

A SECOND-ORDER NECESSARY CONDITION FOR OPTIMALITY IN THE GENERAL NONLINEAR CASE

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1977 Math. USSR Sb. 31 493

(<http://iopscience.iop.org/0025-5734/31/4/A05>)

[The Table of Contents](#) and [more related content](#) is available

Download details:

IP Address: 147.122.5.50

The article was downloaded on 25/06/2009 at 11:04

Please note that [terms and conditions apply](#).

A SECOND-ORDER NECESSARY CONDITION FOR OPTIMALITY IN THE GENERAL NONLINEAR CASE

UDC 519.3

A. A. AGRAČEV

ABSTRACT. In this paper a second-order optimality condition is obtained for the general nonlinear problem with arbitrary boundary conditions and an integral-type minimized functional. Moreover, a second-order sufficient condition for controllability is derived.

Bibliography: 3 titles.

This paper is a direct continuation of [1]. Here, the necessary condition for optimality contained in [1] is extended to the general nonlinear problem with arbitrary boundary conditions and minimized functional of integral type. Moreover, a second-order sufficient condition for controllability is derived as a corollary.

The work consists of four sections. The problem is formulated and necessary notions introduced in §1. In §2 the basic result of the work, i.e. a second-order necessary condition for optimality (Theorem 1), is formulated, and simple consequences are noted. A sufficient condition for controllability, Proposition 2, is formulated in §3. Finally, the outline of the proof of Theorem 1 is presented in §4. The proof of this theorem, except for the changes related to boundary conditions and the functional being more general than that of [1], consists in essence of reducing the problem with the aid of sliding regimes to the case of linear control that was considered in [1].

§§1–3 can be read independently of [1]. On the contrary, §4 depends on [1] in an essential way.

I would like to express my deep gratitude to R. V. Gamkrelidze, under whose guidance this work was done.

§1. Basic definitions

Let there be given a $(1 + m)$ -dimensional, twice continuously differentiable function of two n -dimensional arguments $x_0, x_1 \in \mathbf{R}^n$,

$$Q(x_0, x_1) = \begin{pmatrix} q_0(x_0, x_1) \\ q_1(x_0, x_1) \\ \dots \\ q_m(x_0, x_1) \end{pmatrix} = \begin{pmatrix} q_0(x_0, x_1) \\ q(x_0, x_1) \end{pmatrix} \in \mathbf{R}^{1+m},$$

and the n -dimensional controlled equation

$$\dot{x} = f(x, u), \quad x \in \mathbf{R}^n, \quad u \in U \subset \mathbf{R}^r. \quad (1.1)$$

We assume that the function $f(x, u)$ is infinitely differentiable, and that the set U is bounded and has the form

$$U = V \times W = \left\{ u = \begin{pmatrix} v \\ w \end{pmatrix} : v \in V \subset \mathbb{R}^s, \quad w \in W \subset \mathbb{R}^{r-s} \right\},$$

where W is arbitrary and V is an open subset of \mathbb{R}^s , $0 \leq s < r$.

This decomposition of the control parameter u is convenient insofar as we can derive additional information, corresponding to the "classical" case, when a part of the coordinates of the vector u changes freely.

The problem consists of minimizing the function $q_0(x(0), x(t_1))$ over the set of pairs $(t_1, x(t_1))$, where $t_1 \in \mathbb{R}^1$ and $x(t)$ is a solution of (1.1) (corresponding to a certain admissible control), where the following "boundary conditions" are satisfied:

$$q(x(0), x(t_1)) = 0. \quad (1.2)$$

We fix a (not necessarily optimal) solution of (1.1) that satisfies (1.2),

$$\tilde{x}(t), \quad \tilde{u}(t) = \begin{pmatrix} \tilde{v}(t) \\ \tilde{w}(t) \end{pmatrix}, \quad 0 \leq t \leq a, \quad (1.3)$$

Throughout the presentation, we shall use the following notation:

$$Q_0 = \frac{\partial Q}{\partial x_0}(\tilde{x}(0), \tilde{x}(a)), \quad Q_1 = \frac{\partial Q}{\partial x_1}(\tilde{x}(0), \tilde{x}(a)), \quad y = \begin{pmatrix} t \\ x \end{pmatrix}, \quad \tilde{y}(t) = \begin{pmatrix} t \\ \tilde{x}(t) \end{pmatrix},$$

$$f(y) = f(t, x) = f(x, \tilde{u}(t)), \quad \tilde{f}(t) = f(\tilde{x}(t), \tilde{u}(t)), \quad f_x(t) = \frac{\partial f}{\partial x}(\tilde{x}(t), \tilde{u}(t)),$$

$$f_v(y) = f_v(t, x) = \frac{\partial f}{\partial v}(\tilde{x}, \tilde{v}(t), \tilde{w}(t)), \quad f_{vv}(y) = f_{vv}(t, x) = \frac{\partial^2 f}{\partial v^2}(x, \tilde{v}(t), \tilde{w}(t)).$$

The matrices Q_0 and Q_1 have the dimension $(1 + m) \times n$; adjoining the matrix Q_1 to the matrix $-Q_0$, we obtain the $(1 + m) \times 2n$ matrix $(-Q_0, Q_1)$, with the aid of which the transversality condition is formulated (see (1.7)).

The dimension of the matrix $f_v(t, x)$ is $n \times s$, and we shall consider this matrix as an n -dimensional linear form of the s -dimensional argument δv . Correspondingly, $f_{vv}(t, x)$ is an n -dimensional bilinear form of the s -dimensional arguments δv_1 and δv_2 , and we shall write

$$f_v(t, x) \delta v, \quad f_{vv}(t, x) [\delta v_1, \delta v_2].$$

We denote by Σ the interior of the set of infinite differentiability points of the control $\tilde{u}(t)$.

The convex hull of the set $f(x, \bar{U})$ in \mathbb{R}^n will be denoted by $\text{conv } f(x, \bar{U})$ (\bar{U} is the closure of the set U in \mathbb{R}^n). The faces of the set $\text{conv } f(x, \bar{U})$ are the equivalence classes into which this set is subdivided if two points are said to be equivalent whenever they lie in an open interval contained in $\text{conv } f(x, \bar{U})$.

Let there be given a point $\hat{y} = \begin{pmatrix} \hat{t} \\ \hat{x} \end{pmatrix}$ with $\hat{t} \in \Sigma$. We denote by $\Phi_{\hat{y}}$ the set of infinitely differentiable functions $\varphi: \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ such that, for all $y = \begin{pmatrix} t \\ x \end{pmatrix}$ of a neighborhood $O_{\hat{y}} \subset \mathbb{R}^{n+1}$ of the point \hat{y} (the neighborhood depending on the function), the vector $\varphi(y) = \varphi(t, x)$ lies in the set $\text{conv } f(x, \bar{U})$, and on the same face of this set as the vector $f(t, x)$. Finally, we denote by $\Pi_{\hat{y}}$ the linear hull of the set $\{\varphi(\hat{y}) - f(\hat{y}) : \varphi \in \Phi_{\hat{y}}\}$. Then $\Pi_{\hat{y}}$ is a subspace of \mathbb{R}^n .

