Journal of Mathematical Sciences, Vol. 100, No. 5, 2000

A NOTE ON CONVEXITY IN SMOOTH NONLINEAR SYSTEMS*

S. A. Vakhrameev UDC 517.977.57; 517.977.1

Introduction

This paper is an extended version of the lecture given by the author at the International Conference Dedicated to the 90th Anniversary of the Birth of L. S. Pontryagin (Moscow, Aug. 31–Sept. 6, 1998) [23]. The results presented below in a short form can be found in [24]. It should be emphasized that the author, being one of the students and collaborators of R. V. Gamkrelidze, is, at the same time, one of the so-called indirect students of L. S. Pontryagin, and, therefore, it was very important to him to participate in this conference. Thus, the author expresses his gratitude to the Organizing Committee for their invitation to participate in the conference and give a lecture at it.

Our goal is twofold. First, we present an existence theorem for the two-point, time-optimal control problem associated with a smooth control system

$$\dot{x} = f(x,u), \quad x \in M, \quad u \in U; \tag{0.1}$$

this theorem improves the existence theorem proved by the author in [20] (see also [22]). In (0.1), M is an n-dimensional smooth (of class C^∞) manifold that is regularly embedded into some Euclidean space $\mathbf{R}^d$; U is a compact convex polyhedron in $\mathbf{R}^m$; $\{f(\cdot,u); u \in \mathbf{R}^m\}$ is a family of smooth (of class C^∞) vector fields on M that smoothly depend on the parameter u in the natural topology (see, e.g., [7, 19, 22]).

Let us briefly explain why this theorem is necessary precisely in the statement presented below.

Recall that in 1994, the author had proved the bang-bang theorem in [18] (see also [21, 22]), which generalizes to the nonlinear (in control and state) case the well-known theorem of R. V. Gamkrelidze on the finiteness of the number of switchings (see, e.g., [14]). But this theorem itself has no sense if we do not have at our disposal an existence theorem under its conditions (or under a part of them). This is the main motivation of the present research.

Second, and this is a byproduct of the search for the proof of the above-mentioned existence theorem, we present a test for convexity, which has a very clear geometrical sense. We mention here two tests for convexity of such a type: the famous Motzkin theorem [11] (see also [9]) and the recent result by A. V. Arutyunov [2].[1]

Recall that the Motzkin theorem consists of the following. Denote by dist(x, K) the *Euclidean distance* from a point $x \in \mathbf{R}^m$ to a nonempty set $K \subset \mathbf{R}^m$. Then the theorem asserts that a nonempty closed set $K \subset \mathbf{R}^m$ is convex iff, for each point $x \in \mathbf{R}^n$, there exists a unique point $y \in K$ that is nearest to x, that is,

$$|x-y| = \operatorname{dist}(K,x),$$

where $|\cdot|$ stands for the Euclidean norm on $\mathbf{R}^m$.

The result by A. V. Arutyunov refers to sets in an arbitrary locally convex space E and is formulated as follows. Let a nonempty set $K \subset E$ be connected and locally convex in the topology of K that is induced by E, and let this set be closed in the topology of E. Then K is convex.

* This work was supported by the Russian Foundation for Basic Research, project No. 96-01-00860.

[1] *Added in proof*: As was revealed later, this result (formulated in somewhat different terms) is already known and is called the *Tietze convexity test*; see J. Cel, "A generalization of Tietze's theorem on local convexity for open sets," *Bull. Soc. Roy. Sci. Liedge*, **67**, No. 1–2, 31–33 (1998) and further references given therein.

Translated from Itogi Nauki i Tekhniki, Seriya Sovremennaya Matematika i Ee Prilozheniya, Tematicheskie Obzory. Vol. 60, Pontryagin Conference-1, Optimal Control, 1998.

 1072-3374/00/1005-2470 $25.00

For other tests for convexity of a geometrical and topological nature, see [4].

The paper is organized as follows.

In the first section, we state the main result, our existence theorem, and consider some particular cases where the proof of this theorem can be directly obtained.

In Sec. 2, we present the test for convexity and give the proof of the main theorem.

Section 3 contains auxiliary lemmas that were used in Sec. 1.

In Sec. 4, we present one more application of our test for convexity. Namely, we study conditions for the convexity of images of matrix exponentials and their applications to the geometry of reachable sets of commutative bilinear systems. Here we restrict ourselves to the consideration of the simplest case; the author hopes to give more comprehensive results in the near future.

In conclusion, we note that there are no difficult theorems in convex analysis; all its results are visual, and the problem consists only of their detection. The author found his test for convexity when he translated into English the famous book by A. D. Aleksandrov Intrinsic Geometry of Convex Surfaces (it may published by Gordon and Breach this year). The reading (and translation) of this remarkable book stimulated him to understand the geometrical nature of many facts which he knew only from the analytical point of view and helped him to find the above-mentioned test for convexity.

As for the existence theorem, it demonstrates the phenomenon of the so-called "implicit convexity" (this term was introduced by V. M. Tikhomirov at the Soviet-Poland International Workshop "Mathematical Methods of Optimal Control and Their Applications," Minsk, May, 1989). Also, we recall here that L. S. Pontryagin himself regarded the existence problem as a very difficult one: "However, it should be noted that from the mathematical viewpoint, the question on the existence of an optimal trajectory seems to be very important and difficult" ([13], p. 189). This is also one of the reasons why the author dealt with the existence problem for several years.

We use the standard notation in this paper. Thus, $\mathbf{R}^n$ always stands for the n-dimensional Euclidean space with the norm $|\cdot|$ generated by the standard inner product $(\cdot,\cdot)$. We identify a vector field X on a smooth manifold M with a derivation of the algebra $C^\infty(M)$ of all smooth functions on M. Recall that a *derivation* D of any $\mathbf{R}$-algebra $\mathcal{A}$ is an $\mathbf{R}$-linear mapping that satisfies the *Leibnitz rule*

$$X(ab) = bX(a) + aX(b) \quad \forall a, b \in \mathcal{A}.$$

The set $\mathrm{Vect}\,(M)$ of all vector fields on M has a natural structure of a Lie algebra with Lie bracket

$$[X,Y] = X \circ Y - Y \circ X \quad \forall X, Y \in \mathrm{Vect}\,(M).$$

Any field $X \in \mathrm{Vect}\,(M)$ defines an $\mathbf{R}$-linear operator $\mathrm{ad}\,X : \mathrm{Vect}\,(M) \to \mathrm{Vect}\,(M)$ by

$$\mathrm{ad}\,XY = [X,Y] \quad \forall Y \in \mathrm{Vect}\,(M).$$

In a natural way, one considers the powers (iterations) of this operator:

$$\mathrm{ad}^0 x = \mathrm{Id}, \quad \mathrm{ad}^m X = \mathrm{ad}\,X \circ \mathrm{ad}^{m-1} X = \mathrm{ad}^{m-1} X \circ \mathrm{ad}\,X, \quad m \geqslant 1,$$

where Id is the identity mapping of $\mathrm{Vect}\,(M)$. By $T_x(M)$, we denote the tangent space to M, and TM stands for the tangent bundle of M. Of course, if M is smoothly embedded into a certain Euclidean space, it is convenient to identify the tangent space T_xM with the affine plane $x + L_x$ in this space, where L_x is a hyperplane in $\mathbf{R}^n$ that is uniquely defined by T_xM. Smooth vector fields on M are also smooth sections of the tangent bundle and, in the above case, can be identified with smooth vector-valued functions (of the corresponding dimension) that are tangent to M at each point of M. If $f : M \to N$ is a smooth mapping of smooth manifolds M and N, then by $f_{x,*}$ we denote the tangent mapping (differential) which maps from $T_x(M)$ into $T_{f(x)}(N)$ and thus defines the smooth mapping $f_* : TM \to TN$ of the corresponding tangent bundles. For a subset $K \in \mathbf{R}^k$, we denote by $\mathrm{conv}\,A$ the convex hull of this subset. Also, $\mathrm{span}\,K$ stands for the linear hull of K, but we prefer to denote it also by $\mathrm{aff}\,K = \mathrm{span}\,K$. By $\mathrm{cl}\,A$ we denote the closure of a set A, and $\mathrm{int}\,(K)$ denotes the interior of a set K, while the notation $\mathrm{rel\,int}(K)$ is used for the relative interior of a *convex* set K in a finite-dimensional space; it is well known that the latter set is always nonempty. For a mapping f, $\operatorname{im} f$ always stands for its image, and for $K \subset \operatorname{dom}(f)$, where $\operatorname{dom}(f)$ is the domain of f, $f^{-1}(K)$ stands for the inverse image of the set K under the mapping f.

Acknowledgments

The author express his gratitude to R. V. Gamkrelidze for his attention to this work and to S. M. Aseev and N. A. Bobylev for a fruitful discussion of the results obtained.

