_0, 0)$, this hyperplane is uniquely determined; at the boundary point $(e^{aT}x_0, 0)$ (which is reached by applying the control $u_\star \equiv 0$) all vectors $(p,q)^T \neq (0,0)$ with $q \leq 0$ and $p \geq 0$ are possible. This is shown in Fig. 17.

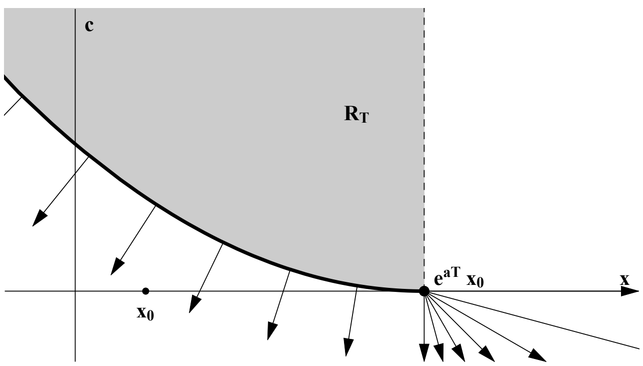

Fig. 17. Reachable set R_T and adjoint vectors attached to different boundary points of R_T if $U = [0, \infty)$.

VII. Three different reference controls

We now want to determine the Boltyanskiĭ cone for each of the following reference controls.

• We leave the insect population to itself, i.e., use the "control" $u_\star \equiv 0$. The resulting system response it shown as the blue trajectory in Fig. 18. Explicitly, we have

$$\begin{aligned} u_\star(t) &= 0, \\ x_\star(t) &= e^{at}x_0, \\ c_\star(t) &= 0. \end{aligned} \tag{51}$$

• We keep the size of the insect population constant so that $x \equiv x_0$, hence $\dot{x} \equiv 0$ and therefore $u_\star \equiv ax - \dot{x} = ax_0$. The system response to this control is shown as the green trajectory in Fig 18. Explicitly, we have

$$\begin{aligned} u_\star(t) &= ax_0, \\ x_\star(t) &= x_0, \\ c_\star(t) &= (a^2x_0^2/2) \cdot t. \end{aligned} \tag{52}$$

• We steer the system in such a way that the target condition $x(T) = 0$ is reached while minimizing the environmental damage $c(T)$. The control $u_\star$ which accomplishes this optimization goal ist obtained by determining the constant σ from letting $x_\star(T) = 0$ in the first equation in (49) and then plugging in $C := 2a\sigma$ into the equation (48); the system response to this control is shown as the red trajectory in Fig. 18. Explicitly, we have

$$\begin{aligned} u_\star(t) &= \frac{ax_0}{\sinh(aT)} \cdot e^{a(T-t)}, \\ x_\star(t) &= \frac{x_0}{\sinh(aT)} \cdot \sinh\big(a(T-t)\big), \\ c_\star(t) &= \frac{ax_0^2 e^{2aT}}{4\sinh(aT)^2} \cdot (1 - e^{-2at}). \end{aligned} \tag{53}$$

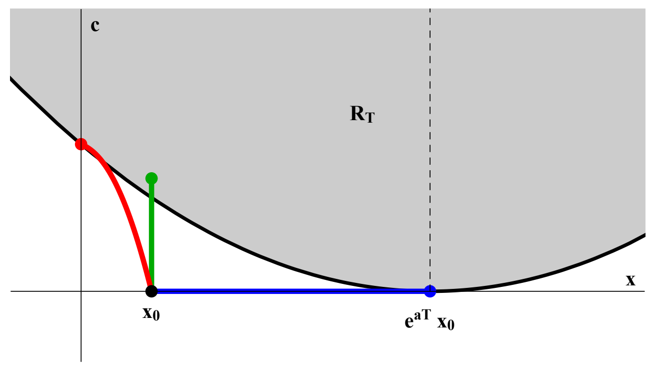

Fig. 18. System responses to the three reference controls considered.

We now consider for each of these three reference controls an arbitrary simple needle variation at time $\tau \in (0, T)$ with control value w.