Any family of functions $\varphi_1(y), \dots, \varphi_k(y)$ belonging to $\Phi_{\hat{y}}$ and such that the vectors

$\varphi_i(\hat{y}) - f(\hat{y}), i = 1, \dots, k$, constitute a basis of the space $\Pi_{\hat{y}}$; will be called a *parametrization at the point \hat{y}* (defined by the control $\tilde{u}(t)$).

A parametrization at the point \hat{y} will also be a parametrization at all points close to \hat{y} if $\dim \Pi_y = \dim \Pi_{\hat{y}}$ for any point y sufficiently close to \hat{y} . These points will be called *regular points* (for $\tilde{u}(t)$).

The set of regular points is open and dense everywhere in $\Sigma \times \mathbb{R}^n$, because, if $\hat{y} \in \Sigma \times \mathbb{R}^n$, then, as it follows easily from the definition of the space Π_y , we have $\dim \Pi_y \geq \dim \Pi_{\hat{y}}$ for all y sufficiently close to \hat{y} . Therefore the function $\dim \Pi_y, y \in \Sigma \times \mathbb{R}^n$, assuming a finite number of values, is lower semicontinuous, which implies the assertion.

We introduce sliding regimes. Let

$$\Lambda = \left\{ \lambda = (\lambda_0, \lambda_1, \dots, \lambda_n) : \lambda_i \geq 0, i = 0, 1, \dots, n, \sum_{i=0}^n \lambda_i = 1 \right\}$$

be an n -dimensional simplex, and let

$$M = \Lambda \times \bar{U}^{n+1}, \quad F(x, \mu) = \sum_{i=0}^n \lambda_i f(x, u_i), \quad \text{where } \mu = (\lambda, u_0, u_1, \dots, u_n) \in M.$$

Clearly, $F(x, M) = \text{conv } f(x, \bar{U})$. In the sequel, we shall make use of the controlled equation

$$\dot{x} = F(x, \mu), \quad \mu \in M. \tag{1.4}$$

Let us indicate a simple method of constructing a parametrization at a given point with the aid of the function $F(x, \mu)$. Let A be a convex set containing the origin. The linear hull of the face of A containing the origin coincides with the linear hull of the set $A \cap (-A)$. We shall utilize this fact in order to find the parametrization at the point $\hat{y} = \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix}$.

Consider the following equation in unknown functions $\mu(t, x)$ and $\mu'(t, x)$:

$$F(x, \mu) + F(x, \mu') - 2f(t, x) = 0, \quad \mu(t, x), \mu'(t, x) \in M \quad \forall t, x. \tag{1.5}$$

Let $(\mu_1, \mu'_1), \dots, (\mu_k, \mu'_k)$ be the maximal family of solutions of (1.5) defined in a neighborhood of the point $\begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix}$ that satisfies the following conditions:

- a) The functions $F(x, \mu_i(t, x)), i = 1, \dots, k$, are infinitely differentiable.
- b) The vectors $F(\hat{x}, \mu_i(\hat{t}, \hat{x})) - f(\hat{t}, \hat{x}), i = 1, \dots, k$, are linearly independent.

Then the functions $F(x, \mu_i(t, x)), i = 1, \dots, k$, give a parametrization at the point $\begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix}$.

Assume now that the solution (1.3) is an extremal, i.e. that it satisfies the maximum principle: there exists a solution $\psi(t), 0 \leq t \leq a$, of the equation $\dot{\psi} = -\psi f_x(t)$ such that, for almost all t , the maximum condition

$$\psi(t) f(t) = \sup_{u \in U} \psi(t) f(\tilde{x}(t), u) = 0 \tag{1.6}$$

holds, and the transversality condition

$$(\psi(0), \psi(a)) = \xi (-Q_0, Q_1), \tag{1.7}$$

where ξ is a nonzero $(1 + m)$ -dimensional row, is satisfied.

In order to guarantee that the maximum principle is nontrivial, we assume that the rank of the matrix $(-Q_0, Q_1)$ is $1 + m$. In this case, the n -dimensional row $\psi(t) \neq 0$.

We denote the set of all $\psi(t)$ of this form by Ψ ,

$$\Psi = \{\psi(t) : \dot{\psi} = -\psi f_x(t), \quad \psi(t) f(t) \stackrel{\text{a.e.}}{=} \sup_{u \in U} \psi(t) f(\tilde{x}(t), u) = 0, \quad (1.8)$$

$$(\psi(0), \psi(a)) = \xi(-Q_0, Q_1), \xi \neq 0\}.$$

For any function $\psi(t) \in \Psi$, we form the $s \times s$ symmetric matrix

$$\psi(t) \ddot{f}_{vv}(t, \tilde{x}(t)) = \psi(t) f_{vv}(t),$$

which we shall consider as a linear mapping of the s -dimensional space \mathbf{R}^s into itself. We denote the intersection of the kernels of all such mappings by P_t ,

$$P_t = \{\delta v \in \mathbf{R}^s : \psi(t) \ddot{f}_{vv}(t) \delta v = 0 \quad \forall \psi(t) \in \Psi\}.$$

It follows from the maximum condition (1.6) that the quadratic forms $\psi(t) \ddot{f}_{vv}(t) [\delta v, \delta v]$, $\psi(t) \in \Psi$, are nonpositive. Therefore the subspace $P_t \subset \mathbf{R}^s$ can be also defined by the condition

$$P_t = \{\delta v \in \mathbf{R}^s : \psi(t) \ddot{f}_{vv}(t) [\delta v, \delta v] = 0 \quad \forall \psi(t) \in \Psi\}.$$

We denote by Σ^0 the set of those $\hat{t} \in \Sigma$ in a neighborhood of which $\dim P_t$ does not change. The same considerations that proved that the set of regular points is open and everywhere dense in $\Sigma \times \mathbf{R}^n$ imply that the set Σ^0 is open and everywhere dense in Σ .

Finally, let $p(t)$ be the orthogonal projection of \mathbf{R}^s onto P_t . Obviously, $p(t)$ on the set Σ^0 is an infinitely differentiable function of t .

We are now able to give the basic definition of this section.

Let $\sigma \in \Sigma^0$, and let $\varphi_1(t, x), \dots, \varphi_{k_\sigma}(t, x)$ be an arbitrary parametrization of the extremal (1.3) at the point $(\tilde{x}(\sigma)) = \tilde{y}(\sigma)$. We introduce the $n \times (k_\sigma + s)$ matrix

$$\Delta^\sigma f(t, x) = (\varphi_1(t, x), \dots, \varphi_{k_\sigma}(t, x), f_v(t, x) p(t)).$$

The first k_σ columns of this matrix consist of the vectors $\varphi_1, \dots, \varphi_{k_\sigma}$, and the remaining s columns are the columns of the $n \times s$ matrix $f_v \cdot p$.