1. Existence Theorem

Let M be a smooth (of class C^∞) manifold smoothly embedded in $\mathbf{R}^d$. Consider a smooth control system

$$\dot{x} = f(x,u), u \in U, x \in M, \tag{1.1}$$

where $f(u), u \in \mathbf{R}^m$, is a family of smooth (of class C^∞) vector fields on M depending smoothly on a parameter $u \in \mathbf{R}$ and U is a *compact convex polyhedron* in $\mathbf{R}^m$. Without loss of generality, we can assume that $\operatorname{aff} U = \mathbf{R}^m$. We assume that there is a function $g : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$$| f(x,u) | \leq g(| x |) \tag{1.2}$$

for all $x \in M, u \in U$ and

$$\limsup_{s\to\infty} \frac{g(s)}{s} < \infty.$$

Under this condition, for any *admissible*, i.e., bounded and measurable control $u(t), t \in \mathbf{R}$, with values in U, the *flow* $p_t(u), t \in \mathbf{R}$, is well defined on the manifold M.

Consider the following *two-point, time-optimal control problem* for system (1.1): *given two points* x_0 *and* x_T, *find an admissible control* $u(t), t \in \mathbf{R}$, *such that the trajectory* $x(t) = x(t; x_0, u(\cdot)) = p_t(u)(x_0), t \in \mathbf{R}$, *of system* (1.1) *corresponding to that control and to the initial position* x_0 ($x(0) = x_0$) *satisfies the condition* $x(T) = x_T$, *and time* T *is the minimal possible.*

The classical existence theorem states that this problem is solvable if the set

$$\mathfrak{F}(x) = \{f(x,u) : u \in U\},$$

which is called the *vectorgram* of the system, is convex for all $x \in M$ (see [6], pp. 203–204). In the present paper, we give the following (explicit) condition for verification of the convexity in the case considered.

Theorem 1.1. *The vectorgram* $\mathfrak{F}(x)$ *is a convex set if the following condition holds: for any edge* Γ *of the polyhedron* U, *there exists a smooth function* $a_\Gamma : M \times U \times \mathbf{R}^m \to \mathbf{R}$ *such that*

$$\frac{\partial^2 f(x,u)}{\partial^2 u}(v,w) = a_\Gamma(x,u,v)\frac{\partial f(x,u)}{\partial u}w \tag{1.3}$$

for all $v \in \mathbf{R}^m, u \in U, x \in M$, *and for any nonzero tangent vector* w *to the relative interior of this edge. Therefore, the existence theorem holds for the considered time-optimal control problem under this condition.*

Remark 1.1. This theorem was proved in [20] (see also [22]) under stronger conditions, in particular, it was required early that condition (1.3) hold for *any* vector $w \in \mathbf{R}^m$. Note that this weakness of the condition is very significant, as the following example shows.

Example 1.1. Consider the following two-dimensional system:

$$\dot{x}^1 = e^u, \quad \dot{x}^2 = e^v,$$

$$(u,v) \in U = \{(u,v) | \, 0 \le u \le 1, 0 \le v \le 1\}.$$

One can directly verify that condition (1.3) holds for this system, while this is not the case when one considers this condition for an *arbitrary* $w \in \mathbf{R}^2$.

Remark 1.2. Note that the function $a_\Gamma(x,u,v)$ depends on the edge Γ but not on a particular nonzero tangent vector w to the relative interior of this edge. Indeed, assume that $a_\Gamma(x,u,v) = a_\Gamma(x,u,v,w)$ in (1.3). Replace the tangent vector w by λw with $\lambda \in \mathbf{R}$. Then

$$\frac{\partial^2 f(x,u)}{\partial^2 u}(v,\lambda w) = \lambda \frac{\partial^2 f(x,u)}{\partial^2 u}(v,w) = \lambda a_\Gamma(x,u,v,w)\frac{\partial f(x,u)}{\partial u}w = a_\Gamma(x,u,v,\lambda w)\frac{\partial f(x,u)}{\partial u}w.$$

If

$$\frac{\partial f(x,u)}{\partial u}w \neq 0$$

(note that we can always assume this; see Proposition 2.1 below), then we have

$$a_\Gamma(x,u,v,\lambda w) = \lambda a_\Gamma(x,u,v,w).$$

Since λ is arbitrary, the latter relation means that the function $a_\Gamma(x,u,v,w)$ does not depend on w, i.e., the function a_Γ is well defined (that is, this function depends only on the edge Γ).

Let us consider some simple cases where the proof of this theorem is a direct verification of the convexity condition. To this end, we take into account the following observations, which will be proved in the next to last section:

Lemma 1.1. *For any two points u' and u'' of the polyhedron U, there is a broken line lying entirely in U and consisting of a finite number of links, each of which is parallel to a certain edge of the polyhedron U that starts at the point u' and terminates at the point u''.*

Lemma 1.2. *For any smooth function $b : \mathbf{R} \to \mathbf{R}$, the boundary-value problem*

$$\begin{cases} \ddot{u} + \dot{u}^2 b(u) = 0, & 0 \le t \le 1, \\ u(0) = 0, & u(1) = 1, \end{cases} \tag{1.4}$$

has a monotonically increasing solution $u(t), 0 \le t \le 1$, satisfying the condition $0 \le u(t) \le 1, 0 \le t \le 1$.

Lemma 1.3. *The equation in the boundary-value problem (1.4) is invariant with respect to the action of any affine transformation of the time axis: if $u(\cdot)$ is a solution to this equation, then, for any function $s(t) = kt + l$, the function $u \circ s(\cdot)$ is also a solution to this equation.*

Lemma 1.4. *If at some point $u \in U$,*

$$\frac{\partial^2}{\partial u^2} f(x,u)(w_1,w_2) \neq 0$$

for two nonzero tangent vectors w_1 and w_2 to the interiors of two edges Γ_1 and Γ_2 of the polyhedron U, respectively, then the vectors

$$\frac{\partial}{\partial u} f(x,u)w_1 \quad \text{and} \quad \frac{\partial}{\partial u} f(x,u)w_2$$

lie on one and the same line $L(x,u) \subset T_x M$. Moreover, this line is independent of $u \in U$:

$$L(x,u) = L(x) \quad \text{for all} \quad u \in U, \quad x \in M.$$

First, we consider the case where the functions a_Γ do not depend on the edge Γ at all:

$$a_\Gamma(x,u,v) = a(x,u,v) \quad \forall x \in M, u \in U, v \in \mathbf{R}^d.$$

Given two arbitrary points u' and u'' in U, we will find a smooth (of class C^2) function $s(\lambda), 0 \le \lambda \le 1$, such that $0 \le s(\lambda) \le 1$, $0 \le \lambda \le 1$, and

$$f(x, s(\lambda)u'' + (1 - s(\lambda))u') = \lambda f(x, u'') + (1 - \lambda) f(x, u'). \tag{1.5}$$

Since $\operatorname{aff} U = \mathbf{R}^m$, we can find, for any $w \in \mathbf{R}^m$, some edges $w_1, \ldots, w_m$ of the polyhedron U such that

$$w = \sum_{i=1}^{m} \alpha_i w_i$$

for certain $\alpha_i \in \mathbf{R}$. Therefore, according to (1.3) with $a_\Gamma(x, u, v) \equiv a(x, u, v)$, we have $\forall v \in \mathbf{R}^m$, $x \in M$,

$$\frac{\partial^2 f(x,u)}{\partial u^2}(v, w) = \frac{\partial^2 f(x,u)}{\partial u^2}(v, \sum_{i=1}^{m} \alpha_i w_i)$$

$$= \sum_{i=1}^{m} \alpha_i \frac{\partial^2 f(x,u)}{\partial u^2}(v, w_i) = \sum_{i=1}^{m} \alpha_i a(x, u, v) \frac{\partial f(x,u)}{\partial u} w_i$$

$$= a(x, u, v) \frac{\partial f(x,u)}{\partial u} \sum_{i=1}^{m} \alpha_i w_i = a(x, u, v) \frac{\partial f(x,u)}{\partial u} w.$$

Thus, condition (1.3) holds for any $w \in \mathbf{R}^m$ with the function $a_\Gamma(x, u, v) \equiv a(x, u, v)$.

Now to find a C^2 function $s(\lambda), 0 \le \lambda \le 1$, for which (1.5) holds, we set $u(\lambda) = s(\lambda)u'' + (1 - s(\lambda)u')$ and differentiate (1.5) with respect to λ:

$$\frac{df(x, u(\lambda))}{d\lambda} = f(x, u'') - f(x, u') = \frac{ds(\lambda)}{d\lambda} \frac{\partial f(x, u(\lambda))}{\partial u}(u'' - u');$$

$$\frac{d^2 f(x, u(\lambda))}{d\lambda^2} = \frac{d^2 s(\lambda)}{d\lambda^2} \frac{\partial f(x, u(\lambda))}{\partial u}(u'' - u') + \Big(\frac{ds(\lambda)}{d\lambda}\Big)^2 \frac{\partial^2 f(x, u(\lambda))}{\partial u^2}(u'' - u', u'' - u').$$

Using (1.3) with $a_\Gamma(x, u, v) \equiv a(x, u, v)$, we have from the latter relation that

$$\Big(\frac{d^2 s(\lambda)}{d\lambda^2} + a(x, u(\lambda), u'' - u')\Big(\frac{ds(\lambda)}{d\lambda}\Big)^2\Big) \frac{\partial f(x, u(\lambda))}{\partial u}(u'' - u') = 0.$$

Now we set $b(u) = a(x, u, u'' - u')$ and use Lemma 1.2. According to this lemma, there is an increasing C^2-function $s(\lambda), 0 \le \lambda \le 1$, $s(0) = 0, s(1) = 1$, such that

$$\frac{d^2 s(\lambda)}{d\lambda^2} + b(s(\lambda))\Big(\frac{ds(\lambda)}{d\lambda}\Big)^2 = 0.$$

Thus, we see that the C^2-curve $u(\lambda) = s(\lambda)u'' + (1 - s(\lambda))u'$ lies entirely in U for $0 \le \lambda \le 1$ and

$$f(x, u(\lambda)) \equiv \lambda f(x, u'') + (1 - \lambda) f(x, u').$$

This proves the convexity of $f(x, U)$ in this simple case.