VIII. Needle variations of the reference controls

• The first reference control and its system response are given by (51). The perturbed control u_ε and its system response $(x_\varepsilon, c_\varepsilon)$ are given as follows.

First phase $(0 \leq t \leq \tau - \varepsilon)$:

$$u_\varepsilon(t) = 0, \qquad x_\varepsilon(t) = e^{at}x_0, \qquad c_\varepsilon(t) = 0 \tag{54}$$

Second phase $(\tau - \varepsilon \leq t \leq \tau)$:

$$\begin{aligned} u_\varepsilon(t) &= w \\ x_\varepsilon(t) &= \frac{w}{a} + \Big(x_0 - \frac{w}{a}e^{-a(\tau-\varepsilon)}\Big)\, e^{at} \\ c_\varepsilon(t) &= \frac{w^2}{2}\big(t - (\tau - \varepsilon)\big) \end{aligned} \tag{55}$$

Third phase $(\tau \leq t \leq T)$:

$$\begin{aligned} u_\varepsilon(t) &= 0 \\ x_\varepsilon(t) &= \Big(x_0 - \frac{w}{a}e^{-a(\tau-\varepsilon)} + \frac{w}{a}e^{-a\tau}\Big)\, e^{at} \\ c_\varepsilon(t) &= \frac{w^2\varepsilon}{2} \end{aligned} \tag{56}$$

The limiting direction from which the curve $\varepsilon \mapsto \big(x_\varepsilon(T), c_\varepsilon(T)\big)$ approaches the point $\big(x_\star(T), c_\star(T)\big) = (e^{aT}x_0, 0)$ as $\varepsilon \to 0$ is given by

$$\frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} \begin{bmatrix} x_\varepsilon(T) \\ c_\varepsilon(T) \end{bmatrix} = \begin{bmatrix} -we^{a(T-\tau)} \\ w^2/2 \end{bmatrix}. \tag{57}$$

Thus the Boltyanskiĭ cone of the reference control considered in this case is spanned by all vectors of the form

$(-we^{a\theta}, w^2/2)^T$ where $\theta \in (0,T)$ and $w \in U$. If $U = \mathbb{R}$ (so that all control values $w \in \mathbb{R}$ are admissible) this is the cone spanned by all vectors $(u,v)^T$ where $v \geq 0$, i.e., the upper half-plane. If $U = [0,\infty)$ (so that only nonnegative control values $w \geq 0$ are admissible) this is the cone spanned by all vectors $(u,v)^T$ where $u \leq 0$ and $v \geq 0$, i.e., the second quadrant. This is fully consistent with the shape of the reachable set as depicted in Fig. 15 and Fig. 17. For one particular choice of w and τ, Fig. 19 shows the trajectories $t \mapsto \big(x_\varepsilon(t), c_\varepsilon(t)\big)$ for various values of ε tending to zero, along with a tangent vector of the curve $\varepsilon \mapsto \big(x_\varepsilon(T), c_\varepsilon(T)\big)$ at $\varepsilon = 0$. (The length of this vector can be arbitrarily scaled by changing the length of the perturbation interval.)

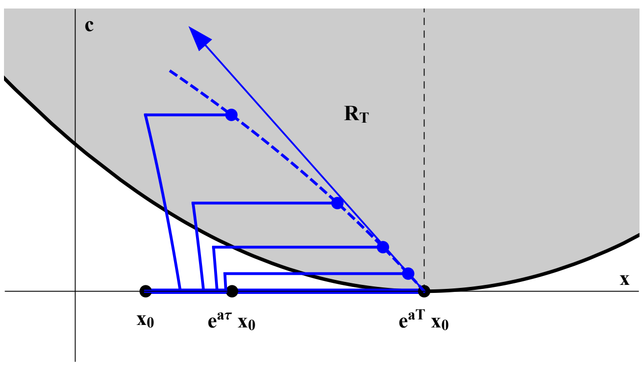

Fig. 19. Results of a needle variation of the reference control $u_\star \equiv 0$.

