A SECOND ORDER OPTIMALITY PRINCIPLE FOR A TIME-OPTIMAL PROBLEM

1976 Math. USSR Sb. 29 547

(http://iopscience.iop.org/0025-5734/29/4/A07)

MAT. SBORNIK
TOM 100(142)(1976), No. 4

MATH. USSR SBORNIK
Vol. 29(1976), No. 4

A SECOND ORDER OPTIMALITY PRINCIPLE FOR A TIME-OPTIMAL PROBLEM

UDC 519.3

A. A. AGRAČEV AND R. V. GAMKRELIDZE

ABSTRACT. In this paper, a general necessary optimality condition of second order is proved for a time-optimal problem. Necessary optimality conditions of second order are applied, in general, in studying singular optimal regimes which can not be found with the aid of the first-order necessary condition for optimality, i.e. with the aid of the Pontrjagin maximum principle.

Bibliography: 4 titles.

Recently, second-order necessary conditions have been intensively studied in view of their importance in the theory of singular extremals. Here, we give a general approach to the solution of this problem for a time-optimal problem with fixed end-points and linear control. For the case under consideration, the necessary condition that we formulate, Theorem 2.2, is apparently definitive. The general nonlinear case can be reduced to the case considered in this work with the aid of sliding regimes. However, this case requires a special investigation and will be published later.

The work consists of four sections. In §1, basic notions are introduced and the methods necessary for subsequent presentation are developed. In §2, the basic result of the work, i.e. the optimality principle in the form of Theorem 2.2, is stated. This principle is proved in §3. Finally, in §4 the basic operator necessary in order to formulate the optimality principle is expressed in terms of the right-hand side of a differential equation with the aid of Lie brackets. The expressions thus obtained are identical with the expressions contained in the work of A. J. Krener [4].

There is an extensive literature devoted to the problem studied here; we have made use of the works [1]–[4].

In conclusion, the senior author would like to express his gratitude to Professors Pavol Brunovský (Czechoslovakia), Claude Lobry (France), Czesław Olech (Poland) and Henry Hermes (USA) for useful discussions at the Banach Center in Warsaw during the winter of 1974.

§1. Legendre families (packets) of perturbations, Legendre representations of the second variation, and Legendre forms and operators

The controlled equation that we consider has the form

AMS (MOS) subject classifications (1970). Primary 49B10.

$$\dot{x} = f(t, x, u) = g(t, x) + G(t, x)u, \quad u \in U \subset \mathbf{R}^r, \tag{1.1}$$

where $g(t, x)$ is an n-dimensional infinitely differentiable column vector and $G(t, x)$ is an $n \times r$ infinitely differentiable matrix.(1) The set U of admissible values of the control parameter is an arbitrary closed convex polyhedron in $\mathbf{R}^r$ (not necessarily compact, so that coincidence with $\mathbf{R}^r$ is not excluded).

An arbitrary measurable function $u(t)$, $t \in R$, square-integrable on every finite interval and assuming values in U will be called a *control*.(2)

Let us fix a solution

$$\tilde{u}(t), \quad \tilde{x}(t), \quad 0 \leqslant t \leqslant a, \tag{1.2}$$

of equation (1.1). We shall denote by $\delta u(t)$ an arbitrary *perturbation* of the control $\tilde{u}(t)$ on $[0, a]$, i.e. a measurable and square-integrable function on $[0, a]$ which satisfies the condition $\tilde{u}(t) + \delta u(t) \in U$ for all $t \in [0, a]$.

Let us write down the Taylor expansion in δx and δu of $f(t, x + \delta x, u + \delta u)$, up to terms of third order:

$$\begin{aligned} f(t, x+\delta x, u+\delta u) &= f(t, x, u) + f_x(t, x, u)\,\delta x + f_u(t, x, u)\,\delta u \\ &+ f_{xu}(t, x, u)[\delta u, \delta x] + \frac{1}{2} f_{xx}(t, x, u)[\delta x, \delta x] + \cdots . \end{aligned}$$

We denote the linear and bilinear forms

$$f_x(t, x, u)\,\delta x, \quad f_u(t, x, u)\,\delta u, \quad f_{xu}(t, x, u)[\delta u, \delta x], \quad f_{xx}(t, x, u)[\delta x', \delta x''],$$

evaluated along the solution (1.2) respectively by

$$\begin{gathered} f_x(t)\,\delta x = f_x(t, \tilde{x}(t), \tilde{u}(t))\,\delta x, \quad f_u(t)\,\delta u = f_u(t, \tilde{x}(t), \tilde{u}(t))\,\delta u, \\ f_{xu}(t)[\delta u, \delta x] = f_{xu}(t, \tilde{x}(t), \tilde{u}(t))[\delta u, \delta x], \\ f_{xx}(t)[\delta x', \delta x''] = f_{xx}(t, \tilde{x}(t), \tilde{u}(t))[\delta x', \delta x'']. \end{gathered}$$

The solutions $\delta_1 x(t)$ and $\delta_2 x(t)$, $0 \leqslant t \leqslant a$, with zero initial values $\delta_1(x) = \delta_2(x) = 0$, of the equations (linear in $\delta_1 x(t)$ and $\delta_2 x(t)$ and having the same matrix $f_x(t)$ in the homogeneous part)

$$\begin{gathered} \delta_1 \dot{x} = f_x(t)\,\delta_1 x + f_u(t)\,\delta u(t), \\ \delta_2 \dot{x} = f_x(t)\,\delta_2 x + f_{xu}(t)[\delta u(t), \delta_1 x(t)] + \frac{1}{2} f_{xx}(t)[\delta_1 x(t), \delta_1 x(t)], \end{gathered}$$

will be called the *first* and *second variation of the trajectory* $\tilde{x}(t)$, $0 \leqslant t \leqslant a$, *corresponding to the perturbation* $\delta u(t)$.

Next, we denote by $\Gamma(t)$, $0 \leqslant t \leqslant a$, the fundamental matrix of the equation $\dot{\Gamma} = f_x(t)\Gamma$, and let $\Gamma(t, \tau) = \Gamma(t)\Gamma^{-1}(\tau)$. Finally, let $h(t)$ with $t \in \mathbf{R}$ be the Heaviside function:

$$h(t) = 0 \ \text{ for } \ t < 0, \quad h(0) = \frac{1}{2}, \quad h(t) = 1 \ \text{ for } \ t > 0.$$

(1) It is assumed that finite-dimensional spaces are arithmetic spaces; consequently matrix notation is used. Vectors denoted by small Roman letters are always columns, while vectors denoted by small Greek letters are always rows; the scalar product is formed only from a row and column of the same dimension. The operation of taking the adjoint matrix is denoted by a star.

(2) Since equation (1.1) is linear in u, for controls we can also take a wider class of functions, integrable on finite intervals. The definition given here has been adopted for purely technical reasons, namely for terminological convenience in investigating the integral quadratic forms that are defined below ("Legendre" forms).

We shall show that the end-point of the second variation $\delta_2 x(a)$ corresponding to the perturbation $\delta u(t)$ can be expressed as an integral quadratic form in $\delta u(t)$:

$$\delta_2 x(a) = \iint\limits_0^a h(t-s) B(t, s) [\delta u(t), \delta u(s)]\, dt\, ds,$$

where $B(t, s)[p_1, p_2]$ is an n-dimensional bilinear form in the r-dimensional columns p_1 and p_2, the explicit expression for which is given below.

The formula for the solution of a linear nonhomogeneous equation yields

$$\delta_2 x(a) = \int\limits_0^a \Gamma(a, t) \left\{ f_{xu}(t) [\delta u(t), \delta_1 x(t)] + \frac{1}{2} f_{xx}(t) [\delta_1 x(t), \delta_1 x(t)] \right\} dt.$$

Substituting here the expression for the first variation

$$\delta_1 x(t) = \int\limits_0^t \Gamma(t, s) f_u(s) \delta u(s)\, ds = \int\limits_0^a h(t-s) \Gamma(t, s) f_u(s) \delta u(s)\, ds,$$

we obtain

$$\delta_2 x(a) = \iint\limits_0^a h(t-s) \Gamma(a, t) f_{xu}(t) [\delta u(t), \Gamma(t, s) f_u(s) \delta u(s)]\, dt ds + \frac{1}{2} \int\limits_0^a \Gamma(a, \tau)\, d\tau$$
$$\times \iint\limits_0^a h(\tau-t) h(\tau-s) f_{xx}(\tau) [\Gamma(\tau, t) f_u(t) \delta u(t), \quad \Gamma(\tau, s) f_u(s) \delta u(s)]\, dt ds.$$

Introducing the bilinear forms

$$\begin{aligned} F_1(t, s)[p_1, p_2] &= f_{xu}(t) [p_1, \Gamma(t, s) f_u(s) p_2], \\ F_2(\tau, t, s)[p_1, p_2] &= f_{xx}(\tau) [\Gamma(\tau, t) f_u(t) p_1, \Gamma(\tau, s) f_u(s) p_2], \end{aligned} \tag{1.3}$$

we write

$$\begin{aligned} \delta_2 x(a) &= \iint\limits_0^a h(t-s) \Gamma(a, t) F_1(t, s) [\delta u(t), \delta u(s)]\, dt ds \\ &+ \frac{1}{2} \int\limits_0^a \Gamma(a, \tau)\, d\tau \iint\limits_0^a h(\tau-t) h(\tau-s) F_2(\tau, t, s) [\delta u(t), \delta u(s)]\, dt ds. \end{aligned} \tag{1.4}$$

Since the form $f_{xx}(\tau)[\delta x', \delta x'']$ is symmetric, the form F_2 satisfies a relation of "selfadjointness with respect to t, s and p_1, p_2":

$$F_2(\tau, t, s)[p_1, p_2] = F_2(\tau, s, t)[p_2, p_1]. \tag{1.5}$$

Therefore the function under the double integral in the second term in (1.4) is symmetric with respect to t and s. Moreover, the identity $h(t - s) + h(s - t) \equiv 1$ holds. Thus

$$\frac{1}{2} \iint\limits_0^a h(\tau-t) h(\tau-s) F_2(\tau, t, s) [\delta u(t), \delta u(s)]\, dt\, ds$$
$$= \iint\limits_0^a h(\tau-t) h(\tau-s) h(t-s) F_2(\tau, t, s) [\delta u(t), \delta u(s)]\, dt\, ds.$$

Hence

$$\frac{1}{2}\int_0^a \Gamma(a,\tau)\,d\tau \iint_0^a h(\tau-t)\,h(\tau-s)\,F_2(\tau,t,s)\,[\delta u(t),\delta u(s)]\,dt\,ds$$

$$=\iint_0^a h(t-s)\,dt ds \int_0^a h(\tau-t)\,h(\tau-s)\,\Gamma(a,\tau)\,F_2(\tau,t,s)\,[\delta u(t),\delta u(s)]\,d\tau$$

$$=\iint_0^a h(t-s)\,dt ds \int_t^a \Gamma(a,\tau)\,F_2(\tau,t,s)\,[\delta u(t),\delta u(s)]\,d\tau.$$

Substituting this expression for the second term in (1.4), we arrive at the required representation

$$\delta_2 x(a)=\iint_0^a h(t-s)\,B(t,s)\,[\delta u(t),\delta u(s)]\,dt\,ds,$$

$$B(t,s)\,[p_1,p_2]=\Gamma(a,t)\,F_1(t,s)\,[p_1,p_2]+\int_t^a \Gamma(a,\tau)\,F_2(\tau,t,s)\,[p_1,p_2]\,d\tau, \tag{1.6}$$

where the n-dimensional bilinear forms F_1 and F_2 are given by (1.3), and F_2 satisfies the selfadjointness condition (1.5).

Generalizing the notions of special variations introduced by Kelley, Kopp, and Moyer [1], and of packets of variations (Gabasov and Kirillova [2]), we shall now define Legendre families, or packets, of perturbations.

Let σ be an arbitrary point of the interval $(0, a)$. We perform a parallel shift of the side of smallest dimension of the polyhedron $U\subset \mathbf{R}^r$ containing the point $\tilde{u}(\sigma)$ to the origin of $\mathbf{R}^r$, and denote by R_σ the subspace spanned by the transferred side. If σ is a point of continuity of the control $\tilde{u}(t)$, we denote by π_σ the orthogonal projection of $\mathbf{R}^r$ onto R_σ; in the contrary case, π_σ denotes the projection of $\mathbf{R}^r$ into the origin. Note that, if $\tilde{u}(\sigma)$ is a vertex of the polyhedron U, then π_σ is the zero mapping.

We denote by $P^{(m)}$, $m\geqslant 0$, the set of all measurable and square-integrable r-dimensional column vectors on $[-1, 1]$ which satisfy the conditions

$$\int_{-1}^{1} t^i p(t)\,dt=0,\quad i=0,\ \ldots,\ m. \tag{1.7}$$

Thus the set $P^{(m)}$ is the subspace of the Hilbert space L_2^r containing all measurable and square-integrable r-dimensional functions $p(t)$ on the interval $[-1, 1]$ which consists of those $p(t)$ all coordinates of which are orthogonal to the first $m+1$ Legendre polynomials on $[-1, 1]$. For convenience, we shall assume that functions in L_2^r are defined for all $t\in\mathbf{R}$ and vanish outside of $[-1, 1]$.

Let there be given an integer $m\geqslant 0$ and an arbitrary point $\sigma\in(0, a)$. For any function $p(t)\in P^{(m)}$ and any two positive functions $\alpha(\varepsilon)$ and $\beta(\varepsilon)$, $\varepsilon>0$, which tend to zero as $\varepsilon\to 0$, the family of functions

$$\delta u(t;\varepsilon)=\alpha(\varepsilon)\,\pi_\sigma\,p\left(\frac{t-\sigma}{\beta(\varepsilon)}\right),\quad 0\leqslant t\leqslant a, \tag{1.8}$$

will be called the *Legendre family of perturbations* or the *packet* (of perturbations) of mth order determined by the point σ, the functions $p(t)\in P^{(m)}$, and $\alpha(\varepsilon)$ and $\beta(\varepsilon)$.

Obviously, for all $\varepsilon > 0$ sufficiently small, every function in the family (1.8) is a perturbation of the control $\tilde{u}(t)$.