Consider now the second particular case. Given two arbitrary point u' and u'', we construct a broken line according to Lemma 1.1; let this line consist of N links $[s_i, s_{i+1}], i = 0, 1, \ldots, N s_0 = u', s_N = u''$. Set $\lambda_i = i/N$, and let

$$u(\lambda) = u_i\Big(\frac{\lambda - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}\Big)s_i + \Big(1 - u_i\Big(\frac{\lambda - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}\Big)\Big)s_{i-1}, \quad \lambda_{i-1} \le \lambda \le \lambda_i,$$

where $u_i(\cdot)$ is a solution to (1.4) with

$$b(t) = a_{\Gamma_i}(x, t s_i + (1 - t)s_{i-1}, s_i - s_{i-1})$$

and Γ_i is the edge of the polyhedron U, which is parallel to the link $[s_{i-1}, s_i]$. Then it is easily seen that the curve $u(\lambda)$, $0 \leq \lambda \leq 1$, is a piecewise-smooth, continuous curve lying entirely in U.

Denote by w_i any nonzero tangent vector to the interior of the edge Γ_i of the polyhedron U. In this particular case, we assume that for any $i = 1, \dots, N-1$,

$$\frac{\partial^2}{\partial u^2} f(x, u(\lambda_i))(w_{i-1}, w_i) \neq 0.$$

According to Lemma 1.4, we obtain that at *each* point of $[0,1]$, including the points of discontinuity of the derivative $\dot{u}(\lambda)$, $\lambda \in [0,1]$, the vector

$$\frac{\partial}{\partial u} f(x, u(\lambda))\dot{u}(\lambda)$$

lies on the line $L(x) \subset T_x M$. Indeed, we have

$$\frac{\partial^2}{\partial \lambda^2} f(x, u(\lambda)) = 0$$

at all points of $[0,1]$, except for the points λ_i, $i = 1, \dots, N-1$, by the choice of $u(\lambda), 0 \leq \lambda \leq 1$. At the points where there is no derivative $\dot{u}(\lambda)$, we have that both right and left derivatives of this function belong to this line by Lemma 4. This implies that the image of the mapping $f(x, u(\lambda)), \lambda \in [0,1]$, is contained in this line $L(x)$. Since $u(\lambda), \lambda \in [0,1]$, is continuous, this image is a segment because of the connectedness of any continuous image of the connected set $[0,1]$. Moreover, by construction, this segment contains the points $f(x,u')$ and $f(x,u'')$ and lies entirely in $f(x,U)$. This proves the theorem in this "nonsingular" case. Note that this segment can be several times covered under the mapping $f(x,\cdot)$, i.e., the curve $f(x, u(\lambda)), \lambda \in [0,1]$, can have self-intersections and parts that overlap each other.

To conclude this section, we state the following theorem, which unifies the above theorem and the previous results of the author obtained in [18, 21] (see also [22]).

We say that system (1.1) satisfies the *general-position condition with respect to the polyhedron* U if, for any nonzero tangent vector w to the relative interior of any edge Γ of the polyhedron U, the vectors

$$\frac{\partial f(x,u)}{\partial u} w, \quad \mathrm{ad}\, f(\cdot,u)\frac{\partial f(\cdot,u)}{\partial u}(x)w, \dots, \mathrm{ad}^{n-1} f(\cdot,u)\frac{\partial f(\cdot,u)}{\partial u}(x)w \tag{1.6}$$

are linearly independent for all $x \in M$ and $u \in U$. Of course, this condition is a generalization of the well-known general-position condition, which was initially discovered by R. V. Gamkrelidze and was later (in a rather simple case) formulated by R. Kalman in his controllability conditions for linear systems.

We say that the *strengthened bang-bang conditions* hold for system (1.1) *on some edge* Γ *of the polyhedron* U if, for any compact set $K \subset M$, there exist linear functionals

$$a_\Gamma(x,u,\cdot) : \mathbf{R}^m \to \mathbf{R},$$

$$b^\alpha_{\beta,\Gamma}(x,u',u'',\cdot) : \mathbf{R}^m \to \mathbf{R}^m, \quad \alpha \geq \beta, \quad \alpha = 0, 1, \dots,$$

depending smoothly on $(x,u) \in K \times U$ and $(x,u',u'') \in K \times U \times U$, respectively, such that

$$\frac{\partial^2 f(x,u)}{\partial u^2}(v,w) = a_\Gamma(x,u,v)\frac{\partial f(x,u)}{\partial u} w, \tag{1.7}$$

$$[\frac{\partial f(\cdot,u'')}{\partial u} v, \mathrm{ad}^\alpha f(\cdot,u')\frac{\partial f(\cdot,u')}{\partial u} w](x) = \sum_{\beta=0}^{\alpha} b^\alpha_{\beta,\Gamma}(x,u',u'',v)\mathrm{ad}^\beta f(\cdot,u)\frac{\partial f(\cdot,u')}{\partial u} w(x) \tag{1.8}$$

for all $(x,u) \in K \times U$, $(x,u',u'') \in K \times U \times U)$, $v \in \mathbf{R}^m, \alpha = 0, 1, \dots$, and for any nonzero tangent vector w to the relative interior of the edge Γ. If this condition holds for every edge Γ of the polyhedron U, then we say that system (1.1) satisfies the *strengthened bang-bang conditions with respect to the polyhedron* U.

In the case of a single-input, real-analytic, control-linear system

$$\dot{x} = f_0(x) + uf_1(x), \quad x \in M, u \in \mathbf{R}, \quad |u| \leq 1,$$

this condition is in fact equivalent to the condition Δ by H. Sussmann, which is as follows. Locally, in a neighborhood of each point of M, for each $m = 0, 1, \ldots$, there exist real-analytic functions a_i, $i = 0, 1, \ldots, m$, and a real-analytic function b, $|b(x)| < 1$, in this neighborhood, such that

$$[f_1, \mathrm{ad}^m f_0 f_1](x) = \sum_{i=0}^{m} a_i(x) \mathrm{ad}^i f_0 f_1(x) + b(x) \mathrm{ad}^{m+1} f_0 f_1(x)$$

for all x in that neighborhood. For more details, see [21].

Recall that an *extremal control* is an admissible control that satisfies the Pontryagin maximum principle for the two-point, time-optimal control problem. By an *admissible control,* we mean a measurable function $u(\cdot) : \mathbf{R} \to U$, i.e., we deal with the maximal possible class of admissible controls.

Theorem 1.2. *Let* (1.1) *be a smooth control system that satisfies the strengthened bang-bang conditions* (1.7) – (1.8) *with respect to the polyhedron U and the general-position condition with respect to this polyhedron. Assume that the growth condition* (1.2) *holds.*

Then, if there exists at least one admissible control that steers an initial state x_0 to a final state x_T at some time T, there exists an optimal control that steers x_0 to x_T at the optimal time $\tilde{T} \leq T$, and, moreover, this control is a piecewise-constant function that assumes its values in the set of vertices of the polyhedron U. This control is an extremal control, and, in addition, any extremal control is a piecewise-constant function that assumes its values in the set of vertices of the polyhedron U.

Controls that possess the last property in the above theorem are usually called *the bang-bang controls with a finite number of switchings.* We see that the set of conditions (1.7)–(1.8) includes condition (1.3) of Theorem 1.1 as a particular case, and, in addition, they imply that for any nonzero tangent vector w to the relative interior of any edge Γ of the polyhedron U, we have

$$\frac{\partial f(x,u)}{\partial u} w \neq 0 \quad \text{for any} \quad u \in U, \quad x \in M.$$

In some sense, this theorem gives us a complete solution of the bang-bang problem in nonlinear system theory. In the author's opinion, it is not possible to weaken the conditions of Theorem 1.2, but it is possible to consider more general objects than polyhedrons.

These objects are called *manifolds with corners* and were introduced by A. A. Agrachev and the author in order to generalize Morse theory in such a way that it can be applied to optimal control problems. We refer the reader to [16, 19, 25] for more details. Here we present only the main definitions and the statement of the corresponding theorem on existence and bang-bang properties of an optimal control in this case.

Let N be an m-dimensional smooth manifold. A closed subset $v \subset N$ is called a *submanifold with corners* in N if, for each point $x_0 \in N$, there exists a chart (O_{x_0}, φ) of the manifold N such that $\varphi(x_0) = 0$ and

$$K_{x_0} = \varphi(O_{x_0} \cap V)$$

is a closed polyhedral cone in $\mathbf{R}^m$ with vertex at the origin.