• The second reference control and its system response are given by (52). The perturbed control u_ε and its system response $(x_\varepsilon, c_\varepsilon)$ are given as follows.

First phase $(0 \leq t \leq \tau - \varepsilon)$:

$$u_\varepsilon(t) = ax_0, \qquad x_\varepsilon(t) = x_0, \qquad c_\varepsilon(t) = \frac{a^2x_0^2\,t}{2} \tag{58}$$

Second phase $(\tau - \varepsilon \leq t \leq \tau)$:

$$\begin{aligned} u_\varepsilon(t) &= w \\ x_\varepsilon(t) &= \frac{w}{a} + \Big(x_0 - \frac{w}{a}\Big)\, e^{-a(\tau-\varepsilon)} \cdot e^{at} \\ c_\varepsilon(t) &= \frac{w^2}{2} \cdot t + \frac{a^2x_0^2 - w^2}{2}(\tau - \varepsilon) \end{aligned} \tag{59}$$

Third phase $(\tau \leq t \leq T)$:

$$\begin{aligned} u_\varepsilon(t) &= ax_0 \\ x_\varepsilon(t) &= x_0 + \Big(x_0 - \frac{w}{a}\Big)\big(-e^{-a\tau} + e^{-a(\tau-\varepsilon)}\big) \cdot e^{at} \\ c_\varepsilon(t) &= \frac{a^2x_0^2}{2} \cdot t + \frac{w^2 - a^2x_0^2}{2} \cdot \varepsilon \end{aligned} \tag{60}$$

The limiting direction from which the curve $\varepsilon \mapsto \big(x_\varepsilon(T), c_\varepsilon(T)\big)$ approaches the point $\big(x_\star(T), c_\star(T)\big) = (x_0, a^2x_0^2T/2)$ as $\varepsilon \to 0$ is given by

$$\frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} \begin{bmatrix} x_\varepsilon(T) \\ c_\varepsilon(T) \end{bmatrix} = \frac{w - ax_0}{2} \cdot \begin{bmatrix} -2e^{a(T-\tau)} \\ w + ax_0 \end{bmatrix} \tag{61}$$

Thus the Boltyanskiĭ cone of the reference control considered in this case is spanned by all vectors of the form $\pm(-2e^{a\theta}, w + ax_0)^T$ where $\theta \in (0,T)$ and $w \in U$. This cone is all of $\mathbb{R}^2$, independently of whether $U = \mathbb{R}$ or $U = [0,\infty)$, which is consistent with the fact that $\big(x_\star(T), c_\star(T)\big)$ is an inner point, not a boundary point of R_T. Fig. 20 shows, for one particular choice of w and τ, the trajectories $t \mapsto \big(x_\varepsilon(t), c_\varepsilon(t)\big)$ for various values of ε tending to zero, along with a tangent vector of the curve $\varepsilon \mapsto \big(x_\varepsilon(T), c_\varepsilon(T)\big)$ at $\varepsilon = 0$. (The length of this vector can be arbitrarily scaled by changing the length of the perturbation interval.)

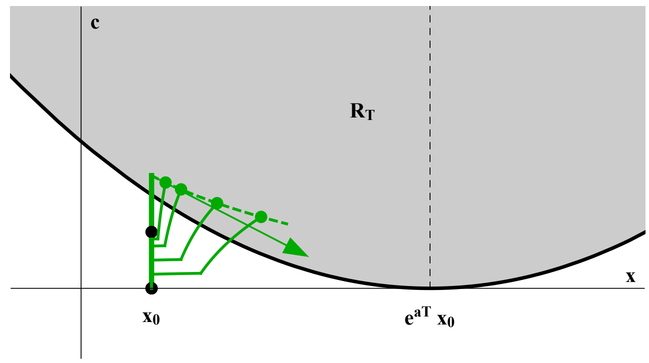

Fig. 20. Results of a needle variation of the reference control $u_\star \equiv ax_0$.