Let us evaluate the end-point of the second variation $\delta_2 x(a)$ "on the packet" (1.8), i.e. let us substitute the packet (1.8) into (1.6) in place of an arbitrary perturbation $\delta u(t)$. Then we obtain

$$\delta_2 x\,(a;\,\varepsilon) = \alpha^2(\varepsilon) \iint_0^a h\,(\tau_1-\tau_2)\, B\,(\tau_1,\,\tau_2)\left[\pi_\sigma p\left(\frac{\tau_1-\sigma}{\beta(\varepsilon)}\right),\,\pi_\sigma p\left(\frac{\tau_2-\sigma}{\beta(\varepsilon)}\right)\right] d\tau_1 d\tau_2,$$

or, introducing the new variables

$$t = \frac{\tau_1-\sigma}{\beta(\varepsilon)}, \qquad s = \frac{\tau_2-\sigma}{\beta(\varepsilon)},$$

and taking into account the relation $\beta(\varepsilon) \to 0$,

$$\delta_2 x\,(a;\,\varepsilon) = \alpha^2(\varepsilon)\,\beta^2(\varepsilon) \iint_{-1}^{1} h(t-s)\, B\,(\sigma+\beta(\varepsilon)\,t,\ \sigma+\beta(\varepsilon)\,s)\,[\pi_\sigma p\,(t),\ \pi_\sigma p\,(s)]\, dt ds. \quad (1.9)$$

One can give a concise expression for the expansion of the bilinear form B under the integral in a power series in $\beta(\varepsilon)t$ and $\beta(\varepsilon)s$ at the point (σ, σ), without mentioning each time the order to which such an expansion is possible (which obviously depends on the order of differentiability of the control $\tilde{u}(t)$ at the point σ). One can do it by adopting the following notation and conventions, which we shall use throughout the presentation.

Any interior point of the set on which all partial derivatives up to order m of a function $\Phi(\tau_1, \tau_2, \ldots)$ of several variables exist and are continuous will be called a *point of m-fold differentiability*, $m \geqslant 0$, of this function.

Next, we introduce the differential operators $D_1, D_2, \ldots$ with respect to the variables $\tau_1, \tau_2, \ldots$. The symbol

$$D_1^{i_1} D_2^{i_2} \ldots \Phi\,(\tau_1,\,\tau_2,\,\ldots)\,|_{\tau_i=\sigma}$$

will mean the corresponding partial derivative at the point $(\sigma, \sigma, \ldots)$ if this point is an $(i_1 + i_2 + \ldots)$-fold point of differentiability of the function $\tilde{u}(t)$; in the contrary case this expression is equal to zero.

We shall make use of formal power series in the commuting operators $D_1, D_2, \ldots$, especially exponential series:

$$e^{\beta(\varepsilon)tD_1} = \sum_{i=0}^{\infty} \frac{(\beta(\varepsilon)\,t)^i}{i!} D_1^i, \qquad e^{\beta(\varepsilon)(tD_1+sD_2)} = e^{\beta(\varepsilon)tD_1} e^{\beta(\varepsilon)sD_2},$$

where the product of two formal power series is meant in the usual (Cauchy) sense.

We replace the kernel B in (1.9) by its formal expansion into an infinite Taylor series in powers of $\beta(\varepsilon)t$ and $\beta(\varepsilon)s$ at the point (σ, σ). We obtain the correspondence

$$\begin{aligned}\delta_2 x\,(a;\,\varepsilon) &\sim \alpha^2(\varepsilon)\,\beta^2(\varepsilon) \iint_{-1}^{1} h\,(t-s)\, e^{\beta(\varepsilon)(tD_1+sD_2)} B\,(\tau_1,\,\tau_2)\,[\pi_\sigma p\,(t),\ \pi_\sigma p\,(s)]\, dt ds \\ &= \alpha^2(\varepsilon)\,\beta^2(\varepsilon) \sum_{i=0}^{\infty} \frac{\beta^i(\varepsilon)}{i!} \iint_{-1}^{1} h\,(t-s)\,(tD_1+sD_2)^i\, B\,(\tau_1,\,\tau_2)\,[\pi_\sigma p\,(t),\ \pi_\sigma p\,(s)]\, dt ds,\end{aligned} \quad (1.10)$$

where the kernel of the $(i + 1)$th term

$$(tD_1 + sD_2)^i B(\tau_1, \tau_2)[p_1, p_2] = B_i(t, s; \sigma)[p_1, p_2], \qquad i = 0, 1, \ldots, \tag{1.11}$$

is an ith order homogeneous polynomial in t and s with n-dimensional coefficients which depend bilinearly on the r-dimensional vectors p_1 and p_2.

Obviously the meaning of this correspondence is that, at every point σ of m-fold differentiability of the control $\tilde{u}(t)$, the series in question yields an asymptotic expansion of the end-point $\delta_2 x(a; \varepsilon)$ of the second variation up to order m with respect to $\beta(\varepsilon)$. This means that the relation

$$\frac{1}{\beta^l(\varepsilon)} \left| \frac{\delta_2 x(a; \varepsilon)}{\alpha^2(\varepsilon)\beta^2(\varepsilon)} - \sum_{i=0}^{l} \frac{\beta^i(\varepsilon)}{i!} \int_{-1}^{1}\!\!\int h(t-s) B_i(t, s; \sigma)[\pi_\sigma p(t), \pi_\sigma p(s)]\, dt ds \right| \to 0$$
$$(\varepsilon \to 0)$$

holds for any $l \leqslant m$.

Independently of the order of differentiability of the control $\tilde{u}(t)$ at the point σ, the correspondence (1.10) will be called the *asymptotic expansion of the end-point of the second variation on the packet* (1.8) *at the point* σ.

Let us represent the kernel B_i as the sum $B_i = S_i + K_i$ of self- and skew-adjoint parts S_i and K_i, where

$$\begin{aligned} S_i(t, s; \sigma)[p_1, p_2] &= \frac{1}{2}(B_i(t, s; \sigma)[p_1, p_2] + B_i(s, t; \sigma)[p_2, p_1]), \\ K_i(t, s; \sigma)[p_1, p_2] &= \frac{1}{2}(B_i(t, s; \sigma)[p_1, p_2] - B_i(s, t; \sigma)[p_2, p_1]). \end{aligned} \tag{1.12}$$

The fact that the bilinear functions S_i and K_i are selfadjoint and skew-adjoint, respectively, means that

$$S_i(t, s; \sigma)[p_1, p_2] = S_i(s, t; \sigma)[p_2, p_1], \; K_i(t, s; \sigma)[p_1, p_2] = -K_i(s, t; \sigma)[p_2, p_1].$$

Replacing the kernels B_i in (1.10) by their skew-adjoint parts, we obtain the correspondence

$$\begin{aligned} \delta_2 x(a; \varepsilon) &\sim \alpha^2(\varepsilon)\beta^2(\varepsilon) \sum_{i=0}^{\infty} \frac{\beta^i(\varepsilon)}{i!} L_i(\sigma)[p(t)], \\ L_i(\sigma)[p(t)] &= \int_{-1}^{1}\!\!\int h(t-s) K_i(t, s; \sigma)[\pi_\sigma p(t), \pi_\sigma p(s)]\, dt ds, \end{aligned} \tag{1.13}$$

which will be called the *Legendre representation* of the end-point of the second variation on the packet (1.8). Since the coefficients of the representation $L_i(\sigma)[p(t)]$ do not depend on the functions $\alpha(\varepsilon)$ and $\beta(\varepsilon)$, but only on σ and $p(t)$, we shall also speak of the representation on a packet defined by the point σ and function $p(t)$, or even of the representation at the point σ, if the coefficients $L_i(\sigma)[p(t)]$ are viewed as quadratic forms of $p(t)$ that depend on the parameter σ.

In order to clarify the meaning of this definition, which is expressed in Proposition 1.2, we shall first prove Proposition 1.1.

Proposition 1.1. *Let $S(t, s)[p_1, p_2]$ be a polynomial of degree $\leqslant m$ $(m \geqslant 0)$ in t and s*

with vector coefficients that depend bilinearly on the r-dimensional column vectors p_1 and p_2, and let this polynomial satisfy the selfadjointness condition $S(t,s)[p_1,p_2] = S(s,t)[p_2,p_1]$. Then the integral quadratic form

$$\iint_{-1}^{1} h(t-s)S(t,s)[\pi_\sigma p(t),\ \pi_\sigma p(s)]\,dtds$$

in $p(t)$ is identically zero on $P^{(m)}$:

$$\iint_{-1}^{1} h(t-s)S(t,s)[\pi_\sigma p(t),\pi_\sigma p(s)]\,dtds = 0 \quad \forall\, p(t)\in P^{(m)}.$$

PROOF. Since, together with $p(t)$, $\pi_\sigma p(t)$ also belongs to the set $P^{(m)}$, and since the degree of the polynomial S does not exceed m,

$$S(t,s)[p_1,p_2] = \sum_{i+j\leqslant m} t^i s^j S_{ij}[p_1,p_2] = \sum_{i+j\leqslant m} S_{ij}[t^i p_1, s^j p_2],$$

we have by (1.7)

$$\begin{aligned}
&\iint_{-1}^{1} S(t,s)[\pi_\sigma p(t),\pi_\sigma p(s)]\,dtds\\
&= \sum_{i+j\leqslant m}\iint_{-1}^{1} S_{ij}[t^i\pi_\sigma p(t), s^j\pi_\sigma p(s)]\,dtds\\
&= \sum_{i+j\leqslant m} S_{ij}\left[\int_{-1}^{1} t^i\pi_\sigma p(t)\,dt,\ \int_{-1}^{1} s^j\pi_\sigma p(s)\,ds\right] = 0.
\end{aligned}$$

Therefore, making use of the identity $h(t-s)\equiv 1-h(s-t)$ and of the fact that S is selfadjoint, we obtain

$$\begin{aligned}
&\iint_{-1}^{1} h(t-s)S(t,s)[\pi_\sigma p(t),\pi_\sigma p(s)]\,dtds\\
&\equiv -\iint_{-1}^{1} h(s-t)S(s,t)[\pi_\sigma p(s),\pi_\sigma p(t)]\,dtds,
\end{aligned}$$

which proves the proposition.

As an immediate corollary, we obtain

PROPOSITION 1.2. *The first $m+1$ coefficients $L_0,\dots,L_m$ in the Legendre representation* (1.13) *of the end-point of the second variation on an arbitrary mth order packet coincide with the first $m+1$ coefficients*

$$\iint_{-1}^{1} h(t-s)B_i(t,s;\sigma)[\pi_\sigma p(t),\pi_\sigma p(s)]\,dtds,\quad i=0,\dots,m,$$

of the corresponding asymptotic expansion (1.10).

We now define the third basic notion of this section, namely the notion of Legendre forms.

We denote by $Q_a^{(1)}$ the convex cone spanned from the origin in $\mathbf{R}^n$ by the set of first variations $\delta_1 x(a)$ that correspond to all possible perturbations $\delta u(t)$ of the control $\tilde u(t)$. This cone will be called the *first order cone for the solution* (1.2). We denote by $N_a\subset\mathbf{R}^n$

the maximal subspace contained in the closure $\bar{Q}_a^{(1)} \subset \mathbf{R}^n$ of the cone $Q_a^{(1)}$, and we denote by $\overset{*}{Q}_a^{(1)}$ the cone dual to $Q_a^{(1)}$ (and therefore dual to $\bar{Q}_a^{(1)}$), i.e. the set of all n-dimensional row vectors χ such that $\chi\delta x \leqslant 0\ \forall \delta x \in Q_a^{(1)}$. The subspace N_a is the intersection of all the supporting hyperplanes to the cone $Q_a^{(1)}$. Therefore we can write

$$N_a = \bigcap_{\chi \in \overset{*}{Q}_a^{(1)}} N_\chi, \tag{1.14}$$

where N_χ is the subspace of $\mathbf{R}^n$ orthogonal to χ. If $Q_a^{(1)} = \mathbf{R}^n$, then $N_a = \mathbf{R}^n$, and the right-hand side of (1.14) is to be understood as the intersection of an empty set of subspaces N_χ.

To every coefficient $L_m(\sigma)[p(t)]$ in the Legendre representation (1.13), there corresponds a family of integral quadratic forms in $p(t)$ depending on the parameter χ:

$$\omega_m(\chi, \sigma)[p(t)] = \iint_{-1}^{1} h(t-s)\, \chi K_m(t, s; \sigma)[\pi_\sigma p(t), \pi_{\bar\sigma} p(s)]\, dt ds, \quad \chi \in \overset{*}{Q}_a^{(1)}, \tag{1.15}$$

which will be called the *Legendre forms of the solution* (1.2) *at the point* σ.

The following proposition gives an explicit expression for the kernel of the Legendre form $\chi K_m(t, s; \sigma)[p_1, p_2]$ in terms of the form $B(t, s)[p_1, p_2]$ in a particular case. Such an expression is necessary for the statement of the optimality principle (Theorem 2.2).

PROPOSITION 1.3. *Let a vector $\chi \in \overset{*}{Q}_a^{(1)}$, an integer $m \geqslant 0$, and r-dimensional vectors p_1 and p_2 be such that all the polynomials $\chi K_l(t, s; \sigma)[p_1, p_2]$ in t and s vanish identically for any point σ of the interval $O \subset (0, a)$ and any $l \leqslant m - 1$. Then*

$$\begin{gathered} \chi K_m(t, s; \sigma)[p_1, p_2] = \frac{1}{2}(s-t)^m \Omega_m(\chi, \sigma)[p_1, p_2], \\ \Omega_m(\chi, \sigma)[p_1, p_2] = \chi\{D_2^m B(\sigma, \tau_2)[p_1, p_2] - D_1^m B(\tau_1, \sigma)[p_2, p_1]\} \quad \forall \sigma \in O, \end{gathered} \tag{1.16}$$

where

$$\Omega_m(\chi, \sigma)[p_1, p_2] = \Omega_m(\chi, \sigma)[p_2, p_1] \tag{1.17}$$

for odd m, and

$$\Omega_m(\chi, \sigma)[p_1, p_2] = -\Omega_m(\chi, \sigma)[p_2, p_1] \tag{1.18}$$

for even m.

PROOF. We denote by $\mathfrak{P}$ the permutation of the variables p_1 and p_2 in the expression for $B(\tau_1, \tau_2)[p_1, p_2]$:

$$\mathfrak{P}B(\tau_1, \tau_2)[p_1, p_2] = B(\tau_1, \tau_2)[p_2, p_1].$$

On the basis of (1.11) and (1.12), one can write

$$\begin{aligned} \sum_{i=0}^{\infty} \frac{\beta^i(\varepsilon)}{i!} K_i(t, s; \sigma)[p_1, p_2] &= \frac{1}{2}\{e^{\beta(\varepsilon)(tD_1+sD_2)} - e^{\beta(\varepsilon)(sD_1+tD_2)}\mathfrak{P}\} B(\tau_1, \tau_2)[p_1, p_2] \\ &= \frac{1}{2} e^{\beta(\varepsilon)t(D_1+D_2)} \{e^{\beta(\varepsilon)(s-t)D_2} - e^{\beta(\varepsilon)(s-t)D_1}\mathfrak{P}\} B(\tau_1, \tau_2)[p_1, p_2], \end{aligned} \tag{1.19}$$

where the derivatives are taken, as before, at the point $\tau_1 = \sigma$, $\tau_2 = \sigma$. Since the application of the operator $D_1 + D_2$ to an arbitrary function $\Phi(\tau_1, \tau_2)$ at the point $\tau_1 = \tau_2 = \sigma$ can be expressed in the form of the total derivative

$$\frac{d}{d\tau}\Phi(\tau, \tau)\big|_{\tau=\sigma} = D_1\Phi(\tau_1, \sigma) + D_2\Phi(\sigma, \tau_2),$$

(1.16) follows from (1.19) by an obvious inductive argument.