Every polyhedral convex cone is given by a finite set of linear inequalities, and, therefore, submanifolds with corners are locally defined by a finite set of nonlinear inequalities

$$g_i(x) \leq 0, \quad i = 1, 2, .., k,$$

on N that satisfies the "straightening condition" that appears in their definition. Note that these objects are *Whitney stratified spaces* (see [10]). Their stratification is naturally defined by *faces* of submanifolds with corners. An *open face* Γ of a submanifold with corners V is a maximal (with respect to inclusion) smooth connected submanifold of N that lies entirely in V. A *closed face* of V is the closure of a certain open face.

Any l-dimensional stratum $\Gamma^{(l)}$ (an l-dimensional open face) of our stratification admits the following explicit description: it coincides with a certain connected component of the set

$$\{x \in V \mid \dim V \cap (-V) = l\}.$$

One-dimensional faces are one-dimensional strata of the Whitney stratification and are called *edges* of V, while zero-dimensional faces (strata) are called *vertices* of V. Of course, any polyhedron in $\mathbf{R}^n$ is a submanifold with corners in this space.

In a natural way, we can consider system (1.1) in the case where U is now a compact submanifold with corners in a certain smooth manifold N, $\dim N = m$. Furthermore, the general-position conditions and the strengthened bang-bang conditions admit a straightforward generalization to the case of such a system: it suffices to mean by an edge now the edge of the submanifold U with corners. In a similar way, a bang-bang control with a finite number of switchings is now a piecewise-constant function that assumes its values in the set of vertices (zero-dimensional strata) of U. Now we state the result which is an obvious generalization of Theorem 1.2.

Theorem 1.3. *Assume that all conditions of Theorem* 1.2 *hold for the polyhedron U replaced by a certain compact connected submanifold with corners in some smooth manifold N, which is denoted by the same letter. Then there exists an optimal control for the corresponding two-point, time-optimal control problem, every optimal control is an extremal one, and each extremal control is a bang-bang control with a finite number of switchings.*

2. Proof of the Main Theorem

The proof of our theorem in the general case is based on the following observation, which itself is of independent interest.

Recall (see [5]) that the *Clarke tangent cone* of a set $K \subset \mathbf{R}^n$ at a point $x \in K$ is the set of points $v \in \mathbf{R}^n$ such that for any monotonically decreasing sequence of real numbers $t_i \to 0$ and for any sequence $\{x_i\}$ with $x_i \to x$, there is a sequence of vectors $\{v_i\}$ such that $x_i + t_i v_i \in K$ and $v_i \to v$ as $i \to \infty$. We denote this cone by $\tilde{T}K(x)$ and set $TK(x) = x + \tilde{T}K(x), x \in K$. As is known, in the case where the set K is a convex body, the set $TK(x)$ is also called the *support cone* of this body and can also be defined as the closure of all secant lines to this body that are drawn from the point x (see, e.g., [3]).

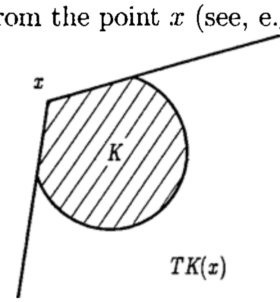

Fig. 1. Support cone to a convex body.

The following "convexity criterion" holds:

Theorem 2.1. (Convexity Criterion). *A closed set K with nonempty interior in $\mathbf{R}^n$ is convex if and only if*

$$K \subset TK(x) \; \forall x \in K.$$

Moreover,

$$K = \bigcap_{x \in K} TK(x).$$

This criterion is an immediate consequence of definitions and the fact that a closed set with nonempty interior is convex iff we can draw a support hyperplane through each of its boundary points (see [9]).

Indeed, the necessity is obvious, since, in this case, the cone $TK(x)$ is the support cone to the closed convex body K and we can refer to [3], where it is proved that $TK(x) \supset K$ for any $x \in K$.

To prove the sufficiency, it suffices to draw a support plane to the cone $TK(x)$ for any $x \in K$. This is always possible, since $TK(x)$ is closed and convex. Since this plane is also a support plane for K, the result follows from Theorem 3.3 in [9].

Remark 2.1. In a private communication, S. M. Aseev told the author that he has proved this test for convexity without the assumption that the interior of the set K is nonempty. Nevertheless, we preserve the original proof, since some of its aspects are of independent interest.

Now the proof of the theorem is a direct verification of this property, where condition (1.3) is substantially used.

It is easy to verify that the Clarke tangent cone to $f(x, U)$ at the point $p = f(x, u) \in f(x, U)$ is

$$\tilde{T}f(x, U)(p) = \frac{\partial f(x, u)}{\partial u}\tilde{T}U(u),$$

where $\tilde{T}U(u)$ is the Clarke tangent cone to U at the point u. This follows from the fact that $f(x, U)$ is a *transversally convex set* (see [17]), since, by condition (1.3), the mapping $u \mapsto f(x, u)$ is a *mapping of constant rank* (see [15]), i.e., the rank of the differential $f_{*,u}$ of this mapping does not depend on $u \in U$.

In fact, we will prove here a more general fact. Obviously, the polyhedron U is a submanifold with corners in $\mathbf{R}^m$, and, therefore, it can be considered as a *Whitney stratified space*: this stratification is defined by (open) faces of U (of all possible dimensions). We assert that the mapping $u \mapsto f(x, u)$ is of constant rank on any stratum $\tilde{U}$ of this stratification, and, moreover, the plane

$$\operatorname{span}\{f_{*,u}w; w \in T_x\tilde{U}\} = \operatorname{im} f_{*,u}|_{\tilde{U}}$$

is independent of $u \in \tilde{U}$.

To prove this, we use the induction on dimension of strata of our stratification.

Indeed, for a one-dimensional stratum (edge of the polyhedron U) this is implied by Lemma 1.4 and Proposition 2.1 below.

For an arbitrary $k, 1 \leq k \leq m$, every $(k+1)$-dimensional stratum (face) of U is obtained from some k-dimensional stratum (face) by adding to it a certain set of vertices that do not belong to this stratum and then by taking the convex hull of this set of vertices and the previous k-dimensional stratum. Therefore, the tangent space to this new $(k+1)$-dimensional stratum is a linear hull of the tangent space of the preceding stratum and a finite set of tangent vectors to some edges of U. But when we apply $f_{*,u}$ to this new tangent space, by Proposition 2.1 and Lemma 1.4, the image of this mapping does not depend on u belonging to this new stratum, since each of the additional vectors $f_{*,u}w$ are either zero or its linear span does not depend on u.

Recall that a set $V \subset N$ of a smooth manifold N is *transversally convex* if for any point $x \in V$, there exist a chart $(O, \phi), \phi : O \to \mathbf{R}^n$, of the manifold N, a closed convex cone $K \subset \mathbf{R}^m$ with vertex at the origin, and a smooth mapping $F : \mathbf{R}^n \to \mathbf{R}^m$ that preserves the origin and is transversal to the cone K, that is,

$$F_{*,x}T_x\mathbf{R}^n + \tilde{T}K(F(x)) = T_{F(x)}\mathbf{R}^m$$

$$\forall x \in \mathbf{R}^n : F(x) \in K,$$

such that

$$\phi(O \cap V) = F^{-1}(K).$$

We note that every submanifold with corners is a transversally convex set; the converse statement is not true in general. The class of transversally convex subsets (which was introduced by the author in order to generalize the Lyusternik–Shnirel'man theory to manifolds with singularities; see [6, 7]) is closed under transversal intersections, contains all closed convex subsets, and, moreover, at least for a compact transversally convex set, the image of any transversally convex set under any surjective submersion is a transversally convex set (see [16, 17, 19]).

Now our conclusion that $f(x, U)$ is a transversally convex subset of M follows from the fact that the image of a transversally convex subset $V \subset N$ under any surjective submersion $F : N \to L$ of smooth manifolds N and L is a transversally convex subset $F(V)$ of L. Indeed, take $N = \mathbf{R}^m$, $V = U$, where U is our polyhedron, which is a manifold with corners, and the more so, a transversally convex subset of $\mathbf{R}^m$ [16]. Also, we see that $f(x, \cdot) : U \to T_xM$ is a mapping of constant rank; thus,

$$f(x, \operatorname{aff} U) = \mathcal{F}(x)$$

is a smooth manifold (see [17], where this fact was proved in the infinite-dimensional setting), and, therefore, the mapping

$$F : u \mapsto f(x, u) : \quad \operatorname{aff} U \to \mathcal{F}(x) \subset T_xM$$

is a surjective submersion.

Since $U \subset \operatorname{aff} U$ is transversally convex, the same is true for its image $f(x, U) = F(U)$.

In the case where $f(x, U)$ is transversally convex, the Clarke tangent cone coincides with the tangent cone defined by using smooth curves lying in this set (see [16]).

To be more precise, the Clarke tangent cone at a point p here is the closure of all vectors tangent to $f(x, U)$ at this point; a vector v is called *tangent* to $f(x, U)$ if there exists a smooth curve $\sigma(\varepsilon), 0 \le \varepsilon \le \varepsilon_0$, such that $\sigma(\varepsilon) \in f(x, U), 0 \le \varepsilon \le \varepsilon_0$, and

$$v = \frac{d}{dt}\sigma(\varepsilon)|_{\varepsilon=0}.$$

Now the above representation for the Clarke tangent cone follows from the chain rule for the differentiation.