• The third reference control and its system response are given by (53). To avoid cumbersome notation, we introduce the following abbreviations:

$$\begin{aligned} A &= \frac{ax_0e^{aT}}{\sinh(aT)}, & B &= \frac{x_0e^{aT}}{2\sinh(aT)}, \\ C &= \frac{-x_0e^{-aT}}{2\sinh(aT)}, & D &= \frac{ax_0^2e^{2aT}}{4\sinh(aT)^2}. \end{aligned} \tag{62}$$

Then the perturbed control u_ε and its system response $(x_\varepsilon, c_\varepsilon)$ are given as follows.

First phase $(0 \leq t \leq \tau - \varepsilon)$:

$$\begin{aligned} u_\varepsilon(t) &= Ae^{-at} \\ x_\varepsilon(t) &= Be^{-at} + Ce^{at} \\ c_\varepsilon(t) &= D(1 - e^{-2at}) \end{aligned} \tag{63}$$

Second phase $(\tau - \varepsilon \leq t \leq \tau)$:

$$\begin{aligned} u_\varepsilon(t) &= w \\ x_\varepsilon(t) &= \frac{w}{a} + \Big(Be^{-2a(\tau-\varepsilon)} - \frac{w}{a}e^{-a(\tau-\varepsilon)} + C\Big) \cdot e^{at} \\ c_\varepsilon(t) &= \frac{w^2}{2}(t - \tau + \varepsilon) + D(1 - e^{-2a(\tau-\varepsilon)}) \end{aligned} \tag{64}$$

Third phase $(\tau \leq t \leq T)$:

$$\begin{aligned} u_\varepsilon(t) &= Ae^{-at} \\ x_\varepsilon(t) &= Be^{-at} + \frac{w}{a}\Big(e^{-a\tau} - e^{-a(\tau-\varepsilon)}\Big) \cdot e^{at} \\ &\quad + \Big(Be^{-2a(\tau-\varepsilon)} + (C - B)\,e^{-2a\tau}\Big) \cdot e^{at} \\ c_\varepsilon(t) &= \frac{w^2\varepsilon}{2} + D\Big(1 - e^{-2a(\tau-\varepsilon)} + e^{-2a\tau} - e^{-2at}\Big) \end{aligned} \tag{65}$$

The limiting direction from which the curve $\varepsilon \mapsto \big(x_\varepsilon(T), c_\varepsilon(T)\big)$ approaches the point $\big(x_\star(T), c_\star(T)\big) = \big(0, c_\star(T)\big)$ as $\varepsilon \to 0$ is given by

$$\frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} \begin{bmatrix} x_\varepsilon(T) \\ c_\varepsilon(T) \end{bmatrix} = \begin{bmatrix} e^{aT}(Ae^{-2a\tau} - we^{-a\tau}) \\ (1/2)(w^2 - A^2 e^{-2a\tau}) \end{bmatrix} \tag{66}$$

which can be rewritten as

$$\frac{ce^{a\theta} - w}{2} \begin{bmatrix} 2e^{a\theta} \\ -(w + ce^{a\theta}) \end{bmatrix} \tag{67}$$

if we introduce the new constants $\theta := T - \tau$ and $c := ax_0/\sinh(aT)$. The Boltyanskiĭ cone associated with the reference control $u_\star$ is then the cone spanned by all vectors of the form (67) where $w \in U$ and $\theta \in (0, T)$; the formation of one such vector is shown in Fig. 21. It is not obvious (but can be checked) that this is the upper half-plane bounded by the tangent line to ∂R_T at $\big(x_\star(T), c_\star(T)\big)$.

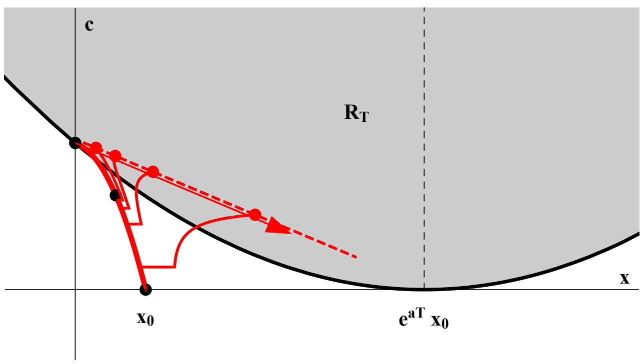

Fig. 21. Results of a needle variation of the reference control $u_\star \equiv Ae^{-at}$.