Since $K_m(t, s; \sigma)[p_1, p_2]$ is a skew-adjoint polynomial in t and s, the equality

$$(s-t)^m \Omega_m(\chi, \sigma)[p_1, p_2] = -(t-s)^m \Omega_m(\chi, \sigma)[p_2, p_1]$$

holds for all $m \geqslant 0$; which is equivalent to (1.17) and (1.18).

The expression (1.16) for Ω_m can be further simplified. In order to do this, we replace the form B in the difference within the braces in (1.16) by the expression for B from (1.6). We obtain

$$D_2^m\Gamma(a, \sigma)F_1(\sigma, \tau_2)[p_1, p_2] - D_1^m\Gamma(a, \tau_1)F_1(\tau_1, \sigma)[p_2, p_1]$$
$$+ D_2^m\int_\sigma^a \Gamma(a, \tau)F_2(\tau, \sigma, \tau_2)[p_1, p_2]\,d\tau - D_1^m\int_{\tau_1}^a \Gamma(a, \tau)F_2(\tau, \tau_1, \sigma)[p_2, p_1]\,d\tau.$$

We introduce the operator D_0 of differentiation with respect to τ_0, and by the symmetry property (1.5) of F_2 we write

$$D_2^m\int_\sigma^a \Gamma(a, \tau)F_2(\tau, \sigma, \tau_2)[p_1, p_2]\,d\tau - D_1^m\int_{\tau_1}^a \Gamma(a, \tau)F_2(\tau, \tau_1, \sigma)[p_2, p_1]\,d\tau$$
$$= D_1^m\int_\sigma^{\tau_1} \Gamma(a, \tau)F_2(\tau, \tau_1, \sigma)[p_2, p_1]\,d\tau$$
$$= D_1^m\int_\sigma^{\tau_1}\sum_{i=0}^{m-1}\frac{(\tau-\sigma)^i}{i!}D_0^i\Gamma(a, \tau_0)F_2(\tau_0, \tau_1, \sigma)[p_2, p_1]\,d\tau$$
$$= D_1^m\sum_{i=1}^{m}\frac{(\tau_1-\sigma)^i}{i!}D_0^{i-1}\Gamma(a, \tau_0)F_2(\tau_0, \tau_1, \sigma)[p_2, p_1].$$

By the identity $\Gamma(a, t) = \Gamma(a, \sigma)\Gamma(\sigma, t)$, we arrive at

$$\Omega_m(\chi, \sigma)[p_1, p_2] = \chi\Gamma(a, \sigma)\{D_2^m F_1(\sigma, \tau_2)[p_1, p_2]$$
$$- D_1^m\Gamma(\sigma, \tau_1)F_1(\tau_1, \sigma)[p_2, p_1]$$
$$+ D_1^m\sum_{i=1}^{m}\frac{(\tau_1-\sigma)^i}{i!}D_0^{(i-1)}\Gamma(\sigma, \tau_0)F_2(\tau_0, \tau_1, \sigma)[p_2, p_1]\},$$

where the function in parentheses is evaluated at the point $\tau_0 = \tau_1 = \tau_2 = \sigma$, and the third term is to be set equal to zero for $m = 0$.

The expression in parentheses is an n-dimensional bilinear form in the r-dimensional vectors p_1 and p_2, and it can be viewed as the result of applying a certain operator $\mathfrak{L}_m$ to the function $f(t, x, u)$ "along the solution $\tilde{x}(t)$, $\tilde{u}(t)$ at the point $t = \sigma$." The operator $\mathfrak{L}_m$ has a "local character"; namely, it is expressed explicitly in terms of the partial derivatives of f at the point $(\sigma, \tilde{x}(\sigma), \tilde{u}(\sigma))$ and the derivatives of $\tilde{u}(t)$ up to order $m + 2$ at the point σ. Indeed, the values $D_1^i D_2^j\Gamma(\tau_1, \tau_2)$ evaluated at $\tau_1 = \tau_2 = \sigma$ are expressed in terms of these derivatives. This follows at once from an obvious inductive argument, since $\Gamma(\tau_1, \tau_2) = \Gamma(\tau_1)\Gamma^{-1}(\tau_2)$, and since the functions $\Gamma(\tau_1)$ and $\Gamma^{-1}(\tau_2)$ satisfy the adjoint differential equations

$$\dot{\Gamma}(\tau_1) = f_x(\tau_1)\Gamma(\tau_1), \quad \dot{\Gamma}^{-1}(\tau_2) = -\Gamma^{-1}(\tau_2)f_x(\tau_2).$$

The operator $\mathfrak{L}_m$ will be called the *Legendre operator of order m* ($\geqslant 0$); and the result of its application to the function $f(t, x, u)$ along the curve $\tilde{x}(t)$, $\tilde{u}(t)$ evaluated at the point σ will be denoted by

$$\mathfrak{L}_m f(\sigma, \tilde{x}(\sigma), \tilde{u}(\sigma))[p_1, p_2] = \mathfrak{L}_m f(\sigma)[p_1, p_2].$$

Thus

$$\begin{aligned} \mathfrak{L}_m f(\sigma, \tilde{x}(\sigma), \tilde{u}(\sigma))[p_1, p_2] &= D_2^m F_1(\sigma, \tau_2)[p_1, p_2] - D_1^m \Gamma(\sigma, \tau_1) F_1(\tau_1, \sigma)[p_2, p_1] \\ &+ \sum_{i=1}^{m} \frac{m!}{i!\,(m-i)!} D_0^{(i-1)} D_1^{(m-i)} \Gamma(\sigma, \tau_0) F_2(\tau_0, \tau_1, \sigma)[p_2, p_1]. \end{aligned} \tag{1.20}$$

In conclusion, we shall prove two auxiliary propositions, which are made use of in the sequel.

PROPOSITION 1.4. *Let $\Omega(t, s)[p_1, p_2]$ be a skew-adjoint homogeneous polynomial in t and s of even degree $m \geqslant 0$ with scalar coefficients that depend bilinearly on p_1 and p_2, $\Omega(t, s)[p_1, p_2] = -\Omega(s, t)[p_1, p_2]$. Assume that the quadratic form in $p(t)$*

$$\omega[p(t)] = \iint_{-1}^{1} h(t-s)\,\Omega(t, s)[p(t), p(s)]\,dt\,ds$$

has a constant sign, e.g. is nonnegative on $P^{(m)}$:

$$\omega[p(t)] \geqslant 0 \quad \forall\, p(t) \in P^{(m)}.$$

Then the form $\omega[p(t)]$ is identically zero on $P^{(m)}$, and therefore, by Proposition 1.5,

$$\Omega(t, s)[p_1, p_2] \equiv 0.$$

PROOF. Since the degree of Ω is m, the condition $p(t) \in P^{(m)}$ yields (see (1.7))

$$\begin{aligned} \omega[p(t)] &= \iint_{-1}^{1} (1 - h(s-t))\,\Omega(t, s)[p(t), p(s)]\,dt\,ds \\ &= -\iint_{-1}^{1} h(s-t)\,\Omega(t, s)[p(t), p(s)]\,dt\,ds. \end{aligned}$$

Further, since m is even,

$$\Omega(t, s)[p_1, p_2] = \Omega(-t, -s)[p_1, p_2].$$

Moreover, together with $p(t)$, the function $\hat{p}(t) = p(-t)$ belongs to the subspace $P^{(m)}$. Therefore, performing the substitution $t = -t'$, $s = -s'$ in the double integral, we obtain

$$\omega[p(t)] = -\iint_{-1}^{1} h(t'-s')\,\Omega(t', s')[\hat{p}(t'), \hat{p}(s')]\,dt'\,ds' = -\omega[\hat{p}(t)].$$

Since ω does not change sign on $P^{(m)}$, we have $\omega[p(t)] = 0\ \forall p(t) \in P^{(m)}$.

PROPOSITION 1.5. *Let $K(t, s)[p_1, p_2]$ be a skew-adjoint homogeneous polynomial in t and s of degree $m \geqslant 0$ with vector coefficients that depend bilinearly on the r-dimensional vectors p_1 and p_2. If, for some $l \geqslant 0$ (which does not depend on m), the form*

$$L[p(t)] = \iint_{-1}^{1} h(t-s)\,K(t, s)[p(t), p(s)]\,dt\,ds = 0 \quad \forall\, p(t) \in P^{(l)},$$

then the kernel $K(t, s)[p_1, p_2] \equiv 0$.

PROOF. It is sufficient to assume that $l \geqslant m$, because $P^{(l'')} \subset P^{(l')}$ for $l' \leqslant l''$. Moreover, it can be assumed that K is a polynomial in t and s with scalar coefficients. Therefore it can be represented in the form

$$K(t, s)[p_1, p_2] = p_1^* K(t, s) p_2,$$

where $K(t, s)$ is an $r \times r$ homogeneous matrix polynomial of order m which satisfies the skew-adjointness condition $K(t, s) = -K^*(s, t)$.

In the Hilbert space L_2^r of r-dimensional square-integrable functions $p(t)$ on $[-1, 1]$, we define the (completely continuous) "Volterra operator" V by the formula

$$Vp(t) = \int_{-1}^{t} K(t, s) p(s) ds. \tag{1.21}$$

Then for any $p_1(t), p_2(t) \in L_2^r$ we have

$$\iint_{-1}^{1} h(t-s) K(t, s)[p_1(t), p_2(s)] dtds = \int_{-1}^{1} p_1^*(t) dt \int_{-1}^{t} K(t, s) p_2(s) ds$$

$$= \int_{-1}^{1} p_1^*(t) V p_2(t) dt = (p_1(t), V p_2(t)),$$

where $(\cdot, \cdot)$ is the scalar product in L_2^r.

It is easy to see that

$$(p_1(t), Vp_2(t)) = (Vp_1(t), p_2(t)) \quad \forall p_1(t), p_2(t) \in P^{(l)}. \tag{1.22}$$

Indeed, since the degree of $K(t, s)$ does not exceed m, and since $p_1(t), p_2(t) \in P^{(l)}$, $l \geqslant m$, it follows that

$$\iint_{-1}^{1} p_1^*(t) K(t, s) p_2(s) dtds = 0.$$

Therefore, making use of the identity $h(t-s) \equiv 1 - h(s-t)$ and of the fact that $K(t, s)$ is skew-adjoint, we obtain the required equality

$$(p_1(t), Vp_2(t)) = -\iint_{-1}^{1} h(s-t) p_2^*(s) K^*(t, s) p_1(t) dtds$$

$$= \iint_{-1}^{1} h(s-t) p_2^*(s) K(s, t) p_1(t) dtds = (p_2(t), Vp_1(t)).$$

We denote by $\mathfrak{N}$ the orthogonal projection of L_2^r onto the subspace $P^{(l)}$. The mapping $\mathfrak{N}V$ takes $P^{(l)}$ into itself, and by virtue of (1.22) it is selfadjoint on $P^{(l)}$. Moreover,

$$(p(t), Vp(t)) = (\mathfrak{N}p(t), Vp(t)) = (p(t), \mathfrak{N}Vp(t)) \quad \forall p(t) \in P^{(l)}.$$

Therefore, the fact that the form $L[p(t)] = (p(t), Vp(t))$ vanishes on $P^{(l)}$ is equivalent to the fact that the selfadjoint operator $\mathfrak{N}V$ vanishes on $P^{(l)}$.

We shall show that

$$\mathfrak{N}Vp(t) = 0 \quad \forall p(t) \in P^{(l)} \tag{1.23}$$

implies $K(t, s) \equiv 0$.

The dimension of the orthogonal complement of $P^{(l)} \subset L_2^r$ is finite (it is equal to $r(m+1)$). Therefore (1.23) implies that the dimension of the subspace $VL_2^r \subset L_2^r$ is also finite (and does not exceed $2r(m+1)$).

The proposition will be proved if we show that the Volterra operator (1.21), whose kernel is a nonzero $r \times r$ homogeneous matrix polynomial, cannot map L_2^r onto a finite-dimensional subspace.

Let $K_{ij}(t, s)$ be a nonzero element of the matrix $K(t, s)$. By virtue of the homogeneity, it can be expressed in the form

$$K_{ij}(t, s) = \sum_{k=l}^{m} F_k s^{m-k} (t-s)^k,$$

where $F_l \neq 0$, $0 \leqslant l \leqslant m$. We take an arbitrary, linearly independent sequence of infinitely differentiable scalar functions $b_1(t)$, $b_2(t), \ldots$ on $[-1, 1]$ which vanish on an interval $-1 \leqslant t \leqslant \varepsilon$, $\varepsilon > 0$. It is easy to see that the Volterra integral equation of the first kind

$$\int_{-1}^{t} K_{ij}(t, s) z_k(s)\, ds = b_k(t) \tag{1.24}$$

is solvable in $z_k(t)$, $-1 \leqslant t \leqslant 1$, for every $k = 1, 2, \ldots$.

Indeed, differentiating both sides of the equation $l + 1$ times, we obtain

$$l!\, F_l t^{m-l} z_k(t) + \int_{-1}^{t} W(t, s) z_k(s)\, ds = b_k^{(l+1)}(t),$$

whose solution $z_k(t)$, $-1 \leqslant t \leqslant 1$, is zero on the interval $-1 \leqslant t \leqslant \varepsilon$, and coincides on the interval $\varepsilon \leqslant t \leqslant 1$ with the solution of the Volterra integral equation of the second kind

$$l!\, F_l t^{m-l} z(t) + \int_{\varepsilon}^{t} W(t, s) z(s)\, ds = b_k^{(l+1)}(t).$$

Therefore the function $z_k(t)$ is a solution of the original equation (1.24).

With the aid of the solutions $z_k(t)$, we construct an infinite sequence of r-dimensional functions

$$p_k(t) = (p_k^1(t), \ \ldots, \ p_k^r(t))^* \in L_2^r, \quad k = 1, 2, \ \ldots,$$

defined by the conditions $p_k^{k'}(t) = 0$ for $k' \not\equiv j$ and $p_k^j(t) = z_k(t)$. The mapping V takes this sequence into a linearly independent sequence $Vp_k(t)$, $k = 1, 2, \ldots$, since the jth coordinate of the vector $Vp_k(t)$ is equal to $z_k(t)$. This completes the proof of the proposition.