Let $q = f(x, v) \in f(x, U)$ be an arbitrary point. We have to prove that $q \in Tf(x, U)(p)$, i.e., that

$$f(x, v) \in f(x, u) + \frac{\partial f(x, u)}{\partial u}\tilde{T}U(u).$$

We have

$$f(x, v) = f(x, u) + \int_0^1 \frac{\partial f(x, sv + (1 - s)u)}{\partial u}(v - u)\, ds.$$

By definition, $w = v - u \in \tilde{T}U(u)$. Let us consider the tangent cone $\tilde{T}U(u)$. Since U is a polyhedron, this cone can be represented as the sum $K_1 \oplus K_2$, where K_1 is a pointed polyhedral cone and $K_2 = \tilde{T}U(u) \cap (-\tilde{T}U(u))$ is the maximal linear subspace lying in $\tilde{T}U(u)$. Both these cones are generated by vectors that have directions of some edges of the polyhedron U, that is, by tangent vectors to the interiors of these edges of the polyhedron U: the cone K_1 is generated by nonnegative linear combinations of some finite set of these vectors, while the cone K_2 is generated by linear combinations of another set of them, and these sets can be chosen so that for the corresponding vectors, the coefficients of these linear combinations are uniquely defined. Let $w_1, \dots, w_k$ be these vectors (the first l of them generate K_1 and the other vectors generate K_2). Thus, any vector $w \in \tilde{K}U(u)$ is represented in the form

$$w = \sum_{i=1}^{k} \alpha_i w_i. \tag{2.1}$$

We now show that there are smooth nonnegative functions $u_j(t), 0 \le t \le 1, i = 1, \dots, k$, such that

$$\frac{\partial f(x, sv + (1 - s)u)}{\partial u}w_j = u_j(s)\frac{\partial f(x, u)}{\partial u}w_j, \quad 0 \le s \le 1. \tag{2.2}$$

We assume first that

$$\frac{\partial f(x,u)}{\partial u} w_j \neq 0, \quad j = 1, \dots, k, \tag{2.3}$$

and then omit this assumption.

Now, differentiating (2.2), we obtain

$$\frac{\partial}{\partial s} f(x, sv + (1-s)u) w_j = \frac{\partial^2}{\partial u^2} f(x, sv + (1-s)u)(w_j, w)$$

$$= a_j(s) \frac{\partial}{\partial u} f(x, sv + (1-s)u) w_j = a_j(s) u_j(s) \frac{\partial}{\partial u} f(x,u) w_j = \dot{u}_j(s) \frac{\partial}{\partial u} f(x,u) w_j,$$

where $w = v - u$ and the functions

$$a_j(s) = a_{\Gamma_j}(x, sv + (1-s)u, w)$$

are taken from condition (1.3). This implies

$$\dot{u}_j(s) = a_j(s) u_j(s), 0 \leq s \leq 1,$$

and, therefore, taking into account the initial condition $u_j(0) = 1$, we obtain

$$u_j(t) = \exp \int_0^s a_j(s)\, ds \geq 0, \ j = 1, \dots, k, \quad 0 \leq t \leq 1. \tag{2.4}$$

Thus, we have

$$f(x,u) = f(x,v) + \int_0^1 \frac{\partial f(x, sv + (1-s)u)}{\partial u} w\, ds$$

$$= f(x,u) + \sum_{i=1}^{k} \int_0^1 \exp \int_0^s a_j(\xi)\, d\xi \frac{\partial f(x,u)}{\partial u} \alpha_j w_j\, ds$$

$$= f(x,u) + \sum_{i=1}^{k} \int_0^1 u_i(s)\, ds \alpha_i \frac{\partial f(x,u)}{\partial u} w_i = f(x,v) + \frac{\partial f(x,u)}{\partial u} \tilde{w},$$

where

$$\tilde{w} = \sum_{i=1}^{k} \tilde{\alpha}_j w_i \in \tilde{T}U(u),$$

since

$$\tilde{\alpha}_j = \alpha_j \int_0^1 \exp \int_0^s a_j(\xi)\, d\xi\, ds = \alpha_j \int_0^1 u_j(s)\, ds$$

are nonnegative for $j = 1, \dots, l$.

Now, to complete the proof, it is necessary to verify that the set $f(x, U)$ has a nonempty relative interior. To be more precise, we show that there exists an affine plane $\Pi(x) \subset \mathbf{R}^d$ that contains $f(x, U)$ in which our set $f(x, U)$ has a nonempty interior. We give the following explicit description of this plane $\Pi(x)$:

$$\Pi(x,u) = f(x,u) + \text{span}\ \{\frac{\partial f(x,u)}{\partial u} w_j; j = 1, ..., N\} = \Pi(x),$$

where w_j are all nonzero tangent vectors to the interiors of all edges of the polyhedron U.

First, using condition (1.3), we verify that the plane $\tilde{\Pi}(x,u) = \text{span}\ \{\frac{\partial f(x,u)}{\partial u} w_j; j = 1, ..., N\}$ does not depend on

$$u \in \text{aff}\ U = \text{span}\ \{w_j; j = 1, ..., N\} = \mathbf{R}^m;$$

this is already proved when we verify that the mapping $u \mapsto f(x,u)$ is a mapping of constant rank on each stratum of U (it suffices to use the concluding step of induction: $\Pi(x,u)$ is the tangent space to $\operatorname{int} U$, the maximum stratum of the stratification, and, by the way, this can also be done in full analogy with [15], pp. 148–150, where this fact was proved in a substantially more complex infinite-dimensional case). Thus, $\tilde{\Pi}(x,u) = \tilde{\Pi}(x)$. Then, arguing as above, we prove that

$$f(x, \mathbf{R}^m) \subset f(x,u) + \tilde{\Pi}(x).$$

Finally, we see that for any $u, v \in \operatorname{aff} U$,

$$f(x,v) + \tilde{\Pi}(x) = f(x,u) + \tilde{\Pi}(x).$$

Indeed, let $z \in f(x,u) + \tilde{\Pi}(x)$; then

$$z = f(x,u) + \sum_{i=0}^{N} \alpha_i \frac{\partial f(x,u)}{\partial u} w_i$$

for certain $\alpha_i \in \mathbf{R}$ and tangent vectors w_i to edges of the polyhedron ($i = 1, \ldots, N$). Then, for any $v \in U$, we have

$$\begin{aligned} z &= f(x,v) + f(x,u) - f(x,v) + \sum_{i=0}^{N} \alpha_i \frac{\partial f(x,u)}{\partial u} w_i \\ &= f(x,v) + \int_0^1 \frac{\partial(x, su + (1-s)v)}{\partial u}(u - v)\, ds + \sum_{i=0}^{N} \alpha_i \frac{\partial f(x,u)}{\partial u} w_i. \end{aligned}$$

Since $(u - v) \in \operatorname{aff} U$, we find $\tilde{\alpha}_i$, $i = 1, \ldots, N$, such that

$$u - v = \sum_{i=1}^{N} \tilde{\alpha}_i w_i.$$

Therefore,

$$z = f(x,v) + \sum_{i=1}^{N} \tilde{\alpha}_i \int_0^1 \frac{\partial f(x, su + (1-s)v)}{\partial u} w_i\, ds + \sum_{i=0}^{N} \alpha_i \frac{\partial f(x,u)}{\partial u} w_i.$$

Proceeding as above, we see that there exist positive smooth functions $u_i(s), 0 \le s \le 1$, such that

$$\frac{\partial f(x, su + (1-s)v)}{\partial u} w_i = u_i(s) \frac{\partial f(x,u)}{\partial u} w_i, \quad i = 1, \ldots, N.$$

Thus,

$$z = f(x,v) + \sum_{i=0}^{N} \beta_i \frac{\partial f(x,u)}{\partial u} w_i \in f(x,u) + \tilde{\Pi}(x),$$

where

$$\beta_i = \alpha_i + \int_0^s u_i(s)\, ds \tilde{\alpha}_i, \quad i = 1, \ldots, N.$$

Consequently,

$$f(x,u) + \tilde{\Pi}(x) \subset f(x,v) + \tilde{\Pi}(x).$$

Interchanging u and v, we obtain the converse inclusion

$$f(x,v) + \tilde{\Pi}(x) \subset f(x,u) + \tilde{\Pi}(x).$$

Therefore,

$$f(x, \mathbf{R}^m) \subset f(x,u) + \tilde{\Pi}(x).$$

Thus, our affine plane $\Pi(x,u) = \Pi(x)$ is well defined. Now, similarly to [17], pp. 164–166 (where once again a more complex infinite-dimensional case is considered), we confirm that $f(x, \operatorname{rel\,int} U)$ is an open subset of Π; here the fact that $x \mapsto f(x,u)$ is a mapping of constant rank plays a crucial role.