Fig. 22 shows the Boltyanskiĭ cones of the three reference controls; each arrow represents a direction determined by a specific choice of a needle variation of the control in question. (The green trajectory was artificially extended to make the picture more suitable graphically.)

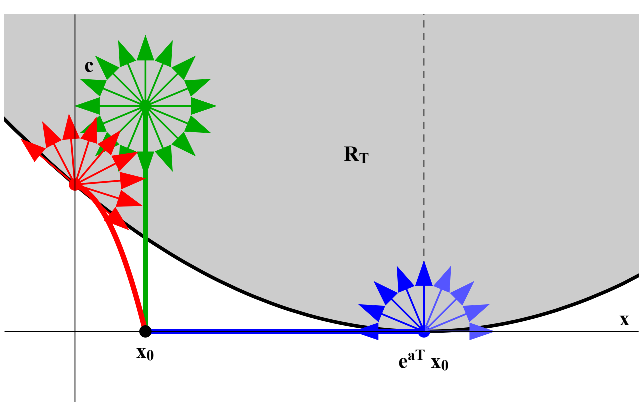

Fig. 22. Boltyanskiĭ cones of the three reference controls.

IX. Concluding Remarks

The chosen examples worked well in classroom use, for several reasons.

• Motivation: The question of identifying the reachable sets of a rather simple control system has the hallmark of a good problem: It is simply stated (to understand it, one does not need more than the concept of an ordinary differential equation), but its solution requires nontrivial tools. Thus the problem provides motivation to develop the necessary theory.

• Visualization: To "see" the action of a boundary control, the effect of needle variations or the formation of a Boltyanskiĭ cone strongly enhances the understanding of the concepts involved and provides geometric intuition for these concepts. (It is not accidental that the words "see" and "understand" can be used synonymously in the English language.)

• Computer use: Doing programming work in class helped in two ways. First, problems in developing software code usually indicated a lack of understanding of underlying concepts, which was overcome by the requirement to write a functioning program. Thus, doing the programming work and understanding the theory worked hand in glove. Second, it was a rewarding experience to finally see the programs run successfully to produce graphic output which, in turn, furthered theoretical understanding.

• Calculations: The understanding of relevant concepts was found to be improved by concrete computations, for example of the effects of different needle variations of a given reference control and of the Boltyanskiĭ cone associated with this reference control. Thus specific examples were not just used to apply the Maximum Principle to see *how* it works, but also to reconstruct its proof in concrete settings to see *why* it works.

The approach of visualizing and clarifying control-theoretical concepts using well-chosen examples turned out to be very beneficial, and the experiences made so far are encouraging for future work (which may address phenomena such as singular controls, abnormal optimizers, chattering control or higher-order optimality criteria).

Acknowledgment

The work presented in this paper is part of a project entitled "Visualization of control-theoretical concepts based on selected examples", which is financially supported by the Klaus Tschira Foundation. This support is gratefully acknowledged.

References

[1] A. A. Agrachev, Yu. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer, Berlin/Heidelberg, 2004
[2] M. Athans, P. L. Falb: *Optimal Control*, McGraw-Hill 1966
[3] E. B. Lee, L. Markus, *Foundations of Optimal Control Theory*, Wiley, New York/London/Sydney, 1967
[4] D. Liberzon, *Calculus of Variations and Optimal Control Theory*, Princeton University Press, Princeton/Oxford 2012
[5] M. Mesterton-Gibbons, *A Primer on the Calculus of Variations and Optimal Control Theory*, American Mathematical Society 2009
[6] L. Naiwert, K. Spindler, *Explicit Calculation of Reachable Sets to Illustrate Concepts in Optimal Control*, 3rd Int. Conf. on Control, Artificial Intelligence, Robotics and Optimization (ICCAIRO 2019), Athens, Dec. 8-10, 2019
[7] L. S. Pontryagin, V. G. Boltyanskiĭ, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Wiley/Interscience, New York, 1962