§2. Statement of the optimality principle

Beginning with this section, we assume that

$$\tilde{u}(t), \ \tilde{x}(t), \quad 0 \leqslant t \leqslant a, \tag{2.1}$$

is a time-optimal solution of the control problem (1.1) with fixed end-points. We denote by Ψ the set of all nonzero solutions $\psi(t)$, $0 \leqslant t \leqslant a$, of the equation

$$\dot{\psi} = -\psi f_x(t, \tilde{x}(t), \tilde{u}(t)) = -\psi f_x(t), \tag{2.2}$$

which satisfy the maximum condition

$$\psi(t) f(t) = \sup_{u \in U} \psi(t) f(t, \tilde{x}(t), u) \tag{2.3}$$

for almost all $t \in [0, a]$. A solution $\psi(t)$ of (2.2) obviously belongs to the set Ψ if and only if $\psi(a) \in \overset{*}{Q}{}_a^{(1)}$ and $\psi(a) \neq 0$.

The basic optimality criterion is expressed by the following theorem.

THEOREM 2.1. *There exists a function* $\tilde{\psi}(t) \in \Psi$ *such that the following two assertions hold for an arbitrary point* $\hat{\sigma} \in (0, a)$:

(A) *The Legendre form* $\omega_0(\tilde{\psi}(a), \hat{\sigma})[p(t)] = \tilde{\psi}(a)L_0(\hat{\sigma})[p(t)]$ *vanishes identically on* $P^{(0)}$:

$$\omega_0(\tilde{\psi}(a), \hat{\sigma})[p(t)] = 0 \quad \forall\, p(t) \in P^{(0)}. \tag{2.4}$$

(B) *If* $m \geqslant 1$, *and if a function* $\hat{p}(t) \in P^{(m)}$ *is such that the equalities*

$$\omega_i(\chi, \sigma)[\pi_{\hat{\sigma}}\hat{p}(t)] = 0 \quad \forall\, \chi \in \overset{*}{Q}{}_a^{(1)},\ \forall\, \sigma \in O_{\hat{\sigma}}, \quad \forall\, i = 0, \ldots, m-1, \tag{2.5}$$

hold for a neighborhood $O_{\hat{\sigma}}$ *of the point* $\hat{\sigma}$, *then*

$$\omega_m(\tilde{\psi}(a), \hat{\sigma})[\hat{p}(t)] \leqslant 0. \tag{2.6}$$

The assertions (A) and (B) are trivial if $\hat{\sigma}$ is a discontinuity point of the control $\tilde{u}(t)$, or if $\tilde{u}(\hat{\sigma})$ is a vertex of the polyhedron U. This is so because in these cases $\pi_{\hat{\sigma}} = 0$, and therefore

$$\omega_i(\chi, \hat{\sigma})[\hat{p}(t)] = \omega_i(\chi, \hat{\sigma})[\pi_{\hat{\sigma}}\hat{p}(t)] = 0$$

for an arbitrary χ and $i = 0, 1, 2, \ldots$. For this reason, Theorem 2.1 does not yield anything new for such points $\hat{\sigma}$, since the condition $\tilde{\psi}(t) \in \Psi$ is the Pontrjagin maximum principle.

The assertions (A) and (B) can be given a simple geometric interpretation.

The assertion (A), i.e. (2.4), means that the first coefficient in the Legendre representation of the end-point of the second variation of the optimal trajectory $\tilde{x}(t)$ on an arbitrary packet of zero order lies on the supporting hyperplane $N_{\tilde{\psi}(a)}$ to the cone $Q_a^{(1)}$ which is orthogonal to the vector $\tilde{\psi}(a)$.

The assertion (B) can be reformulated as follows. If, in the Legendre representation of the end-point of the second variation on the mth order packet defined by an arbitrary point $\sigma \in O_{\hat{\sigma}}$ and a fixed function $\pi_{\hat{\sigma}}\hat{p}(t) \in P^{(m)}$, the first m coefficients

$$L_0(\sigma)[\pi_{\hat{\sigma}}\hat{p}(t)], \ \ldots, \ L_{m-1}(\sigma)[\pi_{\hat{\sigma}}\hat{p}(t)]$$

lie in the subspace N_a (see (1.14)), then the $(m+1)$th coefficient $L_m(\hat{\sigma})[\pi_{\hat{\sigma}}\hat{p}(t)] = L_m(\hat{\sigma})[\hat{p}(t)]$ in the representation defined by the point $\hat{\sigma}$ and the function $\hat{p}(t)$ lies on the same side of the supporting hyperplane $N_{\tilde{\psi}(a)}$ as the cone $Q_a^{(1)}$. If, moreover, the vector $L_m(\hat{\sigma})[\hat{p}(t)]$ does not lie in N_a, then the chain of equalities (2.5) terminates at $i = m - 1$, and nothing can be said a priori about the coefficient $L_{m+1}(\hat{\sigma})[\hat{p}(t)]$. But also in the case $L_m(\hat{\sigma})[\hat{p}(t)] \in N_a$ one cannot draw any conclusions concerning $L_{m+1}(\hat{\sigma})[\hat{p}(t)]$. This is so because in order to pass to the $(m+1)$th coefficient we must have the relations $L_m(\sigma)[\pi_{\hat{\sigma}}\hat{p}(t)] \in N_a$ for any $\sigma \in O_{\hat{\sigma}}$.

In the form presented, the assertion (B) is entirely useless in practical application. Indeed, if the integral quadratic forms in $p(t)$

$$\omega_i(\tilde{\psi}(a), \sigma)[\pi_{\hat{\sigma}}p(t)] = \tilde{\psi}(a)\, L_i(\sigma)[\pi_{\hat{\sigma}}p(t)], \quad \sigma \in O_{\hat{\sigma}}, \quad i = 0, \ldots, m-1,$$

vanish not identically in $p(t) \in P^{(m)}$ but only for particular values of $\hat{p}(t) \in P^{(m)}$, then we have to verify the relations (2.4)–(2.6) for these particular values. However, there are no effective criteria (using, for example, the kernel of the form) that would allow one to find the particular $\hat{p}(t)$ at which the forms vanish, or to verify for these $\hat{p}(t)$ the sign of the form without a direct computation.

Nevertheless, from Theorem 2.1 and the auxiliary assertions proved in §1 we at once easily obtain an optimality principle convenient for actual application, namely Theorem 2.2.

An arbitrary point $\hat{\sigma} \in (0, a)$ will be called a *stationary point of rank no less than m* ($m \geqslant 0$) of the Legendre representation of the second variation end-point at $\hat{\sigma}$ if there exists a neighborhood $O_{\hat{\sigma}}^{(m)}$ such that the first m coefficients of the expansion of L_i satisfy the conditions

$$L_i(\sigma)[\pi_{\hat{\sigma}} p(t)] \in N_a, \quad \forall \sigma \in O_{\hat{\sigma}}^{(m)}, \quad \forall p(t) \in P^{(m)}, \quad \forall i \leqslant m-1,$$

i.e. if

$$\omega_i(\chi, \sigma)[\pi_{\hat{\sigma}} p(t)] \equiv 0 \quad \text{on } P^{(m)}, \quad \forall \chi \in \overset{*}{Q}{}_a^{(1)}, \quad \forall \sigma \in O_{\hat{\sigma}}^{(m)}, \quad \forall i \leqslant m-1.$$

A stationary point $\hat{\sigma}$ of rank no less than m is said to be a *stationary point of rank m* if there exists a row $\hat{\chi} \in \overset{*}{Q}{}_a^{(1)}$ such that $\omega_m(\hat{\chi}, \hat{\sigma})[p(t)] \not\equiv 0$ on $P^{(m)}$.

The set of all stationary points of rank m is denoted by D_m, and the set of all stationary points of rank no less than m, by E_m. The closure of D_m is denoted by $\bar{D}_m$. The sets D_m and E_m are open, $E_m \supset D_m$ and it follows from Proposition 1.5 that $E_{m'} \subset E_{m''}$ for $m' \geqslant m''$. Moreover, $E_0 = (0, a)$.

An obvious induction argument shows that

$$(0, a) \subset \bigcup_{-1 \leqslant i \leqslant m-1} \bar{D}_i \cup \bar{E}_m, \quad \forall m \geqslant 0 \quad (D_{-1} = \varnothing).$$

Therefore, for any $m \geqslant 0$, the open set $\bigcup_{i \leqslant m-1} D_i \cup E_m$ is everywhere dense in the interval $(0, a)$, since this set is obtained from $\bigcup_{i \leqslant m-1} \bar{D}_i \cup \bar{E}_m$ by deleting the boundaries of the open sets D_i.

We shall say that the solution (2.1) *satisfies the optimality criterion of rank* $m \geqslant 0$ *with a function* $\tilde{\psi}(t) \in \Psi$ *at a point* $\hat{\sigma} \in (0, a)$ if there exists a neighborhood $O_{\hat{\sigma}}^{(m)}$ such that for any $l \leqslant m$ the conditions

$$\omega_i(\chi, \sigma)[\pi_{\hat{\sigma}} p(t)] \equiv 0 \quad \text{on } P^{(l)} \quad \forall \chi \in \overset{*}{Q}{}_a^{(1)}, \quad \forall \sigma \in O_{\hat{\sigma}}^{(m)}, \quad \forall i \leqslant l-1, \quad (2.7)$$

imply the inequality

$$\omega_l(\tilde{\psi}(a), \hat{\sigma})[p(t)] \leqslant 0 \quad \forall p(t) \in P^{(l)}, \quad (2.8)$$

which for even l is equivalent to the identity

$$\omega_l(\tilde{\psi}(a), \hat{\sigma})[p(t)] \equiv 0 \quad \text{on } P^{(l)}$$

(see Proposition 1.4).

Theorem 2.1 implies the existence of a function $\tilde{\psi}(t) \in \Psi$ such that, for arbitrary $i \geqslant 0$ and $i \leqslant m$, the optimal solution (2.1) satisfies the optimality criterion of rank m at all points of the sets D_i and E_m. Therefore the successive construction of the sets D_i and E_i, $i = 0, 1, 2, \ldots$, yields at the mth step the open set

$$\bigcup_{i \leqslant m-1} D_i \cup \bar{E}_m, \quad m = 0, 1, 2, \ldots, \quad (2.9)$$

which is everywhere dense in $(0, a)$. The optimality criterion of rank m holds at every point of this set.

If the intersection $E_\infty = \bigcap_0^\infty E_m$ is nonempty, then, at every point $\hat{\sigma} \in E_\infty$,

$$\omega_m(\chi, \hat{\sigma})[p(t)] \equiv 0 \quad \text{on } P^{(m)} \ \forall \chi \in \overset{*}{Q}{}_a^{(1)}, \quad \forall m = 0, 1, 2, \ldots,$$

and the optimality principle cannot yield anything new at these points in comparison with the maximum principle. A typical example of points of this kind is the points in a neighborhood of which the control $\tilde{u}(t)$ is constant and concentrated at a vertex of the polyhedron U, because then $\pi_{\hat{\sigma}} = 0$. Also, interior points of the set of discontinuity points of $\tilde{u}(t)$ are of this kind. Indeed, by the formal definitions that we adopted in §1, all the Legendre forms vanish at these points. In fact, we introduced these definitions in order to obtain uniform statements for the basic Theorems 2.1 and 2.2, and formally not to exclude from consideration those discontinuity points of the control at which the Legendre representations (essentially local) are not meaningful.

We combine what we have said in the following proposition.

PROPOSITION 2.1. *For any optimal solution* (2.1), *there exists a function* $\tilde{\psi}(t) \in \Psi$ *such that for all* $m \geqslant 0$ *the solution* (2.1) *satisfies the optimality criterion of rank* m *with the function* $\tilde{\psi}(t)$ *at every point of the open set* (2.9), *which is dense everywhere in* $(0, a)$.

PROPOSITION 2.2. *The solution* (2.1) *satisfies the optimality criterion of rank* m *at a point* $\hat{\sigma}$ *if and only if the following condition holds at* $\hat{\sigma}$.

For some $l \leqslant m$, *suppose all the bilinear forms in* $p_1, p_2 \in \mathbf{R}^r$

$$\psi(\sigma)\,\mathfrak{L}_i f(\sigma)[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2], \tag{2.10}$$

are identically zero, where $\psi(t)$ *is an arbitrary function in* Ψ, σ *is an arbitrary point near* $\hat{\sigma}$, *and* $i \leqslant l - 1$. *Then the bilinear form in* $p_1, p_2 \in \mathbf{R}^r$

$$\tilde{\psi}(\hat{\sigma})\,\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] \tag{2.11}$$

vanishes identically for even l, *and for odd* l *the quadratic form in* $p \in \mathbf{R}^r$ *with symmetric matrix*

$$(-1)^{\frac{l+3}{2}}\,\tilde{\psi}(\hat{\sigma})\,\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}} p, \pi_{\hat{\sigma}} p] \tag{2.12}$$

is nonpositive on $\mathbf{R}^r$. *Here the* $\mathfrak{L}_i$ *are Legendre operators* (*see* (1.20)).

PROOF. By Propositions 1.5 and 1.3, and by (1.20), the identities (2.7) are equivalent to the identities

$$\chi\Gamma(a, \sigma)\,\mathfrak{L}_i f(\sigma)[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] \equiv 0, \quad \forall \chi \in \overset{*}{Q}{}_a^{(1)},$$

whose left-hand sides range over all the bilinear forms (2.10), since $\psi(t) = \chi\Gamma(a, t)$ is an arbitrary function in Ψ when χ ranges over the cone $\overset{*}{Q}{}_a^{(1)}$. Moreover, also by Proposition 1.3, the kernel of the form (2.8) is equal to

$$\frac{1}{2}(s-t)^l \Omega_l(\tilde{\psi}(a), \hat{\sigma})[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] = \frac{1}{2}(s-t)^l \tilde{\psi}(\hat{\sigma})\,\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2].$$

Therefore the bilinear form (2.11) is zero for even l (Proposition 1.4), and selfadjoint for odd l (see (1.17)).

Thus it remains to prove that for odd l the fact that the quadratic form (2.12) in the r-dimensional argument $p \in \mathbf{R}^r$ is nonpositive is equivalent to the fact that the integral quadratic form (2.8) is nonpositive on $P^{(l)}$.