Now we omit the assumption that

$$\frac{\partial f(x,u)}{\partial u} w \neq 0$$

for the nonzero tangent vectors w to the interior of the corresponding edges of the polyhedron U, which were considered above. Meanwhile, we note that in the framework of Theorem 1.2, this assumption always holds because of the general-position condition. Nevertheless, we assert the following:

Proposition 2.1. *Assume that the conditions of Theorem* 1.1 *hold. Let* Γ *be an arbitrary edge of the polyhedron* U, *and let* w *be a nonzero tangent vector to the interior of this edge. Then either*

$$\frac{\partial f(x,u)}{\partial u} w \neq 0 \tag{2.5}$$

for all $u \in U$ *or*

$$\frac{\partial f(x,u)}{\partial u} w \equiv 0.$$

Proof. We can assume that there is a point $u_0 \in U$ such that (2.5) holds for $u = u_0$ (if there is no such point, there is nothing to prove). Let u_1 be an arbitrary point of the polyhedron U, and let $u_1 \neq u_0$. We have to prove that (2.5) holds for $u = u_1$. Since U is convex, we can connect the points u_0 and u_1 by a segment $u(s) = su_1 + (1-s)u_0, 0 \leq s \leq 1$, so that $u(0) = u_0$, $u_1(1) = 1$, and $u(s) \in U, 0 \leq s \leq 1$.

As above, by using condition (3.1), we see that there is a positive C^1-function $\alpha(s), 0 \leq s \leq 1$, such that

$$\frac{\partial f(x,u(s))}{\partial u} w = \alpha(s) \frac{\partial f(x,u_0)}{\partial u} w. \tag{2.6}$$

Indeed, differentiating (2.6) and using (1.3), we have

$$\frac{d}{ds} \frac{\partial f(x,u(s))}{\partial u} w = \frac{\partial^2 f(x,u(s))}{\partial u} (\dot{u}(s), w)$$

$$= a_\Gamma(x, u(s), u_1 - u_0) \frac{f(x,u(s))}{\partial u} w = \alpha(s) a_\Gamma(x,u(s),u_1-u_0) \frac{\partial f(x,u_0)}{\partial u} w = \dot{\alpha}(s) \frac{\partial f(x,u_0)}{\partial u} w,$$

so that we have the Cauchy problem

$$\dot{\alpha}(s) = \alpha(s) a_\Gamma(x,u(s),u_1-u_0), \quad \alpha(0) = 1,$$

which, obviously, has a unique positive smooth solution

$$\alpha(s) = e^{\int_0^s a_\Gamma(x,u(t),u_1-u_0)\,dt}, \quad 0 \leq s \leq 1.$$

The proposition is proved.

3. Proofs of the Lemmas

Here we prove Lemmas 1.1–1.4.

Proof of Lemma 1.1. Let u' and u'' be two arbitrary points of the polyhedron U. First, we show that there exists a broken line L', each link of which is parallel to a certain edge of U that connects the point u' with a certain vertex e' of the polyhedron U and lies entirely in this polyhedron. Draw any line L_1 through the point u' parallel to some edge of the polyhedron U. This line intersects the boundary of this polyhedron at one point, say u_1, which belongs to the boundary of U. Thus, the point u_1 lies in a certain face U_1 of the polyhedron U; this face is also a polyhedron of dimension less than the dimension of U. Draw a second line L_2 through u_1 that is parallel to a certain edge of U_1 (of course, each edge of U_1 is an edge of U) and consider the point u_2 at which this line intersects the boundary of U_1. Iterating this process, we obtain a finite sequence of points $u_0 = u'$, $u_1, \dots,$ $u_{N'} = e'$, where e' is a certain vertex of U. Thus, we have constructed the desired broken line L': the links of this broken line are $[u_{i-1}, u_i]$, $i = 0, \dots, N'$. Second, proceeding in the same way, we obtain an analogous broken line L'' that connects the point u'' with a certain vertex e'' of the polyhedron U. Finally, according to [12], we can connect two vertices e' and e'' by a broken line L''' whose links are exactly the edges of the polyhedron U. Thus, the union of the broken lines L', L''', and L'' yields the desired broken line L. Lemma 1.1 is proved.

Proof of Lemma 1.2. We will use the direct variational method. Let $W^{1,\infty}[0,1]$ be the (Sobolev) space of all absolutely continuous functions with essentially bounded derivative on the closed interval $[0,1]$ endowed with the usual norm

$$||u||_{W^{1,\infty}} = ||u||_{C[0,1]} + ||\dot{u}||_{L^\infty[0,1]}.$$

Consider the functional

$$J(u) = \frac{1}{2}\int_0^1 (\dot{u})^2 \exp\Big(\int_0^u 2b(u(\theta))d\theta\Big)\, dt$$

with domain $\mathfrak{M} = \{u \in W^{1,\infty}[0,1] : u(0) = 0,\ u(1) = 1\}$. This functional has the integrand

$$L(u, \dot{u}) = \frac{1}{2}(\dot{u})^2 \exp\int_0^u 2b(u(\theta))\, d\theta$$

satisfying the *coercivity condition*

$$\frac{v}{L(u,v)} \to \infty \text{ as } |v| \to \infty,$$

and which is *regular,* i.e., the function L is bounded from below and is convex with respect to the second argument (cf. [6]). Therefore, the variational problem

$$J(u) \to \inf,\ u \in \mathfrak{M},$$

has a solution $u(t)$, $0 \le t \le 1$. First, we show that this solution is of class $C^2[0,1]$. Indeed, for any t, $0 \le t \le 1$,

$$L_{\dot{u}\dot{u}}(u(t), \dot{u}(t)) = \exp\int_0^{u(t)} 2b(u(\theta))\, d\theta > 0;$$

therefore, by the Hilbert theorem [1], $u \in C^2[0,1]$. Hence we can write the Euler–Lagrange equation for this extremal in the differential form

$$L_u(u(t), \dot{u}(t) - \frac{d}{dt}L_{\dot{u}}(u(t), \dot{u}(t)) = 0.$$

A direct computation shows that

$$L_u(u(t),\dot u(t)) - \frac{d}{dt}L_{\dot u\dot u}(u(t),\dot u(t))$$

$$= b(u(t))(\dot u(t))^2 \exp \int_0^{u(t)} 2b(u(\theta))\,d\theta - \ddot u(t) \exp \int_0^{u(t)} 2b(u(\theta))\,d\theta$$

$$-2b(u(t))(\dot u(t))^2 \exp \int_0^{u(t)} 2b(u(\theta))\,d\theta$$

$$= -\exp \int_0^{u(t)} 2b(u(\theta))\,d\theta [b(u(t))(\dot u)^2 + \ddot u(t)] = 0.$$

Since

$$\exp \int_0^{u(t)} 2b(u(\theta))\,d\theta > 0,$$

we have that this extremal is a solution to our boundary-value problem.

Now it remains to prove that this solution $u(t),\ 0 \le t \le 1$, satisfies the condition $0 \le u(t) \le 1,\ 0 \le t \le 1$. To prove this assertion, observe that the complete solution of the Cauchy problem

$$\begin{cases} \ddot u + b(u)(\dot u)^2 &= 0, \quad 0 \le t \le 1, \\ u(0) &= 0 \end{cases}$$

has the closed form of the integral equation

$$x(t) = \int_0^t \frac{d\tau}{\int_0^\tau b(x(\theta))\,d\theta - c},$$

where $c = \text{const}$. Since our solution $u(t),\ 0 \le t \le 1$, is of class $C^2[0,1]$, we have that the first derivative

$$\frac{d}{dt}u(t) = \frac{1}{\int_0^t b(u(\theta))\,d\theta - c_0},$$

where c_0 is the constant corresponding to this solution, exists for any $t,\ 0 \le t \le 1$, but this means that the function

$$\int_0^t b(u(\theta))\,d\theta - c_0,\ 0 \le t \le 1,$$

cannot change its sign because of continuity. Of course, this implies the monotonicity of the function $u(t),\ 0 \le t \le 1$; since $u(0) = 1$ and $u(1) = 1$, we have that $0 \le u(t) \le 1,\ 0 \le t \le 1$, and, certainly, $\dot u(t) > 0$ for $0 \le t \le 1$. This completes the proof of the lemma.

Proof of Lemma 1.3. This is a direct computation:

$$\frac{d}{dt}u(kt+l) = k\dot u(kt+l),$$

$$\frac{d^2}{dt^2}u(kl+l) = k^2\ddot u(kl+t);$$

thus, for $y(t) = u(kl+t)$, we have

$$\ddot y(t) + \dot y(t)^2 b(y(t)) = k^2\ddot u(kt+l) + k^2\dot u^2(kt+l)b(u(kt+l)) = k^2(\ddot u(kt+l) + \dot y^2(kt+l)b(u(kt+l)) = 0.$$

Proof of Lemma 1.4. Let $u \in U$ be such that

$$\frac{\partial^2 f(x,u)}{\partial u^2}(w_1, w_2) \neq 0 \tag{3.1}$$

for certain nonzero tangent vectors w_1 and w_2 to the interior of two edges Γ_1 and Γ_2 of the polyhedron U, respectively. According to condition (1.3), we have

$$\frac{\partial^2 f(x,u)}{\partial u^2}(w_1, w_2) = a_{\Gamma_2}(x,u,w_1)\frac{\partial f(x,u)}{\partial u}w_2$$

$$= \frac{\partial^2 f(x,u)}{\partial u^2}(w_2, w_1) = a_{\Gamma_1}(x,u,w_2)\frac{\partial f(x,u)}{\partial u}w_1.$$

Since $\frac{\partial^2 f(x,u)}{\partial u^2}(w_1, w_2) \neq 0$, we have that both $a_{\Gamma_1}a(x,u,w_2)$ and $a_{\Gamma_2}a(x,u,w_1)$ are different from zero. Therefore, we can divide both sides of the previous relation, e.g., by $a_{\Gamma_1}(x,u,w_2)$. This yields

$$\frac{\partial f(x,u)}{\partial u}w_1 = \lambda\frac{\partial f(x,u)}{\partial u}w_2,$$

where

$$\lambda = \frac{a_{\Gamma_2}(x,u,w_1)}{a_{\Gamma_1}(x,u,w_2)}.$$

This means that the vectors

$$\frac{\partial f(x,u)}{\partial u}w_1 \quad \text{and} \quad \frac{\partial f(x,u)}{\partial u}w_2$$

lie on one and the same line $L(x,u)$.