We now form the following function that depends linearly on $v \in \mathbf{R}^{k_\sigma+s}$:

$$\pi^{\sigma f}(t, x; v) = f(t, x) + \Delta^\sigma f(t, x) v.$$

The controlled equation

$$\dot{x} = \pi^{\sigma f}(t, x; v) = f(t, x) + \Delta^\sigma f(t, x) v, \quad v \in \mathbf{R}^{k_\sigma+s}, \quad (1.9)$$

in which the parameter v enters linearly, will be called the *linearization of the controlled equation (1.1) at the point $\tilde{y}(\sigma)$ corresponding to the chosen parametrization and the trajectory (1.3) under consideration.*

§2. Formulation of a necessary condition for optimality

In order to formulate Theorem 1, we must introduce the Legendre operators $\mathfrak{L}_i, i \geq 0$. The definition of these will be preceded by a brief listing of the notation that we use pertaining to vector fields.

Let an infinitely differentiable function $g: \mathbf{R}^k \rightarrow \mathbf{R}^k$ be considered as a vector field, i.e. as an operator which acts in the set of arbitrary dimensional, infinitely differentiable functions $\varphi(z), z \in \mathbf{R}^k$, according to the formula

$$(g \circ \varphi)(z) = \frac{d\varphi}{dz}(z) g(z).$$

The successive application of the fields g_2 and g_1 yields an operator on the set of all $\varphi(z)$, which is denoted by $g_1 \circ g_2$ (there will be no misunderstanding because of the dual use of the symbol \circ). In general, this operator is not a vector field. However, the Lie

brackets of the fields g_1 and g_2 defined by the expression

$$[g_1, g_2] = g_1 \circ g_2 - g_2 \circ g_1$$

are always a field, i.e. a function, where

$$[g_1, g_2](z) = \frac{dg_2(z)}{dz} g_1(z) - \frac{dg_1(z)}{dz} g_2(z).$$

We denote by $\text{ad } g_1$ the operator acting in the set of vector fields by the formula

$$(\text{ad } g_1) g_2 = [g_1, g_2],$$

and the i th power of the operator $\text{ad } g_1$ will be denoted by $\text{ad}^i g_1$, e.g.

$$(\text{ad}^2 g_1) g_2 = [g_1, [g_1, g_2]].$$

If A is an operator, then $e^{\varepsilon A}$ denotes the formal series of operators of the variable ε ,

$$e^{\varepsilon A} = \sum_{i=0}^{\infty} \varepsilon^i \frac{A^i}{i!}.$$

For arbitrary fields g_1 and g_2 and an infinitely differentiable scalar-valued function $\varphi(z)$, we have

$$(\text{ad } g_1)(\varphi g_2) = (g_1 \circ \varphi) g_2 + \varphi (\text{ad } g_1) g_2. \tag{2.1}$$

Hence by induction we obtain

$$e^{\varepsilon \text{ad } g_1}(\varphi g_2) = (e^{\varepsilon g_1} \circ \varphi) e^{\varepsilon \text{ad } g_1} g_2. \tag{2.2}$$

We now set $z = y = \binom{t}{x}$ and $k = 1 + n$. Let $g(t, x, \nu)$, $\nu \in \mathbf{R}^d$, be an n -dimensional function of the form

$$g(t, x, \nu) = g_0(t, x) + G(t, x) \nu, \tag{2.3}$$

where $g_0(t, x)$ is an n -dimensional function infinitely differentiable with respect to (t, x) , and $G(t, x)$ is an infinitely differentiable $n \times d$ matrix.

We introduce the $(n + 1)$ -dimensional field $\hat{g}_0(y) = \hat{g}_0(t, x)$, and the family of $(n + 1)$ -dimensional fields $\hat{G}(y)\nu = \hat{G}(t, x)\nu$ that depends on ν according to the formulas

$$\hat{g}_0(y) = \hat{g}_0(t, x) = \begin{pmatrix} 1 \\ g_0(t, x) \end{pmatrix}, \quad \hat{G}(y)\nu = \hat{G}(t, x)\nu = \begin{pmatrix} 0 \\ G(t, x) \end{pmatrix} \nu.$$

We define the result of the action of the operator \mathfrak{L}_i , $i = 0, 1, 2, \dots$, on an arbitrary function $g(t, x, \nu)$ of the form (2.3) as the n -dimensional bilinear form $\mathfrak{L}_i g(t, x)[\nu_1, \nu_2]$ of d -dimensional variables ν_1 and ν_2 which is infinitely differentiable with respect to t and x , and is evaluated by the formula

$$\begin{pmatrix} 0 \\ \mathfrak{L}_i g(t, x)[\nu_1, \nu_2] \end{pmatrix} = [(\text{ad}^i \hat{g}_0) \hat{G}\nu_1, \hat{G}\nu_2] \begin{pmatrix} t \\ x \end{pmatrix}. \tag{2.4}$$

This is a correct definition, since it follows easily from the definitions of the fields \hat{g}_0 and \hat{G} , that the vector field in the right-hand side of (2.4) has first coordinate zero.

The operators \mathfrak{L}_i have "local character"; namely, the coefficients of $\mathfrak{L}_i(t, x)[\nu_1, \nu_2]$ at a point t, x are expressed by partial derivatives of $g_0(t, x)$ and $G(t, x)$ at (t, x) . It is convenient to give \mathfrak{L}_i in the form of the "generating series"

$$\sum_{i=0}^{\infty} \frac{\varepsilon^i}{i!} \begin{pmatrix} 0 \\ \mathfrak{L}_i g(t, x)[\nu_1, \nu_2] \end{pmatrix} = [e^{\varepsilon \text{ad } \hat{g}_0} \hat{G}\nu_1, \hat{G}\nu_2] \begin{pmatrix} t \\ x \end{pmatrix};$$

for more detail see [1].

If ξ is an n -dimensional row, then the products $\xi \mathcal{L}_i g(t, x)[v_1, v_2]$, $i = 0, 1, 2, \dots$, are scalar-valued bilinear forms in v_1 and v_2 .

THEOREM 1. *Assume that $q_0(\bar{x}(0), \bar{x}(a))$ is the minimum of the functional $q_0(x(0), x(t_1))$ on solutions of (1.1) under the conditions (1.2). Then there exists a function $\tilde{\psi}(t) \in \Psi$ (see (1.8)) such that, for an arbitrary point $\sigma \in \Sigma^0$ and an arbitrary linearization (1.9) of (1.1), the following assertions hold at the point $(\tilde{x}(\sigma))$ (we assume that $\dim \Pi_\sigma = k_\sigma$):*

1) *If $l \geq 1$ is an odd integer, if v is a fixed vector of $\mathbb{R}^{k_\sigma+s}$, and if the following equalities hold in a neighborhood O_σ of the point σ for all odd $i \leq l - 2$:*

$$\psi(t) \mathcal{L}_i \pi^{\sigma f}(t, \tilde{x}(t)) [v, v] = 0 \quad \forall \psi \in \Psi, \quad \forall t \in O_\sigma,$$

then

$$(-1)^{\frac{l+3}{2}} \tilde{\psi}(\sigma) \mathcal{L}_i \pi^{\sigma f}(\sigma, \tilde{x}(\sigma)) [v, v] \leq 0.$$

2) *If $l \geq 0$ is an even integer, if N is a subspace of $\mathbb{R}^{k_\sigma+s}$ such that, in a neighborhood O_σ of σ and for all $i \leq l - 1$, the bilinear forms*

$$\psi(t) \mathcal{L}_i \pi^{\sigma f}(t, \tilde{x}(t)) [v_1, v_2] \quad \forall \psi \in \Psi, \quad \forall t \in O_\sigma,$$

in $v_1, v_2 \in N$ vanish identically, then so does the bilinear form

$$\tilde{\psi}(\sigma) \mathcal{L}_i \pi^{\sigma f}(\sigma, \tilde{x}(\sigma)) [v_1, v_2].$$

