Control
and Cybernetics
VOL 13 (1984) No. 4

Sufficient conditions for a minimum in a classical optimal control problem

by

ANDRZEJ NOWAKOWSKI

Łódź University
Institute of Mathematics
S. Banacha 22
90-238 Łódź, Poland

Based on the classical methods of the variational calculus a practical method of construction of an optimal feedback control has been given in the paper. In consequence, it yields Weierstrass sufficient conditions for a minimum.

Introduction

The control theory literature has been concerned mainly with sufficient conditions for global optimality, the most well-known result being the Hamilton-Jacobi-Bellman theorem [2]. This, however, requires the existence of a feedback control and a solution to the Hamilton-Jacobi partial differential equation. The local theory treats mainly the exact expression ΔI of a change in a functional due to a change in control [3]. A common technique is to add to the integrand of the functional an exact differential in order to transform the original problem into a simpler one (see, e.g. [2]; on the classical technique [1]).

In the calculus of variations there exists a technique of calculating that exact differential. It is based on the theory of the field.

The present paper examines the case when the differential form ydx-Hdt becomes exact in a control problem, which then gives a method of constructing an optimal feedback control and, in consequence, the Weierstrass sufficiency theorem.

1. Preliminary notes and the algorithm of Huygens for on optimal control

1. We shall be dealing with a Borel subset U of R^r. The assumption that U is Borel is connected with the maximum principle. A Lebesgue measurable function $u\colon [0, 1]\to U$ will be called a control or an admissible control. The choice of the time interval $[0, 1]$ is a normalization for notational convenience.

Let $f(t, x, u)$ be a vector function $f\colon [0,1)\times R^n\times R^r\to R^n$, and let $L(t, x, u)$ be a scalar function defined n $[0, 1]\times R^n\times R_r$.

We suppose, for each t in $[0, 1]$ and u in U, that the functions $(x)\to f(t, x, u)$, $(x)\to L(t, x, u)$ are C^1 in R^n, and the functions $(t, x, u)\to f(t, x, u)$, $(t, x, u)\to L(t, x, u)$ are continuous in $[0, 1]\times R^n\times R^r$.

An admissible trajectory x corresponding to a control u is an absolutely continuous function $x\colon [0, 1]\to R^n$ satisfying

$$\dot{x}(t)=f\big(t, x(t), u(t)\big), \text{ almost everywhere in } [0, 1]. \tag{1.1}$$

2. The class of all admissible pairs $(x(t), u(t))$, $t\in[0, 1]$, such that in the first extremity $x(0)=0$ and in the second $x(1)=e$ where e is a fixed point in R^n, will be denoted by M.

Our main problem is the following:

Find a minimum of the integral

$$I(x, u)=\int_0^1 L\big(\tau, x(\tau), u(\tau)\big)\, d\tau \tag{1.2}$$

"along" the pair $(x(t), u(t))$ from among all pairs of M.

Further, we shall consider not only a single minimum problem, but various minimum problems which arise when the second extremity of a trajectory takes different positions in the space R^{n+1} of the variables (t, x). The integrand L and the first extremity will be kept fixed.

To this effect, we denote by $M(t, x)$ the class of all admisible pairs $(x(\tau), u(\tau))$, $\tau\in[0, t]$, such that $x(0)=0$, $x(t)=x$, $0<t\leqslant 1$, $x\in R^n$. The same symbol x used for a trajectory and a point of R^n will not lead to confusion, for its meaning will be clear from the context. In the minimum problem for this class we take the integral in (1.2) in the interval $[0, t]$, i.e.

$$I^t(x, u)=\int_0^t L\big(\tau, x(\tau), u(\tau)\big)\, d\tau\,. \tag{1.3}$$

3. A pair $(x(\tau), u(\tau))$ of $M(t, x)$ will be said to satisfy the maximum principle (see, e.g. [2]) if there exists an absolutely continuous vector function $y\colon [0, 1]\to R^n$ such that

$$\begin{aligned}
&\text{(i)}\ \ \dot{y}(\tau)=-f_x\big(\tau, x(\tau), u(\tau)\big)\, y(\tau)+\lambda L_x\big(\tau, x(\tau), u(\tau)\big)\,,\\
&\text{(ii)}\ \ y(\tau)\, f\big(\tau, x(\tau), u(\tau)\big)-\lambda L\big(\tau, u(\tau), x(\tau)\big)=\\
&\qquad\qquad =\sup\,\{y(\tau)\, f(\tau, x(\tau), u)-\lambda L(\tau, x(\tau), u)\colon u\in U\}
\end{aligned} \tag{1.4}$$

where we consider only the situation when $\lambda=1$.