Now let us prove that this line does not depend on $u \in U$:

$$L(x,u) = L(x) \quad \forall u \in U, \quad x \in M.$$

Choose two arbitrary points u' and u'' such that at least at one of this point relation (3.1) holds and show that

$$L(x,u') = L(x,u'').$$

To this end, we first note that both $\frac{\partial f(x,u')}{\partial u}w$ and $\frac{\partial f(x,u'')}{\partial u}w$, with w being equal to w_1 or w_2, are different from zero. Indeed, this is implied by Proposition 2.1. Moreover, proceeding as above in Sec. 2 when proving Proposition 2.1, we see that there is a smooth positive function $\alpha(s), 0 \leq s \leq 1$, such that

$$\frac{\partial f(x, su'' + (1-s)u')}{\partial u}w = \alpha(s)\frac{\partial f(x,u')}{\partial u}w, \quad 0 \leq s \leq 1,$$

In particular, the vectors

$$\frac{\partial f(x,u')}{\partial u}w \quad \text{and} \quad \frac{\partial f(x,u'')}{\partial u}w$$

are collinear. But the linear spans of these vectors are the lines $L(x,u')$ and $L(x,u'')$, respectively, and therefore $L(x,u') = L(x,u'')$. The lemma is proved.

4. On the Convexity of Images of Matrix Exponentials and Reachable Sets of Certain Bilinear Systems

In this section, we present one more application of our convexity criterion to control theory, which seems to be of independent interest.

Let $M_n(\mathbf{R})$ be the linear space of all square real matrices of order n.

Recall that for any $n \times n$-matrix A, the *matrix exponential* e^A is defined by the series

$$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!} + \ldots,$$

where I is the identity matrix.

We say that a set $\mathcal{A}$ of $n \times n$ matrices is *commutative* if each $A, B \in \mathcal{A}$ commute with each other:

$$[AB] = BA - AB = 0 \tag{4.1}$$

We take such a matrix commutator in order to put in correspondence the notation used. Indeed, to any $n \times n$ matrix A, there corresponds the linear vector field $\overrightarrow{Ax}$, and, as a direct calculation shows, we have

$$[\overrightarrow{Ax}, \overrightarrow{Bx}] = \overrightarrow{[A, B]x},$$

where on the right-hand side we have the linear vector field on $\mathbf{R}^n$ defined by the matrix commutator in (4.1).

As is well known,

$$e^A e^B = E^{A+B} = e^B e^A \quad \forall A, B \in \mathcal{A}$$

iff $\mathcal{A}$ is a commutative family of matrices.

Theorem 4.1. *Assume that $\mathcal{A} \subset M_n(\mathbf{R})$ is a compact, connected, transversally convex, commutative family of matrices such that for any $A_0 \in \mathcal{A}$,*

$$(A - A_0)^k \subset \tilde{T}\mathcal{A}(A_0)$$

for every $A \in \mathcal{A}$ and $n = 1, 2, \ldots$. Then the set

$$e^{\mathcal{A}} = \{e^A : A \in \mathcal{A}\}$$

is a closed convex subset of $M_n(\mathbf{R})$.

Proof. To prove the theorem, we will use our test for convexity (Theorem 2.1). First, we see that the Clarke tangent cone to the set $e^{\mathcal{A}}$ at any point $e^{A_0} \in e^{\mathcal{A}}$ is equal to

$$\tilde{T}e^{\mathcal{A}}(e^{A_0}) = e^{A_0}\tilde{T}\mathcal{A}(A_0). \tag{4.2}$$

Indeed, the set $e^{\mathcal{A}}$ is a compact subset in $M_n(\mathbf{R})$; moreover, it has a nonempty interior in the so-called *derived orbit* of $\mathcal{A}$, where $\mathcal{A}$ is considered as a family of (invariant) vector fields on $Gl_n(\mathbf{R})$, and is the closure of its interior points. This follows from the general theorems on orbits of vector fields (see, e.g., [7]).

In fact, this set is a transversally convex compact subset of the derived orbits of $\mathcal{A}$, and, therefore, the Clarke tangent cone and the tangent cone generated by tangent vectors defined by smooth curves in $e^{\mathcal{A}}$ coincide with each other. To prove this fact, it suffices to note that the restriction of the mapping e to the *linear space* $M^c_m(\mathbf{R})$ consisting of all commutative matrices is a mapping of constant rank. Therefore, the image $\operatorname{im} e$ of this mapping is a smooth submanifold in the general linear group $Gl_n(\mathbf{R})$ (in fact, in the corresponding derived orbit $\mathfrak{M}$ of $\mathcal{A}$ that passes through the point e^A, where $A \in \mathcal{A}$ is arbitrary). Therefore, e is a smooth submersion of $M^c_n(\mathbf{R})$ onto $\mathfrak{M}$. On the other hand, the set $\mathcal{A}$ is a compact transversally convex set consisting of commutative matrices. Therefore, it is a transversally convex set in $M^c_n(\mathbf{R})$; thus, according to [17], its image $s^{\mathcal{A}}$ is a transversally convex subset of $\mathfrak{M}$.

Therefore, to compute the Clarke tangent cone to this set, it suffices to compute the ordinary tangent vectors defined by smooth curves in $e^{\mathcal{A}}$ and then take their closure. Let $A(\varepsilon), \varepsilon \in [0, \varepsilon_0]$, be a smooth curve in $\mathcal{A}$ such that $A(0) = A_0$. Of course, it defines a smooth curve in $e^{\mathcal{A}}$, and any smooth curve in the latter set can be obtained in such a way. Using the assumption on the commutativity of the set $\mathcal{A}$, we have

$$\frac{d}{d\varepsilon} e^{A(\varepsilon)}|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[e^{A(\varepsilon)} - e^{A_0}\right] = e^{A_0} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[e^{-A_0} e^{A(\varepsilon)} - I\right]$$

$$= e^{A_0} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[e^{A(\varepsilon)-A_0} - I\right] = e^{A_0} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[\varepsilon(A(\varepsilon) - A_0) + o(\varepsilon)] = e^{A_0} \frac{dA(\varepsilon)}{d\varepsilon}|_{\varepsilon=0}.$$

This implies representation (4.2).

Now, for any $A \in \mathcal{A}$, we have

$$e^A = e^{A_0 + (A - A_0)} = e_0^A e^{A - A_0}.$$

By the assumption of the theorem, the matrix $(A - A_0)^k \in \tilde{T}\mathcal{A}(A)$ for any $k = 0, 1, \ldots$ ($A - A_0 \in \tilde{T}\mathcal{A}(A_0)$ by definition). Now

$$e^{A - A_0} = \sum_{i=0}^{\infty} \frac{1}{i!}(A - A_0)^i.$$

Each term of the series

$$\sum_{i=1}^{\infty} \frac{1}{i!}(A - A_0)^i$$

belongs to $\tilde{T}\mathcal{A}(A_0)$ by assumption; in turn, the convexity and closedness of the Clarke tangent cone imply that the sum of this series is also in it. Thus, for any $A, A_0 \in \mathcal{A}$, we have

$$e^A = e^{A_0} e^{A - A_0} = e^{A_0}\Big(I + \sum_{i=1}^{\infty} \frac{1}{i!}(A - A_0)^i\Big) \in e^{A_0} + e^{A_0}\tilde{T}\mathcal{A}(A_0) = Te^{\mathcal{A}}(e^{A_0}),$$

i.e.,

$$e^{\mathcal{A}} \subset e^{A_0} + Te^{\mathcal{A}}(e^{A_0}).$$

All conditions of our test for convexity hold. The theorem is proved.

It seems to be a very interesting problem to describe classes of matrices for which the above condition holds. The author intends to deal with this problem in the near future.

We now apply this result to the following bilinear system (left-invariant system on $\mathrm{GL}_n(\mathbf{R})$):

$$\dot{X} = XA(u), \quad u \in U, \tag{4.3}$$

where U is an arbitrary nonempty compact set in $\mathbf{R}^m$. We assume that the mapping $A : \mathbf{R}^m \to \mathrm{M}_n(\mathbf{R})$ is linear, and, moreover, for each $u', u'' \in U$,

$$[A(u'), A(u'')] = 0.$$

Note that whenever this condition holds our system is also a right-invariant system; such systems are also called *invariant systems* on the matrix Lie group (general linear group $\mathrm{Gl}_n(\mathbf{R})$). We take the set of all measurable essentially bounded functions $u : \mathbf{R} \to U$ as the class $\mathfrak{D}(U)$ of admissible controls.