REMARK. If the control parameter in the linearized equation (1.9) is considered only on the subspace of those vectors v of $\mathbb{R}^{k_\sigma+1}$ whose first k_σ coordinates are zero, and if the corresponding s -dimensional vectors are denoted by \bar{v} , then from the definition of the Legendre operators (2.4) we obtain

$$\begin{pmatrix} 0 \\ \mathcal{L}_i \pi^{\sigma f}(t, \tilde{x}(t)) [v_1, v_2] \end{pmatrix} = [(\text{ad}^i \hat{f}) \hat{f}_{v_1} \bar{v}_1, \hat{f}_{v_2} \bar{v}_2] \begin{pmatrix} t \\ \tilde{x}(t) \end{pmatrix},$$

where

$$\hat{f}(t, x) = \begin{pmatrix} 1 \\ f(t, x) \end{pmatrix}, \quad \hat{f}_\sigma(t, x) = \begin{pmatrix} 0 \\ f_\sigma(t, x) \end{pmatrix}.$$

It will be seen from the proof of the theorem that the corresponding necessary condition is a necessary condition for a “weak” minimum (local in the control, the “classical” case), whereas the theorem in the stated form yields only a necessary condition for a “strong” minimum, where one can compare with $\tilde{u}(t)$ arbitrary controls $u(t)$ that are close to $\tilde{u}(t)$ in a metric weaker than the uniform metric.

If the controlled equation (1.1) is linear in u , and if U is a convex polyhedron, then we obtain from Theorem 1 for the time-optimal problem with fixed endpoints, as a simple corollary, a result whose form is somewhat more general than that of [1]. However, as is easy to see, the results of [1] contain this consequence.

To conclude this section, we shall discuss the problem of the effect of nonuniqueness of the linearized equation on the content of Theorem 1.

The assertions of the theorem for various linearizations (at the same point) are, generally speaking, not equivalent; and there is no linearization for which the assertion of the theorem would be the strongest. Nevertheless, if $\tilde{y}(\sigma) = (\tilde{x}(\sigma))$ is a regular point ($\sigma \in \Sigma^0$), then the assertions of the theorem turn out to be equivalent for all possible

linearizations at a given point $\bar{y}(\sigma)$ in an important particular case. Namely, the following proposition holds.

PROPOSITION 1. Let $\sigma \in \Sigma$, let $\bar{y}(\sigma)$ be a regular point, and let

$$\pi f(t, x, v) = \bar{f}(t, x) + \Delta^\sigma f(t, x)v, \quad \pi^\sigma f'(t, x, v) = f'(t, x) + \Delta^\sigma f'(t, x)v$$

be two different linearizations at the point $\bar{y}(\sigma)$. There exists a constant (not dependent on t or x) nonsingular $(k_\sigma + s) \times (k_\sigma + s)$ matrix B_0 such that, for any $\psi \in \Psi$ and $l \geq 0$, the equalities

$$\psi(\sigma) \mathcal{L}_t \pi^\sigma f'(\sigma, \bar{x}(\sigma)) [v_1, v_2] = 0 \quad \forall v_1, v_2 \in \mathbb{R}^{k_{\sigma+s}}, \quad \forall i \leq l-1, \tag{2.5}$$

imply

$$\begin{aligned} \psi(\sigma) \mathcal{L}_t \pi^\sigma f'(\sigma, \bar{x}(\sigma)) [v_1, v_2] &= \psi(\sigma) \mathcal{L}_t \pi^\sigma f(\sigma, \bar{x}(\sigma)) [B_0 v_1, B_0 v_2] \\ &\quad \forall v_1, v_2 \in \mathbb{R}^{k_{\sigma+s}}, \quad \forall j \leq l. \end{aligned} \tag{2.6}$$

PROOF. Let $\psi \in \Psi$, and let $l \geq 0$ be such that the equalities (2.5) hold. We introduce the $(1+n) \times (k_\sigma + s)$ matrices

$$G(y) = G(t, x) = \begin{pmatrix} 0 \\ \Delta^\sigma \bar{f}(t, x) \end{pmatrix}, \quad G'(y) = G'(t, x) = \begin{pmatrix} 0 \\ \Delta^\sigma f'(t, x) \end{pmatrix}$$

and the $(1+n)$ th row $\hat{\psi}(t) = (0, \psi(t))$. As before,

$$\hat{f}(y) = \hat{f}(t, x) = \begin{pmatrix} 1 \\ f(t, x) \end{pmatrix}$$

is a $(1+n)$ -dimensional vector. With this notation, we obtain from (2.4) the equality

$$\psi(\sigma) \mathcal{L}_t \pi^\sigma f(\sigma, \bar{x}(\sigma)) [v_1, v_2] = \hat{\psi}(\sigma) [(\text{ad}^l \hat{f}) G v_1, G v_2] (\bar{y}(\sigma))$$

and a similar equality for $\psi(\sigma) \mathcal{L}_t \pi^\sigma f'(\sigma, \bar{x}(\sigma)) [v_1, v_2]$.

Since $\bar{y}(\sigma)$ is a regular point, the linear hull of the columns of the matrix $G(y)$, for all y sufficiently close to $\bar{y}(\sigma)$, coincides with the linear hull of the columns of the matrix $g'(y)$ (and with the subspace $(\frac{0}{\Pi})$ of \mathbb{R}^{n+1} ; see §1). Therefore $G'(y) = G(y)B(y)$, where $B(y)$ is a $(k_\sigma + s) \times (k_\sigma + s)$ nonsingular matrix,⁽¹⁾ which is infinitely differentiable with respect to y .

From the maximum condition (1.6), which holds for the functions belonging to the set Ψ , one can easily obtain the identity

$$\hat{\psi}(t) G(\bar{y}(t)) \equiv 0 \tag{2.7}$$

for all t sufficiently close to σ .

Let $v \in \mathbb{R}^{k_\sigma+s}$. Since $\psi(t)$ satisfies the adjoint equation, we have, obviously,

$$\frac{d}{dt} (\hat{\psi}(t) G(\bar{y}(t)) v) = \hat{\psi}(t) ((\text{ad } \hat{f}) G v) (\bar{y}(t)).$$

⁽¹⁾ In fact, we can assume that $B(y)$ has the form

$$\left. \begin{pmatrix} \bar{B}(y) & 0 \\ & 1 & 0 \\ & & \ddots & \\ 0 & 0 & & 1 \end{pmatrix} \right\} s.$$

Therefore successive differentiation of (2.7) with respect to t yields

$$\hat{\psi}(\sigma)(e^{\varepsilon \text{ad} \hat{f}} G v)(\tilde{y}(\sigma)) = 0 \quad \forall v \in \mathbb{R}^{k_{\sigma+s}}. \tag{2.8}$$

We shall prove Proposition 1, transforming the expression

$$\begin{aligned} \hat{\psi}(\sigma)[e^{\varepsilon \text{ad} \hat{f}} G' v_1, G' v_2](\tilde{y}(\sigma)) &= \sum_{i=0}^{\infty} \psi(\sigma) \mathcal{L}_i \pi^{\sigma f'}(\sigma, \tilde{x}(\sigma)) [v_1, v_2] \frac{\varepsilon^i}{i!}, \\ v_1, v_2 &\in \mathbb{R}^{k_{\sigma+s}}, \end{aligned}$$

where, as a rule, the arguments σ and $\tilde{y}(\sigma)$ will be omitted. Besides (2.8), we shall subsequently employ (2.1) and (2.2).