We have

$$\frac{2\omega_l}{l!}(\tilde{\psi}(a), \hat{\sigma})[p(t)] = (-1)\int_{-1}^{1} dt \int_{-1}^{t} \frac{(t-s)^l}{l!}\tilde{\psi}(\hat{\sigma})\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}} p(t), \pi_{\hat{\sigma}} p(s)]\, ds. \quad (2.13)$$

It follows from the convolution formula

$$\int_{-1}^{t} \frac{(t-s)^l}{l!} p(s)\, ds = \int_{-1}^{t} dt_1 \int_{-1}^{t_1} dt_2 \ldots dt_l \int_{-1}^{t_l} p(s)\, ds \quad (2.14)$$

that

$$\int_{-1}^{t} dt_1 \int_{-1}^{t_1} dt_2 \ldots dt_{i-1} \int_{-1}^{t_{i-1}} p(s)\, ds \in P^{(l-i)}, \quad \forall\, p(t) \in P^{(l)}, \quad \forall\, i \leqslant l. \quad (2.15)$$

We integrate (2.13) by parts $(l+1)/2$ times. Making use of (2.14) and (2.15), we obtain

$$\frac{2\omega_l}{l!}(\tilde{\psi}(a), \hat{\sigma})[p(t)] = (-1)^{\frac{l+3}{2}} \int_{-1}^{1} \tilde{\psi}(\hat{\sigma})\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}} q(t), \pi_{\hat{\sigma}} q(t)]\, dt, \quad (2.16)$$

where

$$q(t) = \int_{-1}^{t} dt_1 \int_{-1}^{t_1} dt_2 \ldots dt_{\frac{l-1}{2}} \int_{-1}^{t_{\frac{l-1}{2}}} p(s)\, ds \in P^{(l)}.$$

Since $\tilde{\psi}(\hat{\sigma})\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2]$ is a scalar selfadjoint bilinear form in $p_1, p_2 \in \mathbf{R}^r$, it has r mutually orthogonal eigenvectors in $\mathbf{R}^r$, to which there correspond eigenvalues $\lambda_1, \ldots, \lambda_r$. We denote by $\mathfrak{M}_i$ the orthogonal projection onto the ith direction. It is easy to see that, for all $i = 1, \ldots, r$, $p_i(t) \in P^{(l)}$ can be chosen so that the image $\mathfrak{M}_i q_i(t)$ of the corresponding $q_i(t)$ (formula (2.16)) does not vanish identically on $[-1, 1]$. Therefore we obtain the equalities

$$\begin{aligned} 2\omega_l(\tilde{\psi}(a), \hat{\sigma})[p_i(t)] &= (-1)^{\frac{l+3}{2}} \int_{-1}^{1} \tilde{\psi}(\hat{\sigma})\mathfrak{L}_l f(\hat{\sigma})[\pi_{\hat{\sigma}}\mathfrak{M}_i q_i(t), \pi_{\hat{\sigma}}\mathfrak{M}_i q_i(t)]\, dt \\ &= \lambda_i \int_{-1}^{1} (\mathfrak{M}_i q_i(t))^* \mathfrak{M}_i q_i(t)\, dt, \end{aligned}$$

which imply the proposition.

Combining Propositions 2.1 and 2.2, we arrive at the following basic theorem of this paper. In its statement we employ the convention

$$\mathfrak{L}_{-1} f(\sigma, \tilde{x}(\sigma), \tilde{u}(\sigma)) = 0 \quad \forall \sigma \in (0, a).$$

THEOREM 2.2 (OPTIMALITY PRINCIPLE). *Let* (2.1) *be a time-optimal solution of the control problem* (1.1) *with fixed end-points. Then there exists a function* $\tilde{\psi}(t) \in \Psi$ *such that, if all the bilinear forms in* $p_1, p_2 \in \mathbf{R}^r$

$$\psi(\sigma)\,\mathfrak{L}_i f(\sigma, \dot{\tilde{x}}(\sigma), \tilde{u}(\sigma))[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] \quad \forall \psi(t) \in \Psi, \quad \forall \sigma \in O_{\hat{\sigma}}, \quad \forall i = -1, \dots, m-1, \tag{2.17}$$

where $O_{\hat{\sigma}}$ is a neighborhood of $\hat{\sigma}$, vanish identically for an arbitrary $\hat{\sigma} \in (0, a)$ and a given $m \geqslant 0$, then the bilinear form in $p_1, p_2 \in \mathbf{R}^r$

$$\tilde{\psi}(\hat{\sigma})\,\mathfrak{L}_m f(\hat{\sigma}, \tilde{x}(\hat{\sigma}), \tilde{u}(\hat{\sigma}))[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] \equiv 0 \quad \text{on } \mathbf{R}^r,$$

for even m, and the quadratic form in $p \in \mathbf{R}^r$

$$(-1)^{\frac{m+3}{2}}\tilde{\psi}(\hat{\sigma})\,\mathfrak{L}_m f(\hat{\sigma}, \tilde{x}(\hat{\sigma}), \tilde{u}(\hat{\sigma}))[\pi_{\hat{\sigma}} p, \pi_{\hat{\sigma}} p] \leqslant 0 \quad \forall p \in \mathbf{R}^r$$

for odd m.

We denote by E_m the open set of points $\hat{\sigma} \in (0, a)$ for which all the forms (2.17) are zero, and by D_m the open subset of E_m whose points satisfy the following additional condition: there exists a function $\hat{\psi}(t) \in \Psi$ such that the form in $p_1, p_2 \in \mathbf{R}^r$

$$\hat{\psi}(\hat{\sigma})\,\mathfrak{L}_m f(\hat{\sigma}, \tilde{x}(\hat{\sigma}), \tilde{u}(\hat{\sigma}))[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] \not\equiv 0.$$

Then the open set $\bigcup_{i \leqslant m-1} D_i \cup E_m$ ($D_{-1} = \varnothing$) is everywhere dense in $(0, a)$ for any $m \geqslant 0$.

The family Ψ consists of all nonzero solutions of (2.2) which satisfy (2.3). An explicit expression for the Legendre operators $\mathfrak{L}_m$ is given by (1.20).

The statement of the optimality principle given here combines the Pontrjagin maximum principle (the family Ψ is nonempty) with a second-order necessary condition for optimality.

If $m = 0$, then the forms (2.17) are zero by definition ($\mathfrak{L}_{-1} = 0$), and we obtain a necessary condition for optimality for all $\hat{\sigma} \in (0, a)$:

$$\tilde{\psi}(\hat{\sigma})\,\mathfrak{L}_0 f(\hat{\sigma}, \tilde{x}(\hat{\sigma}), \tilde{u}(\hat{\sigma}))[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2] \equiv 0.$$

As can be seen from (1.20), this condition is equivalent to the condition that the form $\tilde{\psi}(\hat{\sigma})F_1(\hat{\sigma}, \hat{\sigma})[\pi_{\hat{\sigma}} p_1, \pi_{\hat{\sigma}} p_2]$ be selfadjoint. The latter condition is automatically satisfied when the control u is a scalar ($r = 1$). Therefore in this case we can apply the optimality criterion beginning with $m = 1$ at every point $\hat{\sigma}$.

§3. Proof of Theorem 2.1

Let there be given a point $\hat{\sigma} \in (0, a)$ and a function $\hat{p}(t) \in P^{(m)}$. For brevity, we denote by $\delta z^{(m)}(\hat{\sigma}, \hat{p}(t))$ the $(m + 1)$th coefficient $L_m(\hat{\sigma})[\hat{p}(t)]$ in the corresponding Legendre representation (1.13), if there exists a neighborhood $O_{\hat{\sigma}}$ such that

$$L_i(\sigma)[\pi_{\hat{\sigma}} \hat{p}(t)] \in N_a \quad \forall \sigma \in O_{\hat{\sigma}}, \quad \forall i \leqslant m - 1,$$

where the subspace N_a is given by (1.14). In the contrary case, we set $\delta z^{(m)}(\hat{\sigma}, \hat{p}(t)) = 0$.

We denote by T the union of the vectors $\delta z^{(m)}(\hat{\sigma}, \hat{p}(t))$ over all possible $\hat{\sigma}$ and $\hat{p}(t) \in P^{(m)}$ with $m = 0, 1, 2, \dots$.

The cone spanned from the origin by the convex hull conv $(Q_a^{(1)} \cup T)$ of the union of the first-order cone $Q_a^{(1)}$ and the set T will be called the *second-order cone* $Q_a^{(2)} \subset \mathbf{R}^n$ of the trajectory $\tilde{x}(t)$ at the end-point $\tilde{x}(a)$.

Theorem 2.1 will be proved if we show that the convex cone $Q_a^{(2)}$ does not coincide with the entire space $\mathbf{R}^n$:

$$Q_a^{(2)} \neq \mathbf{R}^n. \tag{3.1}$$

Indeed, let χ be a vector orthogonal to a supporting hyperplane of the cone $Q_a^{(2)}$, directed away from $Q_a^{(2)}$. Obviously $\chi \in \overset{*}{Q}{}_a^{(a)}$. If $\tilde{\psi}(t)$, $0 \leqslant t \leqslant a$, is the solution of the differential equation

$$\dot{\psi} = -\psi f_x(t, \tilde{x}(t), \tilde{u}(t)),$$

satisfying the boundary condition $\tilde{\psi}(a) = \chi$, then

$$\chi \delta z^{(m)}(\hat{\sigma}, \hat{p}(t)) = \omega_m(\tilde{\psi}(a), \hat{\sigma})[\hat{p}(t)] \leqslant 0,$$

for all $m = 0, 1, 2, \ldots$, i.e. the assertion of Theorem 2.1.

Turning to the proof of (3.1), we note, first of all, that, if $\delta z^{(m)}(\hat{\sigma}, \hat{p}(t)) \neq 0$, then the equality $\pi_\sigma \pi_{\hat{\sigma}} = \pi_{\hat{\sigma}}$ $\forall \sigma \in O_{\hat{\sigma}}$ (which holds for a sufficiently small neighborhood $O_{\hat{\sigma}}$) implies

$$\delta z^{(m)}(\sigma, \pi_{\hat{\sigma}} \hat{p}(t)) = L_m(\sigma)[\pi_{\hat{\sigma}} \hat{p}(t)] \qquad \forall \sigma \in O_{\hat{\sigma}}.$$

Further, it is clear that for fixed $\hat{\sigma}$ and $\hat{p}(t)$ we have

$$\delta z^{(m)}(\sigma, \pi_{\hat{\sigma}} \hat{p}(t)) = L_m(\sigma)[\pi_{\hat{\sigma}} \hat{p}(t)] \to L_m(\hat{\sigma})[\pi_{\hat{\sigma}} \hat{p}(t)] = L_m(\hat{\sigma})[\hat{p}(t)] = \delta z^{(m)}(\hat{\sigma}, \hat{p}(t))$$

as $\sigma \to \hat{\sigma}$.

Therefore, if we are given an arbitrary number of nonzero vectors from $\delta z^{(m_j)}(\hat{\sigma}_j, p_j(t))$, $p_j(t) \in P^{(m_j)}$, $j = 1, \ldots, l$, from T, we can transfer the points $\hat{\sigma}_1, \ldots, \hat{\sigma}_l$ into distinct points $\sigma_1, \ldots, \sigma_l$ by an arbitrarily small shift, and thus obtain the vectors $\delta z^{(m_j)}(\sigma_j \pi_{\hat{\sigma}_j} p_j(t))$, $j = 1, \ldots, l$, from T which differ arbitrarily little from the corresponding initial vectors.

Let us show that the equality $Q_a^{(2)} = \mathbf{R}^n$ leads to a contradiction.

If the equality holds, then there exist $1 + k + l$ nonzero vectors

$$\delta x_0, \ldots, \delta x_k \in Q_a^{(1)}, \quad \delta z_1 = \delta z^{(m_1)}(\sigma_1, p_1(t)), \ldots, \delta z_l = \delta z^{(m_l)}(\sigma_l, p_l(t)), \quad p_j(t) \in P^{(m_j)},$$

such that the origin of $\mathbf{R}^n$ is an interior point of the convex hull of these points,

$$\operatorname{conv}[\delta x_0, \ldots, \delta x_k; \delta z_1, \ldots, \delta z_l] \subset \mathbf{R}^n. \tag{3.2}$$

Moreover, by what we have said, we can assume that the points $\sigma_1, \ldots, \sigma_l$ are distinct.

We shall now apply the following basic lemma, whose proof we postpone to the end of this section in order not to interrupt the presentation.

LEMMA. *Let there be given vectors $\delta x_0, \ldots, \delta x_k \in Q_a^{(1)}$, l distinct points $\sigma_1, \ldots, \sigma_l$, and the same number of functions $p_j(t) \in P^{(m_j)}$ satisfying the following condition: for each $j = 1, \ldots, l$, the first m_j coefficients $L_i(\sigma_j)[p_j(t)]$, $i \leqslant m_j - 1$, of the corresponding Legendre representation lie in the subspace N_a:*

$$L_i(\sigma_j)[p_j(t)] \in N_a, \quad \forall i \leqslant m_j - 1. \tag{3.3}$$

Denote by Λ the $(k + 1)$-dimensional simplex

$$\Lambda = \{\lambda = (\lambda_0, \ldots, \lambda_{k+l}) \mid \lambda_i \geqslant 0, \lambda_0 + \ldots + \lambda_{k+l} = 1\}.$$

Then there exists a family of perturbations $\delta u(t, \lambda, \varepsilon)$ of the control $\tilde{u}(t)$ which is defined for $\lambda \in \Lambda$ and all sufficiently small $\varepsilon > 0$, and is such that the differential equation

$$\dot{x} = f(t, x, \tilde{u}(t) + \delta u(t; \lambda, \varepsilon)) = g(t, x) + G(t, x)\tilde{u}(t) + G(t, x)\delta u(t; \lambda, \varepsilon)$$

has a solution

$$x(t; \lambda, \varepsilon), \quad 0 \leqslant t \leqslant a; \quad x(0; \lambda, \varepsilon) \equiv \tilde{x}(0),$$

which depends continuously on $(t, \lambda) \in [0, a] \times \Lambda$ *for a fixed* ε *and satisfies the condition*

$$\left| x(a; \lambda, \varepsilon) - \tilde{x}(a) - \varepsilon \left(\sum_{i=0}^{k} \lambda_i \delta x_i + \sum_{j=1}^{l} \lambda_{j+k+1} L_{m_j}(\sigma_j)[p_j(t)] \right) \right| \leqslant o(\varepsilon), \forall \lambda \in \Lambda, \quad (3.4)$$

where $o(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$.

Since $\delta z_j \neq 0$, we have $\delta z_j = L_{m_j}(\sigma_j)[p_j(t)]$, $j = 1, \ldots, l$, and (3.4) can be expressed in the form

$$\left| x(a; \lambda, \varepsilon) - \tilde{x}(a) - \varepsilon \left(\sum_{i=0}^{k} \lambda_i \delta x_i + \sum_{j=1}^{l} \lambda_{j+k+1} \delta z_j \right) \right| \leqslant o(\varepsilon) \qquad \forall \lambda \in \Lambda. \qquad (3.5)$$

We define the family of continuous mappings $Y(\lambda, \varepsilon)$; $\Lambda \to \mathbf{R}^n$, $\varepsilon > 0$, by the formulas

$$Y(\lambda; \varepsilon) = \frac{x(a; \lambda, \varepsilon) - \tilde{x}(a)}{\varepsilon}, \quad \varepsilon > 0,$$

$$Y(\lambda; 0) = \sum_{i=0}^{k} \lambda_i \delta x_i + \sum_{j=1}^{l} \lambda_{j+k+1} \delta z_j.$$

Elementary geometric considerations allow us to choose in Λ an $(n-1)$-dimensional polyhedral sphere S^{n-1} (composed of the sides of the simplex Λ) such that the mapping $Y(\lambda; 0)$ is a (piecewise linear) homeomorphism of S^{n-1} onto the boundary of the convex set (3.2). Since the origin of $\mathbf{R}^n$ is an interior point of the set (3.2), the image of S^{n-1} under the mapping $Y(\lambda; 0)$ touches the origin of $\mathbf{R}^n$. Therefore, by (3.5), the image of the sphere also touches the origin under any mapping $Y(\lambda; \varepsilon)$ with $\varepsilon > 0$ sufficiently small.