We restrict ourselves to this simple case in order to illustrate our results only; a wider scope of examples will be presented in forthcoming publications.

Now we want to find conditions under which the *reachable set*

$$\mathfrak{A}(t; U) =$$

$$\{X(t; u) | X(\cdot; u) \quad \text{is the trajectory of (4.3) corresponding to an admissible control} \quad (\cdot) \in \mathfrak{D}(U)|_{[0,t]}\}$$

at time $t > 0$ from the identity matrix I is a closed convex subset of $\mathrm{M}_n(\mathbf{R})$.

Since system (4.3) is commutative, we have that the reachable set of this system at time t from the identity matrix is exactly the set

$$\mathfrak{A}(t;U) = \{e^{\int_0^t A(u(s))\,ds} | u(\cdot) \in \mathfrak{D}(U)|_{[0,t]}\}.$$

Now we use the linearity of the mapping $A : \mathbf{R}^m \to \mathrm{M}_n(\mathbf{R})$ and the A. A. Lyapunov theorem on the range of a vector measure (see, e.g., [8]) in order to show that, in fact,

$$\mathfrak{A}(t;U) = \{e^{tA} | A \in \operatorname{conv} A(U)\}, \tag{4.4}$$

and, moreover, that the set $\operatorname{cl}\operatorname{conv} A(u) = \mathcal{A}$ consists of pairwise-commuting matrices.

First, we see that for any $u(\cdot) \in \mathfrak{D}|_{[0,t}$,

$$\int_0^t A(u(s))\,ds = A\int_0^t u(s)\,ds,$$

since the mapping $A : \mathbf{R}^m \to \mathrm{M}_n(\mathbf{R})$ is linear. Therefore,

$$\{\int_0^t A(u(s))\,ds;\ u \in \mathfrak{D}|_{[0,t]}\} = A\Big(\int_0^t U\,ds\Big) = tA(\operatorname{conv}\,(U)).$$

Since U is compact, $\operatorname{conv} A$ is compact, and since A is continuous, the set $A(\operatorname{conv} U)$ is also compact and convex. Let us show that

$$A(\operatorname{conv} U) = \operatorname{conv} A(U).$$

Indeed, $A(\operatorname{conv} U) \supset A(U)$ and is convex; therefore,

$$A(\operatorname{conv} U) = \operatorname{conv} A(\operatorname{convU}) \supset \operatorname{conv} A(U).$$

On the other hand, any $u \in \operatorname{conv} U$ can be represented as a convex combination

$$u = \sum_{i=1}^m \alpha_i u_i, \quad \alpha_i \geq 0, \quad \sum_{\alpha=1}^m \alpha_i = 1, \quad u_i \in U,$$

by the Caratheodory theorem (see, e.g., [6]).

Consequently, by the linearity of A,

$$A(u) = A(\sum_{i=1}^m \alpha_i u_i) = \sum_{i=1}^m \alpha_i A(u_i) \in \operatorname{conv} A(U),$$

i.e.,

$$A(\operatorname{conv} U) \subset \operatorname{conv} A(U),$$

and, thus,

$$A(\operatorname{conv} U) = \operatorname{conv} A(U).$$

Let us show that the family $\operatorname{conv} A(U)$ is commutative whenever the same property has the family $A(U)$.

Indeed, using the Caratheodory theorem again, we see that for any

$$A = \sum_{i=1}^{m^2+1} \alpha_i A(u_i) \in \operatorname{conv} A(U) \quad \text{and} \quad B = \sum_{i=1}^{m^2+1} \beta_i A(v_i) \in \operatorname{conv} A(U)$$

where

$$u_i, v_i \in U, \quad \sum_{i=1}^{m^2+1} \alpha_i = \sum_{i=1}^{m^2+1} \beta_i = 1, \quad \alpha_i \geq 0, \beta_i \geq 0,$$

we have

$$\Big[\sum_{i=1}^{m^2+1} \alpha_i A(u_i), \sum_{i=1}^{m^2+1} \beta_i A(v_i)\Big] = \sum_{i=1}^{m^2+1} \sum_{j=1}^{m^2+1} \alpha_i \beta_j [A(u-_i), B(v_j)] = 0.$$

This proves formula (4.4).

Finally, using Theorem 4.1, we conclude that the following assertion holds.

Theorem 4.2. *Under the above assumptions, the reachable set $\mathfrak{A}(t; U)$ of system* (4.3) *at any time $t > 0$ from the identity matrix is convex if for any $A_0 \in \operatorname{conv} A(U)$,*

$$(A - A_0)^k \in \tilde{T}\operatorname{conv} A(U)(A_0)$$

for all $A \in \operatorname{conv} A(U)$ and $k = 1, 2, \ldots$.

REFERENCES

1. N. I. Akhiezer, *Calculus of Variations* [in Russian], Vysshaya Shkola, Khar'kov (1981).

2. A. V. Arutyunov, "Convex properties of the Legendre transform," *Mat. Zametki*, **28**, No. 2, 255–263.

3. V. G. Boltyanskii, *Optimal Control of Discrete Systems* [in Russian], Nauka, Moscow (1973).

4. Yu. D. Burago and V. A. Zalgaller, "Sufficient conditions for convexity," *Zap. Nauchn. Semin. Leningr. Otd. Mat. Inst. Akad. Nauk SSSR*, **45**, 3–52 (1974).

5. F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York (1973).

6. R. V. Gamkrelidze, *Foundations of Optimal Control Theory* [in Russian], Tbilisi Univ., Tbilisi (1975).

7. R. V. Gamkrelidze, A. A. Agrachev and S. A. Vakhrameev, "Ordinary differential equations on vector bundles and the chronological calculus," in: *Progress in Science and Technology, Series on Contemporary Problems in Mathematics, Latest Achievements* [in Russian], Vol. 35, All-Union Institute for Scientific and Technical Information (VINITI), Akad. Nauk SSSR, Moscow (1989), pp. 3–107.

8. A. D. Ioffe and V. M. Tikhomirov, *Theory of Extremal Problems*[in Russian], Nauka, Moscow (1974).

9. K. Leichtweiss, *Convex Sets* [Russian translation], Nauka, Moscow (1985).

10. Yung-Chen Lu, *Singularity Theory and an Introduction to Catastrophe Theory*, Springer, New York (1976).

11. T. Motzkin,"Sur quelque proprietes characteristique des ensembles bornes non convexes," *Rend. Acad. Lincei*, **21**, 773–779 (1935).

12. L. S. Pontryagin, *Foundations of Combinatorial Topology* [in Russian], Nauka, Moscow (1986).

13. L. S. Pontryagin, *Selected Mathematical Works*, Vol. 2 [in Russian], Nauka, Moscow (1988).

14. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *Mathematical Theory of Optimal Processes* [in Russian], Nauka, Moscow (1976).

15. S. A. Vakhrameev, "Smooth control systems of constant rank and linearizable systems," in: *Progress in Science and Technology, Series on Contemporary Problems in Mathematics, Latest Achievements* [in Russian], Vol. 35, All-Union Institute for Scientific and Technical Information (VINITI), Akad. Nauk SSSR, Moscow (1989), pp. 135–177.

16. S. A. Vakhrameev, "Morse theory and the Lyusternik Shnirel'man theory in geometrical control theory," in: *Progress in Science and Technology, Series on Contemporary Problems in Mathematics, Latest Achievements* [in Russian], Vol. 39, All-Union Institute for Scientific and Technical Information (VINITI), Akad. Nauk SSSR, Moscow (1991), pp. 41–117.

17. S. A. Vakrameev, "On transversal convexity of reachable sets of a certain class of smooth control systems of a constant rank," *Dokl. Ross. Akad. Nauk*, **338**, No. 1, 7-9 (1994).

18. S. A. Vakhrameev, "A theorem on the finiteness of the number of switchings for smooth control systems," *Usp. Mat. Nauk*, **49**, No. 4, 197–198 (1994).

19. S. A. Vakhrameev, "Geometrical and topological methods in optimal control theory," *J. Math. Sci.*, **76**, No. 5, 2555–2719 (1995)

20. S. A. Vakhrameev, "An existence theorem for the nonlinear time-optimal control problem in the class of bang-bang controls with a finite number of switchings," *Usp. Mat. Nauk,* **51**, No. 2, 151–152 (1996).

21. S. A. Vakhrameev, "A bang-bang theorem with a finite number of switchings for smooth control systems," *J. Math. Sci.,* **85**, No. 3, 2002–2016 (1997)

22. S. A. Vakhrameev, "Bang-bang theorems and related topics," *Tr. Mat. Inst. Ross. Akad. Nauk*, **220**, 49–112 (1998).

23. S. A. Vakhrameev, "An existence theorem for the nonlinear time-optimal control problems," in: *International Conference Dedicated to the 90th Anniversary of the Birth of L.S. Pontryagin, Abstract of Reports, Optimal Control and Supplement*, MGU, Moscow (1998), pp. 198–220.

24. S. A. Vakhrameev, "An existence theorem for the nonlinear time-optimal control problem," *Differents. Uravn.* (in press).

25. S. A. Vakhrameev, "Morse lemmas for smooth functions on manifolds with corners," *J. Math. Sci.* (in press).