Let $B_0 = B(\tilde{y}(\sigma))$. The expansion in a Taylor series of the functions $B(y)v_i$ ($i = 1, 2$) at the point $\tilde{y}(\sigma)$ can be expressed as $B(y)v_i \sim \sum_{j=0}^{\infty} \varphi^j(y)\eta_i^j$, where $\varphi^j(y)$ are scalar-valued polynomials of y , and η^j are constant (independent of y) $(1+n)$ -vectors, where $\varphi^0(y) \equiv 1$, $\eta_i^0 = B_0 v_i$ and $\varphi^j(\tilde{y}(\sigma)) = 0$ for $j \geq 1$. We have

$$\begin{aligned} \hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} G' v_1, G' v_2] &= \hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} G B v_1, G B v_2] = \sum_{i,j=0}^{\infty} \hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} \varphi^i G \eta_1^i, \varphi^j G \eta_2^j] \\ &= \sum_{i,j=0}^{\infty} \hat{\psi}[(e^{\varepsilon \hat{f}} \circ \varphi^i) e^{\varepsilon \text{ad} \hat{f}} G \eta_1^i, \varphi^j G \eta_2^j] = \sum_{i,j=0}^{\infty} \{(e^{\varepsilon \hat{f}} \circ \varphi^i) \varphi^j \hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} G \eta_1^i, G \eta_2^j] \\ &\quad + (e^{\varepsilon \hat{f}} \circ \varphi^i) ((e^{\varepsilon \text{ad} \hat{f}} G \eta_1^i) \circ \varphi^j) \hat{\psi} G \eta_2^j - \varphi^j ((G \eta_2^j) \circ e^{\varepsilon \hat{f}} \circ \varphi^i) \hat{\psi} e^{\varepsilon \text{ad} \hat{f}} G \eta_1^i\} \\ &= \sum_{i=0}^{\infty} (e^{\varepsilon \hat{f}} \circ \varphi^i) \hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} G \eta_1^i, G B_0 v_2]. \end{aligned}$$

Thus

$$\hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} G' v_1, G' v_2] = \sum_{i=0}^{\infty} (e^{\varepsilon \hat{f}} \circ \varphi^i) \hat{\psi}[e^{\varepsilon \text{ad} \hat{f}} G \eta_1^i, G B_0 v_2].$$

The last equality and (2.5) at once imply the desired equality (2.6).

§3. A sufficient condition for controllability

The method of proving Theorem 1 (described in §4) allows one to obtain also a sufficient condition (of second order) for local controllability, which will be formulated here. In so doing, we restrict ourselves to the consideration of controllability in a neighborhood of a fixed point for an equation with a one-dimensional control, in which case the condition has a particularly simple form.

Thus, we are given the equation

$$x = f(x, u), \quad x \in \mathbb{R}^n, u \in U \subset \mathbb{R}, \quad 0 \in \text{int } U, f(0, 0) = 0. \tag{3.1}$$

Let $a > 0$; we denote by X_a the set of attainability at the time a ,

$$X_a = \{x(a) \in \mathbb{R}^n: x(t) \quad (0 \leq t \leq a)$$

is an admissible trajectory of equation (3.1), $x(0) = 0\}$.

We also use the following notation: $f_x = f_x(0) = (\partial f / \partial x)(0, 0)$ is an $n \times n$ matrix, $f(x) = f(x, 0)$ and $f_u(x) = (\partial f / \partial u)(x, 0)$ are vector fields in \mathbb{R}^n .

Let $H^{(1)}$ be the linear hull of the vectors $f_u(0), f_x \cdot f_u(0), \dots, f_x^{n-1} \cdot f_u(0)$ (a subspace of

\mathbf{R}^n). A well-known condition (of first order) for controllability states that

If $H^{(1)} = \mathbf{R}^n$, then the origin is an interior point of the set X_a for any $a > 0$.

If $H^{(1)} \neq \mathbf{R}^n$, let k be the smallest integer κ such that at least one of the n vectors

$$f_x^i [(ad^{2\kappa-1}f) f_u, f_u](0), \quad i = 0, 1, \dots, n-1,$$

does not lie in the subspace $H^{(1)}$; it is assumed in this connection that

$$[(ad^{-1}f) f_u, f_u](0) = \frac{\partial^2 f}{\partial u^2}(0, 0).$$

If there are no such κ , there is also no second-order condition.

Clearly, the k thus defined is the least integer κ such that the curve

$$e^{-t f_x} [(ad^{2\kappa-1}f) f_u, f_u](0), \quad t \geq 0,$$

does not lie entirely in $H^{(1)}$.

Finally, we define $H_a^{(2)}$ as the cone spanned in \mathbf{R}^n by the convex hull

$$\text{conv} (H^{(1)} \cup \{(-1)^{k+1} e^{-t f_x} [(ad^{2k-1}f) f_u, f_u](0); 0 \leq t \leq a\}).$$

PROPOSITION 2. *Let $a > 0$. If $H_a^{(2)} = \mathbf{R}^n$, then the origin is an interior point of the set of attainability X_a .*

Let us clarify how this result can be obtained from the proof of Theorem 1. If we set $Q(x_0, x_1) = x_1 - e^{af_x} x_0$, $\tilde{x}(t) \equiv 0$, $\tilde{u}(t) \equiv 0$, then we arrive at the situation considered in §§1, 2 and 4 (only instead of (1.1) we consider (3.1)). In this case, the closure of the second-order cone $K_a^{(2)}$ (see p. 504) contains the set $e^{af_x} H_a^{(2)}$. Therefore, since the matrix e^{af_x} is nonsingular, the assertion being proved follows from the arguments of §4.

REMARK. Proposition 2 can be strengthened in the following way. Let $x \in \mathbf{R}^n$ and $a > 0$. If the vector x lies in the interior of the cone $e^{af_x} H_a^{(2)}$ then the vector ϵx is an interior point of the set X_a for all $\epsilon > 0$ sufficiently small.⁽²⁾

On the basis of this remark, we obtain the following corollary.

COROLLARY. *If the linear hull of $2n$ vectors*

$$f_x^i f_u(0), \quad 0 \leq i \leq n-1, \quad f_x^i [(ad^{2k-1}f) f_u, f_u](0), \quad 0 \leq i \leq n-1,$$

coincides with \mathbf{R}^n , then the set of attainability X_a has a nonempty interior for any $a > 0$.

§4. An outline of the proof of Theorem 1

In this section, we shall freely use the definitions and results of [1].