We identify the lower base $\{0\} \times S^{n-1}$ of the cylinder $[0, a] \times S^{n-1}$ with a single point. We denote by C^n the n-dimensional ball thus obtained, and we consider the family of continuous mappings

$$(t, \lambda) \longmapsto X(t; \lambda, \varepsilon) = \frac{x(t; \lambda, \varepsilon) - \tilde{x}(t)}{\varepsilon}, \quad (t, \lambda) \in [0, a] \times S^{n-1}, \quad \varepsilon > 0,$$

of the cylinder $[0, a] \times S^{n-1}$ into $\mathbf{R}^n$.

By virtue of the condition $x(0; \lambda, \varepsilon) \equiv \tilde{x}(0)$, $\forall \lambda \in \Lambda$, $\varepsilon > 0$, this family can be viewed as a family of continuous mappings of C^n into $\mathbf{R}^n$. Since the mapping $X(t; \lambda, \varepsilon)$ coincides on the boundary $\partial C^n = \{a\} \times S^{n-1}$ of the ball C^n with the mapping $Y(\lambda; \varepsilon)$, $\varepsilon > 0$, the image of the boundary ∂C^n under the mapping $X(t; \lambda, \varepsilon)$ touches the origin for all sufficiently small $\varepsilon > 0$. Therefore the image of the entire ball C^n under the mapping $X(t; \lambda, \varepsilon)$ covers the origin of $\mathbf{R}^n$ for all sufficiently small $\varepsilon > 0$. Since the image of the boundary cannot contain the origin, for any sufficiently small $\varepsilon > 0$ there exist a $t_\varepsilon < a$ and a $\lambda_\varepsilon \in S^{n-1} \subset \Lambda$ such that

$$X(t_\varepsilon; \lambda_\varepsilon, \varepsilon) = \frac{x(t_\varepsilon; \lambda_\varepsilon, \varepsilon) - \tilde{x}(a)}{\varepsilon} = 0,$$

or $x(t_\varepsilon;\, \lambda_\varepsilon,\, \varepsilon) = \tilde{x}(a)$. This equality contradicts the assumption of the optimality of the solution (2.1).

PROOF OF THE LEMMA. Let $\alpha(\varepsilon)$ be an arbitrary positive function of $\varepsilon > 0$ which tends to zero as $\varepsilon \to 0$. For every function $p_j(t)$, we take the m_jth order packet

$$\alpha(\varepsilon)\, \pi_{\sigma_j} p_j \left(\frac{t - \sigma_j}{\varepsilon^{\frac{2}{2+m_j}}} \right) \tag{3.6}$$

and prove the existence of a family of functions $\delta v_1^{(j)}(t, \varepsilon)$, uniformly bounded in absolute value, which is such that the end-point of the first variation of the trajectory $\tilde{x}(t)$ evaluated on this packet and added to the end-point of the first variation evaluated on the family of perturbations $\alpha(\varepsilon)\varepsilon^2 \delta v_1^{(j)}(t, \varepsilon)$ is zero:

$$\begin{aligned} &\delta_1 x \left(a;\, \alpha(\varepsilon)\, \pi_{\sigma_j} p_j \left(\frac{t - \sigma_j}{\varepsilon^{\frac{2}{2+m_j}}} \right) \right) + \delta_1 x \left(a;\, \alpha(\varepsilon)\, \varepsilon^2 \delta v_1^{(j)}(t, \varepsilon) \right) \\ &\quad = \alpha(\varepsilon)\, \varepsilon^{\frac{2}{2+m_j}} \int_{-1}^{1} A\left(\sigma_j + t\varepsilon^{\frac{2}{2+m_j}} \right) \pi_{\sigma_j} p_j(t)\, dt \\ &\quad + \alpha(\varepsilon)\, \varepsilon^2 \int_0^a A(t)\, \delta v_1^{(j)}(t, \varepsilon)\, dt \equiv 0 \quad \forall j = 1\ \ldots, l, \end{aligned} \tag{3.7}$$

where

$$A(t) = \Gamma(a, t) f_u(t) = \Gamma(a, t)\, G(t, \tilde{x}(t)).$$

Expanding the kernel A in the first integral in powers of $t\varepsilon^{2/(2+m_j)}$, and integrating term by term with respect to t, by the condition $p_j(t) \in P^{(m_j)}$ we can write (see (1.7))

$$\alpha(\varepsilon)\, \varepsilon^{\frac{2}{2+m_j}} \int_{-1}^{1} A\left(\sigma_j + t\varepsilon^{\frac{2}{2+m_j}} \right) \pi_{\sigma_j} p_j(t)\, dt = \alpha(\varepsilon)\, \varepsilon^2 \delta y^{(j)}(\varepsilon),$$

where the family $\delta y^{(j)}(\varepsilon)$ is uniformly bounded in absolute value and

$$\pm\, \delta y^{(j)}(\varepsilon) \in Q_a^{(1)}, \tag{3.8}$$

This is so because

$$\pm\, \alpha(\varepsilon)\, \pi_{\sigma_j} p_j \left(\frac{t - \sigma_j}{\varepsilon^{\frac{2}{2+m_j}}} \right)$$

is the family of perturbations of the control $\tilde{u}(t)$.

The existence of a family $\delta v_1^{(j)}(t, \varepsilon)$ follows easily. Indeed, let us take an arbitrary simplex of maximum dimension contained in $Q_a^{(1)}$ and with center at the origin:

$$\left\{ \delta y \;\middle|\; \delta y = \sum_{i=0}^{r} \nu_i \delta y_i,\ \nu_i \geqslant 0,\ \nu_0 + \ldots + \nu_r = 1 \right\}.$$

The condition (3.8) and the uniform boundedness of $\delta y^{(j)}(\varepsilon)$ imply that there exist bounded nonnegative functions $\nu_i^{(j)}(\varepsilon)$ such that

$$-\delta y^{(j)}(\varepsilon) = \sum_{i=0}^{r} \nu_i^{(j)}(\varepsilon)\, \delta y_i.$$

Since $\delta y_i \in Q_a^{(1)}$, there exist perturbations $\delta w_i(t)$ satisfying the equalities

$$\delta y_i = \int_0^a A(t)\,\delta w_i(t)\,dt,$$

Therefore we can set

$$\delta v_1^{(j)}(t, \varepsilon) = \sum_{i=0}^{r} \nu_i^{(j)}(\varepsilon)\,\delta w_i(t).$$

We now construct a uniformly bounded family of functions $\delta v_2^{(j)}(t, \varepsilon)$ which satisfy the condition

$$\left| \delta_1 x\left(a; \alpha(\varepsilon)\,\varepsilon^2 \delta v_2^{(j)}(t, \varepsilon)\right) + \alpha^2(\varepsilon)\,\varepsilon^{\frac{4}{2+m_j}} \sum_{i=0}^{m_j-1} \varepsilon^{\frac{2i}{2+m_j}} \frac{1}{i!} L_i(\sigma_j)\,[p_j(t)] \right| = o(\alpha^2(\varepsilon)\,\varepsilon^2). \quad (3.9)$$

This condition asserts that the end-point of the first variation evaluated on the family of perturbations $\alpha(\varepsilon)\varepsilon^2\delta v_2^{(j)}(t, \varepsilon)$, added to the sum of the first m_j terms of the Legendre representation of the end-point of the second variation on the packet (3.6), is a quantity of higher order than $\alpha^2(\varepsilon)\varepsilon^2$.

In order to accomplish the construction, we recall that the subspace N_a (see (1.14)) is contained in the closure of the cone $Q_a^{(1)}$. Therefore the conditions (3.3) imply the existence of a nonnegative function $\varphi_j(\varepsilon)$ (possibly increasing very rapidly as $\varepsilon \to 0$) and of a family of perturbations $\delta w^j(t, \varepsilon)$ such that the sum

$$\delta_1 x\left(a; \varphi_j(\varepsilon)\,\delta w^{(j)}(t, \varepsilon)\right) + \varepsilon^{\frac{4}{2+m_j}} \sum_{i=0}^{m_j-1} \varepsilon^{\frac{2i}{2+m_j}} \frac{1}{i!} L_i(\sigma_j)\,[p_j(t)]$$

tends to zero arbitrarily fast as $\varepsilon \to 0$. We need only the estimate

$$\left| \delta_1 x\left(a; \varphi_j(\varepsilon)\,\delta w^{(j)}(t, \varepsilon)\right) + \varepsilon^{\frac{4}{2+m_j}} \sum_{i=0}^{m_j-1} \varepsilon^{\frac{2i}{2+m_j}} \frac{1}{i!} L_i(\sigma_j)\,[p_j(t)] \right| = o(\varepsilon^2). \quad (3.10)$$

We choose a function $\alpha(\varepsilon)$ which tends to zero as $\varepsilon \to 0$ so fast that, for each $j = 1, \ldots, l$,

$$\alpha(\varepsilon)\,\varphi_j(\varepsilon) \leqslant \frac{\varepsilon^2}{1 + |\delta w^{(j)}(t, \varepsilon)|}, \quad \alpha(\varepsilon) \leqslant \varepsilon^3. \quad (3.11)$$

Moreover, we assume that $\alpha(\varepsilon)$ is monotone.

It is easy to see that we can set

$$\delta v_2^{(j)}(t, \varepsilon) = \frac{\alpha(\varepsilon)\,\varphi_j(\varepsilon)}{\varepsilon^2}\,\delta w^{(j)}(t, \varepsilon),$$

since the first of the inequalities (3.11) yields

$$|\delta v_2^{(j)}(t, \varepsilon)| \leqslant \frac{|\delta w^{(j)}(t, \varepsilon)|}{1 + |\delta w^{(j)}(t, \varepsilon)|} < 1,$$

and (3.9) follows from (3.10).

The families of perturbations

$$\delta u^{(j)}(t;\lambda_j,\varepsilon)=\alpha(\varepsilon)\sqrt{\lambda_j m_j!}\,\pi_{\sigma_j}p_j\left(\frac{t-\sigma_j}{\varepsilon^{\frac{2}{2+m_j}}}\right)$$

$$+\alpha(\varepsilon)\,\varepsilon^2\sqrt{\lambda_j}\,\delta v_1^{(j)}(t,\varepsilon)+\alpha(\varepsilon)\,\varepsilon^2\lambda_j\delta v_2^{(j)}(t,\varepsilon),\quad 0\leqslant\lambda_j\leqslant 1,\quad j=1,\ldots,l,$$

have a number of properties which are important for the proof of the lemma and which can be verified directly. We shall now enumerate these properties.

The end-point of the second variation evaluated on the family $\delta u^{(j)}(t;\lambda_j,\varepsilon)$ differs from the end-point of the second variation evaluated on the packet

$$\alpha(\varepsilon)\sqrt{\lambda_j m_j!}\,\pi_{\sigma_j}p_j\left(\frac{t-\sigma_j}{\varepsilon^{\frac{2}{2+m_j}}}\right)$$

by $o(\alpha^2(\varepsilon)\varepsilon^2)$. Therefore

$$\delta_2 x\,(a;u^{(j)}(t;\lambda_j,\varepsilon))=\delta_2 x\left(a;\alpha(\varepsilon)\sqrt{\lambda_j m_j!}\,\pi_{\sigma_j}p_j\left(\frac{t-\sigma_j}{\varepsilon^{\frac{2}{2+m_j}}}\right)\right)+o\,(\alpha^2(\varepsilon)\,\varepsilon^2)$$

$$=\alpha^2(\varepsilon)\,\varepsilon^{\frac{4}{2+m_j}}\lambda_j m_j!\sum_{i=0}^{m_j-1}\varepsilon^{\frac{2i}{2+m_j}}\frac{1}{i!}L_i(\sigma_j)\,[p_j(t)]+\alpha^2(\varepsilon)\,\varepsilon^2\lambda_j L_{m_j}(\sigma_j)\,[p_j(t)]+o\,(\alpha^2(\varepsilon)\,\varepsilon^2).$$

Thus (3.7) and (3.9) imply that the sum of the end-points of the first and second variations evaluated on $\delta u^{(j)}(t;\lambda_j,\varepsilon)$ differs from the vector $\alpha^2(\varepsilon)\varepsilon^2\lambda_j L_{m_j}(\sigma_j)[p_j(t)]$ by a value of order $o(\alpha^2(\varepsilon)\varepsilon^2)$:

$$|\delta_1 x\,(a;\delta u^{(j)}(t;\lambda_j,\varepsilon))+\delta_2 x\,(a;\delta u^{(j)}(t;\lambda_j,\varepsilon))-\alpha^2(\varepsilon)\,\varepsilon^2\lambda_j L_{m_j}(\sigma_j)\,[p_j(t)]|\leqslant o\,(\alpha^2(\varepsilon)\,\varepsilon^2).\quad(3.12)$$

Next, making use of the condition $\sigma_{j'}\neq\sigma_{j''}$ for $j'\neq j''$ and the fact that

$$\delta_1 x\left(\vartheta;\alpha(\varepsilon)\,\pi_{\sigma_j}p_j\left(\frac{t-\sigma_j}{\varepsilon^{\frac{2}{2+m_j}}}\right)\right)=O\,(\alpha(\varepsilon)\,\varepsilon^2)\qquad\forall\vartheta\neq\sigma_j,$$

we can also directly verify the equality

$$\delta_2 x\left(a;\sum_{j=1}^{l}\delta u^{(j)}(t;\lambda_j,\varepsilon)\right)=\sum_{j=1}^{l}\delta_2 x\,(a;\delta u^{(j)}(t;\lambda_j,\varepsilon))+o\,(\alpha^2(\varepsilon)\,\varepsilon^2).$$

Hence, making use of (3.12), we obtain

$$\begin{aligned}&\delta_1 x\left(a;\sum_{j=1}^{l}\delta u^{(j)}(t;\lambda_j,\varepsilon)\right)+\delta_2 x\left(a;\sum_{j=1}^{l}\delta u^{(j)}(t;\lambda_j,\varepsilon)\right)\\&=\sum_{j=1}^{l}[\delta_1 x\,(a;\delta u^{(j)}(t;\lambda_j,\varepsilon))+\delta_2 x\,(a;\delta u^{(j)}(t;\lambda_j,\varepsilon))]+o\,(\alpha^2(\varepsilon)\,\varepsilon^2)\qquad(3.13)\\&=\alpha^2(\varepsilon)\,\varepsilon^2\sum_{j=1}^{l}\lambda_j L_{m_j}(\sigma_j)\,[p_j(t)]+o\,(\alpha^2(\varepsilon)\,\varepsilon^2).\end{aligned}$$