Again, as in §§1 and 2, we consider (1.1) and its solution (1.3). Let $\sigma \in \Sigma^0$, let $\varphi_1(t, x), \dots, \varphi_{k_\sigma}(t, x)$ be a parametrization at the point (\tilde{x}_σ^0) , and let

$$\Delta^\sigma f(t, x) = (\varphi_1(t, x), \dots, \varphi_{k_\sigma}(t, x), f_\sigma(t, x) p(t))$$

(see §1). Moreover, let $\nu_\epsilon(\cdot): \mathbf{R} \rightarrow \mathbf{R}^{k_\sigma+s}$ be a family of measurable functions depending on $\epsilon > 0$, where $\nu_\epsilon(t)$ tends to zero uniformly in $t \in \mathbf{R}$ as $\epsilon \rightarrow 0$. We denote the first k_σ coordinates of $\nu_\epsilon(t)$ by $\nu_\epsilon^1(t), \dots, \nu_\epsilon^{k_\sigma}(t)$, the vector composed of the remaining s coordinates by $\nu_\epsilon^p(t)$, and we set

⁽²⁾ In order to prove this assertion, one needs to change somewhat the arguments of §4. However, we shall not dwell on it in detail.

$$\|v_\varepsilon^\sigma\|_\infty = \operatorname{ess\,sup}_{t \in \mathbb{R}} |v_\varepsilon^\sigma(t)|.$$

If the functions $v_\varepsilon(t)$ are nonzero only in a sufficiently small neighborhood of σ , then for all small ε the solution of the equation

$$\dot{x} = \sum_{i=1}^{k_\sigma} v_\varepsilon^i(t) \varphi_i(t, x) + \frac{1}{2} f(t, x) + \frac{1}{2} f(x, \tilde{v}(t) + 2p(t) v_\varepsilon^\sigma(t), \tilde{w}(t)), \quad x(0) = \tilde{x}(0), \quad (4.1)$$

is determined on the entire interval $[0, a]$, and it is an admissible trajectory of equation (1.4) on this interval.

Indeed, since $v_\varepsilon(t) \rightarrow 0$ as $\varepsilon \rightarrow 0$, the solution of (1.4) tends to $\tilde{x}(t)$ uniformly in $t \in [0, a]$ as $\varepsilon \rightarrow 0$. Moreover, if the functions $v_\varepsilon^i(t)$ are small and distinct from zero only close to σ , then, for (t, x) sufficiently close to the curve $(t, \tilde{x}(t))$ with $0 < t < a$, we have

$$\left(2 \sum_{i=1}^{k_\sigma} v_\varepsilon^i(t) \varphi_i(t, x) + f(t, x) \right) \in F(x, M)$$

(see the definition of parametrization). Therefore the vector in the right-hand side of (4.1) also belongs to $F(x, M)$. Thus, the solution of (4.1) is an admissible trajectory of (1.4).

We expand the last term in the right-hand side of (4.1) in powers of $v_\varepsilon^\sigma(t)$:

$$\begin{aligned} \dot{x} = f(t, x) + \Delta^\sigma f(t, x) v_\varepsilon(t) + f_{\sigma\sigma}(t, \tilde{x}(t)) [p(t) v_\varepsilon^\sigma(t), p(t) v_\varepsilon^\sigma(t)] \\ + \|v_\varepsilon^\sigma\|_\infty^3 r(t, x; \varepsilon), \quad x(0) = \tilde{x}(0), \end{aligned} \quad (4.2)$$

where $r(t, x; \varepsilon)$ is uniformly bounded in ε .

Let $\alpha(\varepsilon)$ and $\beta(\varepsilon)$ be real-valued positive functions of the positive argument ε , with $\alpha(\varepsilon), \beta(\varepsilon) \searrow 0$ as $\varepsilon \rightarrow +0$. Further, let $\eta(\tau)$ ($\tau \in \mathbb{R}$) be a $(k_\sigma + s)$ -dimensional function that is measurable, bounded, and distinct from zero only for $\tau \in [-1, 1]$, and let $\eta^v(\tau)$ be the vector composed of the last s coordinates of $\eta(\tau)$.

We denote by $x(t; \varepsilon)$ the solution of (4.1) which we obtain on setting

$$v_\varepsilon(t) = \alpha(\varepsilon) \eta\left(\frac{t - \sigma'}{\beta(\varepsilon)}\right).$$

Here σ' is a point of Σ^0 . It follows from what we have said that, if σ' is sufficiently close to σ , then the curve $x(t; \varepsilon)$ with $0 < t < a$ is an admissible trajectory of (1.4) for all small ε .

We now consider the equation

$$\dot{x}_\sigma = f(t, x_\sigma) + \Delta^\sigma f(t, x_\sigma) v, \quad x_\sigma \in \mathbb{R}^n, \quad v \in \mathbb{R}^{(k_\sigma + s)}, \quad (4.3)$$

which is linear in the control. Obviously the curve $\tilde{x}_\sigma(t) = \tilde{x}(t)$ is the solution (4.3) corresponding to the control $v \equiv 0$.

The family of controls $v_\varepsilon(t) = \alpha(\varepsilon) \eta((t - \sigma')/\beta(\varepsilon))$ is a packet of perturbations of this control (see [1], (1.8)). Let $x_\sigma(t; \varepsilon)$ be the family of trajectories obtained as a result of applying this packet of perturbations to equation (4.3), i.e.

$$\dot{x}_\sigma(t; \varepsilon) = f(t, x_\sigma(t; \varepsilon)) + \Delta^\sigma f(t, x_\sigma(t; \varepsilon)) \alpha(\varepsilon) \eta\left(\frac{t - \sigma'}{\beta(\varepsilon)}\right), \quad x_\sigma(0; \varepsilon) = \tilde{x}(0).$$

We denote by $\Gamma(t)$ the fundamental matrix of the equation $\dot{\Gamma} = f_x(t)\Gamma$, $\Gamma(t, \tau) =$

$\Gamma(t)\Gamma^{-1}(\tau)$. Let

$$R(t; \varepsilon) = \int_0^t \Gamma(t, \tau) f_{vv}(\tau, \tilde{x}(\tau)) \left[p(\tau) \eta^v \left(\frac{\tau - \sigma'}{\beta(\varepsilon)} \right), p(\tau) \eta^v \left(\frac{\tau - \sigma'}{\beta(\varepsilon)} \right) \right] d\tau.$$

It follows from the definition of $p(\tau)$ that $\psi(a)R(a; \varepsilon) = 0$ for any $\psi \in \Psi$ (obviously, $\psi(\tau) = \psi(a)\Gamma(a, \tau)$). Finally, from (4.2) one can derive the equality

$$x(t; \varepsilon) = x_\sigma(t; \varepsilon) + \alpha^2(\varepsilon) R(t; \varepsilon) + O(\alpha^3(\varepsilon)), \tag{4.4}$$

where $O(\alpha^3(\varepsilon))/\alpha^3(\varepsilon)$ is uniformly bounded in ε . The formula (4.4) allows one to apply to (1.4) the method developed in [1] for equations linear in the control. In order to pass to (1.1), one must also use the ‘‘approximation lemma’’ of R. V. Gamkrelidze, from which it follows that any family of trajectories of (1.4) depending continuously on a parameter can be approximated to within any accuracy by a family of trajectories of (1.1) that depends continuously on a parameter (for the exact definitions and proofs, see [2] and [3]).