We define the family of perturbations

$$\delta\hat{u}\,(t;\lambda,\varepsilon)=\alpha^2(\varepsilon)\,\varepsilon^2\sum_{i=0}^{k}\mu_i\delta u^{(i)}(t)+\sum_{j=1}^{l}\delta u^{(j)}(t;\lambda_j,\varepsilon),$$

where $\lambda = (\mu_0, \ldots, \mu_k; \lambda_1, \ldots, \lambda_l)$, $0 \leqslant \mu_i \leqslant 1$, $0 \leqslant \lambda_j \leqslant 1$, and the perturbations $\delta u^{(i)}(t)$ are chosen so that

$$\delta_1 x(a; \delta u^{(i)}(t)) = \delta x_i, \quad i = 0, \ldots, k.$$

Then (3.13) implies

$$\begin{gathered}\delta_1 x(a; \delta\hat{u}(t; \lambda, \varepsilon)) + \delta_2 x(a; \delta\hat{u}(t; \lambda, \varepsilon)) \\ = \alpha^2(\varepsilon)\varepsilon^2 \left\{\sum_{i=0}^{k} \mu_i \delta x_i + \sum_{j=1}^{l} \lambda_j L_{m_j}(\sigma_j)[p_j(t)]\right\} + o(\alpha^2(\varepsilon)\varepsilon^2).\end{gathered} \tag{3.14}$$

We now note that the functions $g(t, x)$ and $G(t, x)$ are three times continuously differentiable, and that the family of perturbations $\delta\hat{u}(t, \lambda, \varepsilon)$ is continuous in λ, uniformly with respect to t and ε, and satisfies the estimate

$$\max_{0 \leqslant t \leqslant a} |\delta\hat{u}(t; \lambda, \varepsilon)| = R(\lambda, \varepsilon) \leqslant \text{Const} \cdot \alpha(\varepsilon).$$

Therefore, by a standard theorem on the dependence of a solution of a differential equation on the right-hand side, for all $\varepsilon > 0$ sufficiently small and all λ with $0 \leqslant \mu_i \leqslant 1$, $0 \leqslant \lambda_j \leqslant 1$, the perturbed equation

$$x = g(t, x) + G(t, x)\tilde{u}(t) + G(t, x)\,\delta\hat{u}(t; \lambda, \varepsilon)$$

has a solution

$$\hat{x}(t; \lambda, \varepsilon), \quad 0 \leqslant t \leqslant a, \quad \hat{x}(0; \lambda, \varepsilon) \equiv \tilde{x}(0),$$

which depends continuously on the point

$$(t, \lambda) \in \{(t, \lambda) \mid 0 \leqslant t \leqslant a,\ 0 \leqslant \mu_i \leqslant 1,\ 0 \leqslant \lambda_j \leqslant 1\}$$

for every fixed $\varepsilon > 0$, and satisfies the estimate

$$\begin{gathered}\max_{0 \leqslant t \leqslant a} |\hat{x}(t; \lambda, \varepsilon) - \tilde{x}(t) - \delta_1 x(t; \delta\hat{u}(\tau; \lambda, \varepsilon)) \\ - \delta_2 x(t; \delta\hat{u}(\tau; \lambda, \varepsilon))| \leqslant O(R^3(\lambda, \varepsilon)) = O(\alpha^3(\varepsilon)).\end{gathered}$$

Hence, by (3.14) and the second of the inequalities (3.11), for $t = a$ we obtain the final estimate

$$\left|\hat{x}(a; \lambda, \varepsilon) - \tilde{x}(a) - \alpha^2(\varepsilon)\varepsilon^2 \left\{\sum_{i=0}^{k} \mu_i \delta x_i + \sum_{j=1}^{l} \lambda_j L_{m_j}(\sigma_j)[p_j(t)]\right\}\right| \leqslant o(\alpha^2(\varepsilon)\varepsilon^2),$$

which is equivalent to (3.14). Indeed, introducing the new parameter $\varepsilon' = \alpha^2(\varepsilon)\varepsilon^2$ and solving this equation for ε (by assumption, the function $\alpha(\varepsilon)$ is monotone), $\varepsilon = \gamma(\varepsilon')$, we define the required perturbation by the equality $\delta u(t; \lambda, \varepsilon') = \delta\hat{u}(t; \lambda, \gamma(\varepsilon'))$. The corresponding trajectory is $x(t; \lambda, \varepsilon') = \hat{x}(t; \lambda, \gamma(\varepsilon'))$.

REMARK. We say that the necessary condition for optimality just proved is a *second-order optimality principle*, since it has been obtained as a result of studying the Legendre representation of the end-point of the second variation of a trajectory. A similar method applied to the end-point of the first variation leads at once to necessary conditions for optimality that are direct consequences of the maximum principle. An attempt to obtain necessary conditions of higher order by the method presented, e.g., by the decomposition of the end-point of the third variation of the trajectory on packets, does not lead to a successful result without additional assumptions.

§4. The expression of the Legendre operator in terms of Lie brackets(3)

This section can be viewed as a direct continuation of §1. Here, we shall express the operator $\mathfrak{L}_m$ in terms of Lie brackets (formula (4.11)), whereas in (1.20) this operator was expressed in terms of the fundamental matrix $\Gamma(t, \tau)$. Naturally, both formulas are equivalent, although it is difficult to say which one will turn out to be more convenient for computations.

In order to simplify the formulas, we assume that we are considering the solution

$$\tilde{u}(t) \equiv 0, \quad \tilde{x}(t), \quad 0 \leqslant t \leqslant a,$$

of the autonomous equation

$$\dot{x} = f(x, u) = g(x) + G(x)u, \tag{4.1}$$

so that $\dot{\tilde{x}}(t) = g(\tilde{x}(t))$. The corresponding equation for $\psi(t)$ has the form

$$\dot{\psi} = -\psi g_x(\tilde{x}(t)). \tag{4.2}$$

Further, we assume that $U = \mathbf{R}^r$. This saves us the necessity of introducing the corresponding projection operator π_σ. The case of an arbitrary control $\tilde{u}(t)$ reduces to the case $\tilde{u}(t) \equiv 0$ by a standard method. It is sufficient to add the scalar equation $dt/d\tau = 1$ to (4.1), assuming that t is an additional phase coordinate and τ is the new time.

We begin with several commonly adopted definitions.

Let $v(x)$, $x \in \mathbf{R}^n$, be an n-dimensional, infinitely differentiable column-vector. It can be viewed as an operator acting on the set of all infinitely differentiable vector-valued functions $g(x)$ (of an arbitrary, given dimension) by the formula

$$v \circ q(x) = \frac{\partial q(x)}{\partial x} v(x).$$

This operator is said to be the *differentiation of* $q(x)$ *by virtue of the equation* $\dot{x} = v(x)$, or the *field* of the function $v(x)$. The function $v(x)$ can be reconstructed by means of the operator v according to the formula $v(x) = v \circ x$. The value of the function $v(x)$ at the point $\tilde{x}(\sigma)$ will also be expressed as $v \circ \tilde{x}(\sigma)$.

The successive application of two fields v and w to $q(x)$ yields an operation $v \circ w$ in the set of all $q(x)$ (of a given dimension), which, generally speaking, is not a field:

$$v \circ w \circ q(x) \neq \frac{\partial q(x)}{\partial x}(v \circ w \circ x).$$

However, as can easily be verified by direct computation, the operator

$$[v, w] = v \circ w - w \circ v,$$

which is called the *Lie bracket of the fields* v *and* w, is always a field:

$$[v, w] \circ q(x) = \frac{\partial q(x)}{\partial x}[v, w] \circ x \quad \forall q(x).$$

The set of all n-dimensional fields $v(x)$ will be denoted by $\mathfrak{B}$.

We introduce the operator ad v, acting in $\mathfrak{B}$ (and depending on the choice of $v \in \mathfrak{B}$)

(3) This section was added to the initial text of the paper after the authors became acquainted with the work of A. J. Krener [4], from which they took the idea of employing the notion of field for the corresponding calculations.

by the formula

$$(\mathrm{ad}\, v)\, w = [v, w].$$

A basic property of Lie brackets, besides the obvious property of *skew-symmetry*

$$[v, w] = -[w, v],$$

is the *Jacobi identity*

$$(\mathrm{ad}\, v)\, [v_1, v_2] = [(\mathrm{ad}\, v)\, v_1, v_2] + [v_1, (\mathrm{ad}\, v)\, v_2],$$

which can be verified directly. Hence, denoting the ith power of the operator ad v by $\mathrm{ad}^i\, v$, we obtain the (Leibniz) formula

$$(\mathrm{ad}^i v)\, [v_1, v_2] = \sum_{j_1+j_2=i} \frac{i!}{j_1!\, j_2!} [(\mathrm{ad}^{j_1} v)\, v_1, (\mathrm{ad}^{j_2} v)\, v_2]$$

by induction.

We shall consider formal power series of operators; in particular, exponential series

$$e^v = I + v + \frac{v^2}{2!} + \ldots, \qquad e^{\mathrm{ad}\, v} = I + \mathrm{ad}\, v + \frac{\mathrm{ad}^2 v}{2!} + \ldots .$$

Let us prove the simple formula

$$e^v \circ w \circ e^{-v} = e^{\mathrm{ad}\, v} w. \tag{4.3}$$

We introduce the operator-valued function of the scalar argument t

$$\varphi(t) = e^{tv} \circ w \circ e^{-tv} = \varphi(0) + t \frac{d\varphi(0)}{dt} + \frac{t^2}{2!} \frac{d^2\varphi(0)}{dt^2} + \ldots .$$

We have

$$\frac{d\varphi(t)}{dt} = v \circ e^{tv} \circ w \circ e^{-tv} - e^{tv} \circ w \circ v \circ e^{-tv} = e^{tv} \circ (v \circ w - w \circ v) \circ e^{-tv} = e^{tv} \circ (\mathrm{ad}\, v)\, w \circ e^{-tv},$$

and, by induction,

$$\frac{d^i \varphi(t)}{dt^i} = e^{tv} \circ (\mathrm{ad}^i v)\, w \circ e^{-tv}.$$

Therefore we obtain the equality

$$\frac{d^i \varphi(0)}{dt^i} = (\mathrm{ad}^i v)\, w,$$

which implies the formula being proved.

Finally, for any solution $\psi(t)$ of (4.1) and any n-dimensional infinitely differentiable column vector $q(x)$, an obvious induction yields

$$\frac{d^i}{dt^i} \psi(t)\, q(\tilde{x}(t)) = \psi(t)\, (\mathrm{ad}^i g)\, q \circ \tilde{x}(t). \tag{4.4}$$

We return to our basic problem of evaluating $\mathfrak{L}_m f$.

Let $p(t)$, $t \in \mathbf{R}$, be an arbitrary, n-dimensional square-summable column vector which vanishes outside the interval $[-1, 0]$, and let σ be a point of $(0, a)$. We consider the family of perturbations $p((t-\sigma)/\varepsilon)$ of the control $\tilde{u}(t) \equiv 0$ and the corresponding family of perturbed equations

$$x = g(x) + Gp((t-\sigma)/\varepsilon).$$

We denote by $x_\varepsilon(t)$, with $0 \leqslant t \leqslant a$, the solution of the perturbed equation with the initial condition $x_\varepsilon(0) \equiv \tilde{x}(0)$. We have $x_\varepsilon(\sigma - \varepsilon) \equiv \tilde{x}(\sigma - \varepsilon)$.

Let $\delta_2 x(t, \varepsilon)$, $0 \leqslant t \leqslant a$, be the second variation of the trajectory $x_\varepsilon(t)$, $0 \leqslant t \leqslant a$, corresponding to the perturbation $p((t - \sigma)/\varepsilon)$. We shall now concern ourselves with the Legendre representation of the end-point $\delta_2 x(a; \varepsilon)$ by a method different from that in §1, and we shall obtain an expression different in form from the corresponding expression (1.13).

First we find an asymptotic expansion in powers of ε of the value of the trajectory $x_\varepsilon(t)$ at the point $t = \sigma$.

We denote by f_τ, with $-1 \leqslant \tau \leqslant 0$, the family of fields which depends on the parameter τ and is determined by the family of the n-dimensional functions of x

$$f_\tau(x) = g(x) + G(x)\,p(\tau), \quad -1 \leqslant \tau \leqslant 0.$$

Introducing the "fast" time $\tau = (t - \sigma)/\varepsilon$, we express the differential equation for $x_\varepsilon(\sigma + \varepsilon\tau)$ in the form

$$\frac{dx_\varepsilon}{d\tau} = \varepsilon\,(g(x_\varepsilon) + G(x_\varepsilon)\,p(\tau)), \quad -1 \leqslant \tau \leqslant 0, \quad x_\varepsilon(\sigma - \varepsilon) \equiv \tilde{x}(\sigma - \varepsilon).$$

For an arbitrary function $q(x)$ and an arbitrary integer m we have

$$\begin{aligned}
q(x_\varepsilon(\sigma)) &= q(x_\varepsilon(\sigma-\varepsilon)) + \varepsilon \int_{-1}^{0} \frac{\partial q(x_\varepsilon(\sigma+\varepsilon\tau))}{\partial x}\{g(x_\varepsilon(\sigma+\varepsilon\tau)) + G(x_\varepsilon(\sigma+\varepsilon\tau))p(\tau)\}\,d\tau \\
&= q(x_\varepsilon(\sigma-\varepsilon)) + \varepsilon \int_{-1}^{0} d\tau_1 f_{\tau_1} \circ q(x_\varepsilon(\sigma+\varepsilon\tau_1)) = q(x_\varepsilon(\sigma-\varepsilon)) \\
&\quad + \varepsilon \int_{-1}^{0} d\tau_1 f_{\tau_1} \circ q(x_\varepsilon(\sigma-\varepsilon)) + \varepsilon^2 \int_{-1}^{0} d\tau_1 \int_{-1}^{\tau_1} d\tau_2 f_{\tau_2} \circ f_{\tau_1} \circ q(x_\varepsilon(\sigma+\varepsilon\tau_2)) \\
&= \ldots = q(x_\varepsilon(\sigma-\varepsilon)) + \sum_{i=1}^{m} \varepsilon^i \int_{-1}^{0} d\tau_1 \int_{-1}^{\tau_1} d\tau_2 \ldots \int_{-1}^{\tau_{i-1}} d\tau_i f_{\tau_i} \circ f_{\tau_{i-1}} \circ \ldots \circ f_{\tau_1} \circ q(x_\varepsilon(\sigma-\varepsilon)) \\
&\quad + \varepsilon^{m+1} \int_{-1}^{0} d\tau_1 \int_{-1}^{\tau_1} d\tau_2 \ldots \int_{-1}^{\tau_m} d\tau_{m+1} f_{\tau_{m+1}} \circ f_{\tau_m} \circ \ldots \circ f_{\tau_1} \circ q(x_\varepsilon(\sigma+\varepsilon\tau_{m+1})).
\end{aligned}$$

Setting $q(x) = x$, we obtain the required asymptotic expansion

$$x_\varepsilon(\sigma) \sim x_\varepsilon(\sigma-\varepsilon) + \sum_{i=1}^{\infty} \varepsilon^i \int_{-1}^{0} d\tau_1 \int_{-1}^{\tau_1} d\tau_2 \ldots \int_{-1}^{\tau_{i-1}} d\tau_i f_{\tau_1} \circ \ldots \circ f_{\tau_i} \circ x_\varepsilon(\sigma-\varepsilon). \tag{4.5}$$

This makes it easy for us to obtain the asymptotic expansion in powers of ε of an arbitrary variation $\delta_j x(t; \varepsilon)$ at the point $t = \sigma$. We restrict ourselves to the evaluation of the variation $\delta_2 x(\sigma; \varepsilon)$, which is of interest to us.