We shall now formulate an assertion (Theorem 1’), from which Theorem 1 can be derived in the same way as Theorem 2.2 was derived from Theorem 2.1 in [1]. We turn again to equation (4.3). All the definitions of [1] can be applied to this equation. For example, for this equation one can write the Legendre representation (see [1], (1.13)) of the endpoint of the second variation $\delta_2 x_\sigma(a; \varepsilon)$ of the trajectory $\tilde{x}_\sigma(t) = \tilde{x}(t)$ on the packet $\alpha(\varepsilon)\eta((t - \sigma')/\beta(\varepsilon))$.

We denote this Legendre representation as

$$\delta_2 x_\sigma(a; \varepsilon) \sim \alpha^2(\varepsilon) \beta^2(\varepsilon) \sum_{i=0}^{+\infty} \frac{\beta^i(\varepsilon)}{i!} L_i^\sigma(\sigma') [\eta(t)]. \tag{4.5}$$

Obviously, the Legendre representation (4.5) depends on the choice of a linearization at the point $(\tilde{x}_\sigma^\sigma)$ (the choice of a linearization (1.9) determines (4.3)), although this dependence is not noted explicitly.

Let $L_\infty^{(k_\sigma + s)}$ be the linear space of all bounded measurable $(k_\sigma + s)$ -dimensional functions on $[-1, 1]$, and let $\mathfrak{K}_\sigma^{(l)}$ be the family of functions $\eta(t)$ from $L_\infty^{(k_\sigma + s)}$ such that

$$\int_{-1}^1 t^i \eta(t) dt = 0, \quad i = 0, \dots, l.$$

THEOREM 1’. *Assume that $q_0(\tilde{x}(0), \tilde{x}(a))$ is the minimum of the functional $q_0(x(0), x(t_1))$ over the solutions of equation (1.1) subject to the conditions (1.2). Then there exists a function $\tilde{\psi} \in \Psi$ such that, for an arbitrary point $\sigma \in \Sigma^0$, for $\eta(t) \in \mathfrak{K}_\sigma^{(l)}$, and for an arbitrary linearization of (1.1) at the point $(\tilde{x}_\sigma^\sigma)$, the following assertion holds.*

If $l \geq 0$ is an integer, and if in a neighborhood O_σ of σ

$$\psi(a) L_i^\sigma(\sigma') [\eta(t)] = 0 \quad \forall \psi \in \Psi, \forall \sigma' \in O_\sigma, \forall i \leq l-1, \tag{4.6}$$

then $\tilde{\psi}(a) L^\sigma(\sigma) [\eta(t)] \leq 0$.

Let us describe the main steps of the proof of Theorem 1’. We return to (1.4) and introduce the sets

$$\delta_1 X(t) = \left\{ \int_0^t \Gamma(t, \tau) \varphi(\tau) d\tau : \varphi(\tau) \text{ is measurable,} \right. \\ \left. (\varphi(\tau) + f(\tau)) \in F(\bar{x}(\tau), M), \quad 0 \leq \tau \leq t \right\}.$$

Thus $\delta_1 X(t)$ is the set of the first variations of the trajectory $\bar{x}(\tau)$ at the instant of time t .

Subsequently, we shall assume that a is a Lebesgue point of the function $\bar{u}(t)$ (otherwise the constructions presented below become somewhat more complicated). We denote by $K_a^{(1)}$ the cone spanned by the origin of \mathbf{R}^{1+m} and the convex set

$$\{[Q_0 + Q_1 \Gamma(a, 0)] \delta x_0 + Q_1 f(a) \delta t + Q_1 \delta_1 x(a) : \delta x_0 \in \mathbf{R}^n, \\ \delta t \in \mathbf{R}, \delta_1 x(a) \in \delta_1 X(a)\} \subset \mathbf{R}^{1+m}.$$

This cone will be called the *first-order cone of the solution $\bar{x}(t)$, $\bar{u}(t)$ at the point $\bar{x}(a)$* .

Let $K_a^{(1)*}$ be the polar of the cone $K_a^{(1)}$, i.e. the set of those $(1+m)$ rows ζ for which $\zeta z < 0$ for all $z \in K_a^{(1)}$. Further, let

$$N_a = \{z \in \mathbf{R}^{1+m} : \zeta z = 0 \quad \forall \zeta \in K_a^{(1)*}\}.$$

Then N_a is the maximal linear subspace of $\overline{K_a^{(1)}}$ (the closure of the cone $K_a^{(1)}$).

LEMMA 1. A function $\psi(t) \in \Psi$ if and only if $\psi(a) = \zeta Q_1$, where $\zeta \in K_a^{(1)*}$, $\zeta \neq 0$.

The assertion of the lemma can be easily derived from the definitions of Ψ and $K_a^{(1)*}$.

Let $\sigma \in \Sigma^0$, and let there be chosen a linearization at the point $(\bar{x}(\sigma))$. We introduce the set (we do not explicitly show the dependence on the linearization)

$$T_\sigma = \left\{ \frac{1}{l!} Q_1 L_i^\sigma(\sigma) [\eta(t)] : l \geq 0 \text{ is an integer, } \eta(t) \in \mathcal{X}_\sigma^{(l)}, Q_1 L_i^\sigma(\sigma') [\eta(t)] \in N_a \right.$$

$$\left. \text{for all } i < l \text{ and } \sigma' \text{ sufficiently close to } \sigma \right\} \subset \mathbf{R}^{1+m}$$

The union of the sets T_σ over all points $\sigma \in \Sigma^0$ and all possible linearizations will be denoted by T (by analogy with [1], §3).

The cone spanned by the origin of \mathbf{R}^{1+m} and $\text{conv}(K_a^{(1)} \cup T)$ (the convex hull of the union of the first-order cone $K_a^{(1)}$ and T) will be called the *second-order cone $K_a^{(2)}$ of the solution $\bar{x}(t)$, $\bar{u}(t)$ at the point $\bar{x}(a)$* .

We obtain by Lemma 1 that Theorem 1' is equivalent to the assertion that the polar of the cone $\overline{K_a^{(2)}}$ is nonempty, or, in other words, the assertion that the cone $K_a^{(2)}$ does not coincide with \mathbf{R}^{1+m} .

Thus, it is sufficient to show that the equality $K_a^{(2)} = \mathbf{R}^{1+m}$ leads to a contradiction. This equality implies the existence of $d_1 + d_2$ vectors

$$\delta_1 z_1, \dots, \delta_1 z_{d_1} \in K_a^{(1)}, \quad \delta_2 z_1, \dots, \delta_2 z_{d_2} \in T,$$

such that the origin of \mathbf{R}^{1+m} is an interior point of the set

$$\text{conv} \{ \delta_1 z_1, \dots, \delta_1 z_{d_1}, \delta_2 z_1, \dots, \delta_2 z_{d_2} \}.$$

We have

$$\delta_2 z_j = \frac{1}{l_j!} Q_1 L_{l_j}^{\sigma_j}(\sigma_j) [\eta_j(t)], \quad j = 1, \dots, d_2,$$

and $Q_1 L_i^{\sigma_j}(\sigma_j) [\eta_j(t)] \in N_a$ for $i < l_j$ and σ_j' close to σ_j . Since $L_i^{\sigma}(\sigma') [\eta(t)]$ depends

continuously on σ' , the convex hull of the vectors

$$\delta_1 z_i, \quad i = 1, \dots, d_1, \quad \delta_2 z'_j = \frac{1}{l_j!} Q_1 L_{l_j}^{\sigma'_j}(\sigma'_j)[\eta_j(t)], \quad j = 1, \dots, d_2,$$

for σ' close to σ also contains the origin in its interior, i.e.

$$0 \in \text{int}(\text{conv} \{ \delta_1 z_1, \dots, \delta_1 z_{d_1}, \delta_2 z'_1, \dots, \delta_2 z'_{d_2} \}). \tag{4.7}$$

Obviously, moreover, the points $\sigma'_j, j = 1, \dots, d_2$, can be chosen pairwise distinct, whereas the σ_j are not, in general, pairwise distinct.

The central point of the proof is the following lemma (compare with the lemma in [1], §3).

BASIC LEMMA. *Let there be given vectors $\delta_1 z_1, \dots, \delta_1 z_{d_1} \in K_a^{(1)}$ and points $\sigma_j, \sigma'_j \in \Sigma^0, j = 1, \dots, d_2$, let there be chosen linearizations at the points $\sigma_j, j = 1, \dots, d_2$, and let there be given vectors*

$$\delta_2 z'_j = \frac{1}{l_j!} Q_1 L_{l_j}^{\sigma'_j}(\sigma'_j)[\eta_j(t)],$$

where $\eta_j(t) \in \mathcal{K}_{\sigma_j}^{(l_j)}$ and $Q_1 L_{l_j}^{\sigma'_j}(\sigma'_j)[\eta_j(t)] \in N_a$ for $i \leq l_j - 1$ and $j = 1, \dots, d_2$. Let

$$\mathcal{Q} = \left\{ \alpha = (\alpha_1, \dots, \alpha_{d_1+d_2}) : \alpha_i \geq 0, \sum_{i=1}^{d_1+d_2} \alpha_i = 1 \right\}$$

be a $(d_1 + d_2)$ -dimensional simplex.

If the points $\sigma'_j, j = 1, \dots, d_2$, are pairwise distinct, and if the numbers $|\sigma'_j - \sigma_j|$ are sufficiently small, then for all $\alpha \in \mathcal{Q}$ and all $\varepsilon > 0$ sufficiently small one can construct controls $u(t; \alpha, \varepsilon), u(t; \alpha, \varepsilon) \in U$ with $t \in \mathbf{R}$, perturbations of initial conditions $\delta x_0(\alpha, \varepsilon) \in \mathbf{R}^n$, and perturbations $\delta t(\alpha, \varepsilon)$ of the terminal instant of time such that the following assertion holds.

The equation $\dot{x} = f(x, u(t; \alpha, \varepsilon))$ with the initial condition $x(0) = \tilde{x}(0) + \delta x_0(\alpha, \varepsilon)$ has the solution $x(t; \alpha, \varepsilon), t \in [0, a + \delta t(\alpha, \varepsilon)]$, which depends continuously on $(t, \alpha) \in \{(t, \alpha) : \alpha \in \mathcal{Q}, 0 \leq t \leq a + \delta t(\alpha, \varepsilon)\}$ for every ε , and the following equality holds:

$$\left| Q(\tilde{x}(0) + \delta x_0(\alpha, \varepsilon), x(a + \delta t(\alpha, \varepsilon); \alpha, \varepsilon)) - Q(\tilde{x}(0), \tilde{x}(a)) \right. \tag{4.8}$$

$$\left. - \varepsilon \left(\sum_{i=1}^{d_1} \alpha_i \delta_1 z_i + \sum_{j=1}^{d_2} \alpha_{d_1+j} \delta_2 z'_j \right) \right| = \omega(\alpha; \varepsilon),$$

where $\omega(\alpha; \varepsilon)/\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$ uniformly in $\alpha \in \mathcal{Q}$.

Since (4.4) holds, the proof of the basic lemma follows the same pattern as that of the corresponding lemma in [1], §3. It is true that, since we consider the problem with moving endpoints, the cones $K_a^{(1)}$ and $K_a^{(2)}$ are constructed in a somewhat different way than in [1] (and even in a different space). However, this does not change the essence of the matter, and in this sense the proof of Theorem 1' is related to the proof presented in [1] in the same way as the derivation of the maximum principle for the problem with moving endpoints is related to its proof for the problem with fixed endpoints. One more peculiarity (as compared with [1]) consists of the necessity to return from a convex

problem (equation (1.4)) to the initial problem (equation (1.1)). This transfer is accomplished with the aid of the already mentioned "approximation lemma" in the same way as in the proof of the maximum principle with the aid of generalized controls or sliding regimes (see [2] and [3]).

Making use of the basic lemma, one can easily finish the proof of Theorem 1'. Indeed, by the Brouwer fixed point theorem one can derive the following assertion from (4.7) and (4.8):

For all sufficiently small $\varepsilon > 0$, the set

$$Z_\varepsilon = \{Q(\tilde{x}(0) + \delta x_0(\alpha, \varepsilon), x(a + \delta t(\alpha, \varepsilon); \alpha, \varepsilon)) - Q(\tilde{x}(0), \tilde{x}(a)) : \alpha \in \mathcal{Q}\}$$

contains the origin of \mathbf{R}^{1+m} as an interior point.

Thus the intersection of Z_ε with the ray

$$\{z = (z_0, z_1, \dots, z_m) \in \mathbf{R}^{1+m} : z_0 < 0, z_1 = \dots = z_m = 0\}$$

is nonempty. Therefore for $\hat{\alpha} \in \mathcal{Q}$ we have

$$q_0(x(0; \hat{\alpha}, \varepsilon), x(a + \delta t(\hat{\alpha}, \varepsilon); \hat{\alpha}, \varepsilon)) - q_0(\tilde{x}(0), \tilde{x}(a)) < 0,$$

and

$$q_i(x(0; \hat{\alpha}, \varepsilon), x(a + \delta t(\hat{\alpha}, \varepsilon); \alpha, \varepsilon)) = q_i(\tilde{x}(0), \tilde{x}(a)) = 0, \quad i = 1, \dots, m,$$

which contradicts the assumption that $q_0(\tilde{x}(0), \tilde{x}(a))$ is the minimum.

Received 7/OCT/76

BIBLIOGRAPHY

1. A. A. Agračev and R. V. Gamkrelidze, *A second order optimality principle for a time-optimal problem*, Mat. Sb. **100** (142) (1976), 610-643; English transl. in Math. USSR Sb. **29** (1976).
2. R. V. Gamkrelidze and G. L. Haratišvili, *Extremal problems in linear topological spaces*, Izv. Akad. Nauk SSSR Ser. Mat. **33** (1969), 781-839; English transl. in Math. USSR Izv. **3** (1969).
3. R. V. Gamkrelidze, *Foundations of optimal control*, Izdat. Tbiliss. Univ., Tbilisi, 1975. (Russian)

Translated by K. MAKOWSKI