In order to do this, we obviously need to substitute for f_{τ_i} in (4.5) the corresponding expression $g + Gp(\tau_i)$, and to separate the terms bilinear in $Gp(\tau_i)$ and $Gp(\tau_j)$, placing them in increasing powers of ε:

$$\delta_2 x(\sigma;\varepsilon) \sim \varepsilon^2 \sum_{i=0}^{\infty} \varepsilon^i \sum_{j_1+j_2+j_3=i-1} \int_{-1}^{0} d\tau_1 \int_{-1}^{\tau_1} d\tau_2 \ldots d\tau_{j_1} \int_{-1}^{\tau_{j_1}} dt \int_{-1}^{t} d\theta_1 \int_{-1}^{\theta_1} \ldots d\theta_{j_2}$$

$$\times \int_{-1}^{\theta_{j_2}} ds \int_{-1}^{s} d\vartheta_1 \int_{-1}^{\vartheta_1} \ldots d\vartheta_{j_3-1} \int_{-1}^{\vartheta_{j_3-1}} d\vartheta_{j_3} g^{j_3} \circ Gp(s) \circ g^{j_2} \circ Gp(t) \circ g^{j_1} \circ \tilde{x}(\sigma-\varepsilon).$$

With the aid of the convolution formula (2.14), the interior sum can be written in the form

$$\sum_{j_1+j_2+j_3=i} \int_{-1}^{0}\int_{-1}^{t} \frac{((1+s)g)^{j_3}}{j_3!} \circ Gp(s) \circ \frac{((t-s)g)^{j_2}}{j_2!} \circ Gp(t) \circ \frac{((-t)g)^{j_1}}{j_1!} \circ \tilde{x}(\sigma-\varepsilon)\, dt ds.$$

Therefore, by (4.3) and the equality $p(t) = 0$ for $t > 0$,

$$\delta_2 x(\sigma;\varepsilon) \sim \varepsilon^2 \sum_{i=0}^{\infty} \sum_{j_1+j_2+j_3=i} \int_{-1}^{0}\int_{-1}^{t} \frac{(\varepsilon(1+s)g)^{j_3}}{j_3!} \circ Gp(s) \circ \frac{(\varepsilon(t-s)g)^{j_2}}{j_2!} \circ Gp(t)$$

$$\circ \frac{(\varepsilon(-t)g)^{j_1}}{j_1!} \circ \tilde{x}(\sigma-\varepsilon)\, dtds = \varepsilon^2 \int_{-1}^{0}\int_{-1}^{t} e^{\varepsilon(1+s)g} \circ Gp(s) \circ e^{\varepsilon(t-s)g} \circ Gp(t) \circ e^{-\varepsilon t g} \circ \tilde{x}(\sigma-\varepsilon)\, dtds \tag{4.6}$$

$$= \varepsilon^2 e^{\varepsilon g} \circ \iint_{-1}^{1} h(t-s)\, e^{\varepsilon s \operatorname{ad} g} Gp(s) \circ e^{\varepsilon t \operatorname{ad} g} Gp(t) \circ \tilde{x}(\sigma-\varepsilon)\, dtds.$$

For an arbitrary function $q(x)$ we have

$$\frac{d^i}{d\tau^i} q(\tilde{x}(\sigma+\varepsilon\tau)) = \varepsilon^i g^i \circ q(\tilde{x}(\sigma+\varepsilon\tau)).$$

Therefore, the following asymptotic representation holds:

$$q(\tilde{x}(\sigma)) \sim \sum_{i=0}^{\infty} \frac{\varepsilon^i g^i}{i!} \circ q(\tilde{x}(\sigma-\varepsilon)) = e^{\varepsilon g} \circ q(\tilde{x}(\sigma-\varepsilon)),$$

which allows us to give (4.6) the form

$$\delta_2 x(\sigma;\varepsilon) \sim \varepsilon^2 \iint_{-1}^{1} h(t-s)\, e^{\varepsilon s \operatorname{ad} g} Gp(s) \circ e^{\varepsilon t \operatorname{ad} g} Gp(t) \circ \tilde{x}(\sigma)\, dtds. \tag{4.7}$$

We show that, if $p(t) \in P^{(m)}$, then the end-point of the second variation $\delta_2 x(a;\varepsilon)$ and its value $\delta_2 x(\sigma;\varepsilon)$ at the point σ are connected by the relation

$$|\delta_2 x(a;\varepsilon) - \Gamma(a,\sigma)\,\delta_2 x(\sigma;\varepsilon)| = O(\varepsilon^{2(2+m)}). \tag{4.8}$$

Hence, by (4.7), we can assert that for $p(t) \in P^{(m)}$ the first $m+1$ terms of the series

$$\varepsilon^2 \Gamma(a,\sigma) \iint_{-1}^{1} h(t-s)\, e^{\varepsilon s \operatorname{ad} g} Gp(s) \circ e^{\varepsilon t \operatorname{ad} g} Gp(t) \circ \tilde{x}(\sigma)\, dtds \tag{4.9}$$

express the end-point of the second variation $\delta_2 x(a;\varepsilon)$ to within $O(\varepsilon^{m+3})$.

In order to prove (4.8), we write

$$\delta_1 x(\sigma;\varepsilon) = \int_0^\sigma \Gamma(\sigma, t)\, G(\tilde{x}(t))\, p\left(\frac{t-\sigma}{\varepsilon}\right) dt$$

$$= \varepsilon \int_{-1}^0 \Gamma(\sigma, \sigma+\varepsilon\tau)\, G(\tilde{x}(\sigma+\varepsilon\tau))\, p(\tau)\, d\tau$$

$$= \varepsilon \int_{-1}^0 e^{\varepsilon\tau D}\Gamma(\sigma, \theta)\, G(\tilde{x}(\theta))\, p(\tau)\, d\tau = \varepsilon^{2+m} y(\varepsilon).$$

Since $\delta_1 x$ satisfies the homogeneous equation

$$\delta_1 \dot{x} = g_x(\tilde{x}(t))\, \delta_1 x$$

on the interval $\sigma < t < a$, we have

$$\delta_1 x(t;\sigma) = \varepsilon^{2+m}\Gamma(t,\sigma)\, y(\varepsilon), \qquad \sigma \leqslant t \leqslant a.$$

Further, for $\sigma < t < a$ we have

$$\delta_2 \dot{x} = g_x(\tilde{x}(t))\, \delta_2 x + \frac{1}{2} g_{xx}(\tilde{x}(t))\, [\delta_1 x(t), \delta_1 x(t)],$$

and we obtain (4.8):

$$\delta_2 x(a;\varepsilon) = \Gamma(a,\sigma)\, \delta_2 x(\sigma;\varepsilon) + \frac{\varepsilon^{2(2+m)}}{2} \int_\sigma^a g_{xx}(\tilde{x}(t))\, [\Gamma(t,\sigma)\, y(\varepsilon), \Gamma(t,\sigma)\, y(\varepsilon)]\, dt.$$

By Proposition 1.1, the property of the series (4.9) formulated above is retained if this series is replaced by its skew-adjoint part, i.e. by the series

$$\frac{1}{2}\varepsilon^2\Gamma(a,\sigma) \iint_{-1}^{1} h(t-s)\, [e^{\varepsilon s \operatorname{ad} g} Gp(s), e^{\varepsilon t \operatorname{ad} g} Gp(t)] \circ \tilde{x}(\sigma)\, dt ds.$$

Thus the latter is the Legendre representation of the end-point of the second variation (1.13) (for $\alpha(\varepsilon) = 1$ and $\beta(\varepsilon) = \varepsilon$). Hence, by an obvious generalization([4]) of Proposition 1.5, we arrive at the basic formula

$$\frac{1}{2}\Gamma(a,\sigma)\, [e^{\varepsilon s \operatorname{ad} g} Gp_1, e^{\varepsilon t \operatorname{ad} g} Gp_2] \circ \tilde{x}(\sigma) = \sum_{i=0}^{\infty} \frac{\varepsilon^i}{i!} K_i(t,s;\sigma)\, [p_1, p_2]. \tag{4.10}$$

We now note that one can derive the relation

$$\Gamma(a,\sigma)\, \mathfrak{L}_m f(\sigma)\, [p_1, p_2] = \frac{2}{m!} \frac{\partial^m K_m}{\partial s^m}(t,s;\sigma)\, [p_1, p_2]$$

from (1.19) and the definition of $\mathfrak{L}_m$. Utilizing (4.10), we obtain from this the equalities

([4]) The generalization consists in the fact that the identity $K(t,x)[p_1,p_2] \equiv 0$ follows from the equality

$$\iint_{-1}^{1} h(t-s) K(t,s)\, [p(t), p(s)]\, dt ds = 0$$

for arbitrary $p(t) \in P^{(l)}$ which vanish outside of a fixed interval in $[-1, 1]$, but not necessarily for all $p(t) \in P^{(l)}$.

$$\Gamma(a,\sigma)\,\mathfrak{L}_m f(\sigma)[p_1,p_2] = \frac{2}{\varepsilon^m}\frac{\partial^m}{\partial s^m}\sum_{i=0}^{\infty}\frac{\varepsilon^i}{i!}K_i\Big|_{\varepsilon=0}$$
$$= \frac{1}{\varepsilon^m}\Gamma(a,\sigma)\frac{\partial}{\partial s^m}[e^{\varepsilon s \operatorname{ad} g}Gp_1, e^{\varepsilon t \operatorname{ad} g}Gp_2]\circ\tilde{x}(\sigma)\Big|_{\varepsilon=0}$$
$$= \Gamma(a,\sigma)\,[(\operatorname{ad}^m g)\,Gp_1, Gp_2]\circ\tilde{x}(\sigma),$$

which yield the required expression for $\mathfrak{L}_m f$ found by Krener in [4]:

$$\mathfrak{L}_m f(\sigma)[p_1,p_2] = [(\operatorname{ad}^m g)\,Gp_1, Gp_2]\circ\tilde{x}(\sigma). \tag{4.11}$$

The operator series

$$\sum_{i=1}^{\infty}\frac{\varepsilon^i}{i!}[(\operatorname{ad}^i g)\,Gp_1, Gp_2] = [e^{\varepsilon \operatorname{ad} g}Gp_1, Gp_2]$$

will be called the *generating series* for the Legendre operators $\mathfrak{L}_m f$.

We denote by

$$G_i = \frac{\partial}{\partial u^i}(g+Gu) = \frac{\partial}{\partial u^i}f(x,u)$$

the ith column of the $n\times r$ matrix G, $i=1,\dots,r$. By (4.11), the $r\times r$ matrix $\|l_{ij}\|$ of the scalar form $\psi(\sigma)\mathfrak{L}_m f(\sigma)[p_1,p_2]$ can be expressed in the form

$$\|l_{ij}(\sigma)\| = \|\psi(\sigma)\,[(\operatorname{ad}^m g)\,G_i, G_j]\circ\tilde{x}(\sigma)\|.$$

Let us give a somewhat different representation of elements $l_{ij}(\sigma)$, which, in essence, is contained in the work of Kelley, Kopp, and Moyer [1]. We have

$$\left(\frac{\partial}{\partial u^i}\operatorname{ad}^{m+1}(g+Gu)\Big|_{u=0}\right)G_j = \sum_{k=0}^{m}(\operatorname{ad}^k g)\,(\operatorname{ad} G_i)\,(\operatorname{ad}^{m-k} g)\,G_j.$$

Hence by (4.4) we can write

$$\psi(\sigma)\left(\frac{\partial}{\partial u^i}\operatorname{ad}^{m+1}f(x,u)\Big|_{u=0}\right)\frac{\partial f}{\partial u^j}\circ\tilde{x}(\sigma)$$
$$= \psi(\sigma)\,[G_i,(\operatorname{ad}^m g)\,G_j]\circ\tilde{x}(\sigma) + \sum_{k=1}^{m}\frac{d^k}{d\sigma^k}\psi(\sigma)\,[G_i,(\operatorname{ad}^{m-k} g)\,G_j]\circ\tilde{x}(\sigma).$$

If the identities (in p_1, p_2)

$$\psi(\sigma')\,\mathfrak{L}_k f(\sigma')[p_1,p_2] = \psi(\sigma')\,[(\operatorname{ad}^k g)\,Gp_1, Gp_2]\circ\tilde{x}(\sigma')\equiv 0\ \forall\sigma'\in O_\sigma,\ \forall k=1,\dots,m-1,$$

hold for a neighborhood O_σ of σ, then

$$\psi(\sigma)\left(\frac{\partial}{\partial u^i}\operatorname{ad}^{m+1}f\right)\Big|_{u=0}\frac{\partial f}{\partial u^j}\circ\tilde{x}(\sigma) = \psi(\sigma)\,[G_i,(\operatorname{ad}^m g)\,G_j]\circ\tilde{x}(\sigma). \tag{4.12}$$

Since the matrix $\|l_{ij}(\sigma)\|$ is symmetric if the conditions (4.12) are satisfied, we obtain the required expressions

$$l_{ij}(\sigma) = -\psi(\sigma)\frac{\partial}{\partial u^i}(\operatorname{ad}^{m+1}f)\,\frac{\partial f}{\partial u^j}\circ\tilde{x}(\sigma)\Big|_{u=0},$$

which hold at any point σ that satisfies (4.12).

Received 11/MAR/76

BIBLIOGRAPHY

1. Henry J. Kelley, Richard E. Kopp and H. Gardner Moyer, *Singular extremals*, Topics in Optimization, Academic Press, New York, 1967, pp. 63–101. MR **35** #5996.

2. R. Gabasov and F. M. Kirillova, *Singular optimal controls*, "Nauka", Moscow, 1973. (Russian)

3. Arthur J. Krener, *The high order maximum principle*, Geometric Methods in Systems Theory (D. Q. Mayne and R.W. Brockett, editors), Reidel, Dordrecht, 1973, pp. 174–183.

4. ______, *The high order maximal principle and its application to singular extremals*, SIAM J. Control Optimization **15** (1977), 256–293.

Translated by K. MAKOWSKI