4. Suppose now that, in any way (for example from the maximum principle and the existence theorems), we are able to suspect that, for the pair $(x^0(\tau), u^0(\tau))$,

$\tau \in [0, t]$ of $M(t, x)$, integral (1.3) attains its minimum in the class $M(t, x)$. We denote the value of this suspected minimum by $S(t, x)$.

More generally, we can define $S(t, x) = \inf_{(x, u) \in M(t, x)} I^t(x, u) = \inf_{M(t, x)} \int_0^t L(\tau, x(\tau), u(\tau))\, d\tau$. The function $S(t, x)$ is then always defined in $M(t, x)$ and called the value function. The $S(t, x)$ may have the values $-\infty$ and $+\infty$, the latter whenever the class $M(t, x)$ is empty. We shall assume, however, in the whole paper that $M \neq \emptyset$ and so is $M(t, x)$ for some (t, x). It is clear that the pair $(x^0(\tau), u^0(\tau))$, $\tau \in [0, t]$ and the value $S(t, x)$ depend on $(t, x) \in [0, 1] \times R^n$. Thus $S(t, x)$ now becomes a function of (t, x). Evidently, it vanishes at the point $(0, 0)$, so that we can also write (1.3) along $(x^0(\tau), u^0(\tau))$, $\tau \in [0, t]$, as the difference $S(t, x) - S(0, 0)$. We shall write $\Phi_0 = \Phi_0(t, x, u)$ for the exact derivative $S_t + \dot{x} S_x = S_t + f S_x$ of the function $S(t, x)$ if it exists. Then $S(t, x) - S(0, 0)$ is along $(x^0(\tau), u^0(\tau))$ of $M(t, x)$ the integral of Φ_0. Of course, the pair $(x^0(\tau), u^0(\tau))$, $\tau \in [0, 1]$ must satisfy the inequality $I(x, u) \geqslant I(x^0, u^0)$, i.e., if Φ_0 exists, there is

$$\int_0^1 \{L(\tau, x(\tau), u(\tau)) - \Phi_0(\tau, x(\tau), u(\tau))\}\, d\tau \geqslant 0 \tag{1.5}$$

for all pairs (x, u) of M. This inequality is certainly true for all admissible triplets (t, x, u),

$$L(t, x, u) \geqslant \Phi_0(t, x, u). \tag{1.6}$$

5. We shall now reformulate the consideration of 4. in the following algorithm of Huygens (comp. [5]). Let M be a class of pairs (x, u) described in 2., and let (x^0, u^0) be a member of that class. If there exists an exact derivative Φ_0 such that $L = \Phi_0$ along (x^0, u^0) and that $L \geqslant \Phi_0$ along all pairs (x, u) of M, then $I(x^0, u^0)$ is the minimum of $I(x, u)$ for $(x, u) \in M$.

The proof of this algorithm is evident. Indeed, $I(x, u)$ is, for $(x, u) \in M$, not less than the corresponding integral of Φ_0 and we can take the latter along (x^0, u^0) without altering its value which is then $I(x^0, u^0)$.

2. A method of feedback control

Denote by V the set in R^{n+1} covered by graphs of all admissible trajectories, and let $D_{\bar{u}}$ be any domain in V containing the interval $[0, 1]$. Let, further, $\bar{u}(t, x)$ denote a vector-valued function defined in $D_{\bar{u}}$ with values in U.

Definition 2.1. We call $\bar{u}(t, x)$ an optimal feedback control if there exists an exact derivative

$$\Phi_0 = S_t + \dot{x} S_x = S_t + f S_x \tag{2.1}$$

such that, for all admissible triplets (t, x, u) (i.e. $(t, x) \in D_{\bar{u}}$ and those $u(t) \in U$ for which the graphs of the corresponding trajectories lie in $D_{\bar{u}}$);

$$L(t, x, u) \geqslant \Phi_0(t, x, u) \tag{2.2}$$

with equality when $u = \bar{u}$.

Let $M_{D_{\bar{u}}}$ denote a subclass of M of those pairs (x, u) whose graphs of trajectories lie in $D_{\bar{u}}$. Then, by the algorithm of Huygens, it is evident, if an optimal feedback control $\bar{u}(t, x)$ exists, that each pair $(x^0(t), \bar{u}(t, x^0(t)))$ of $M_{D_{\bar{u}}}$ affords a minimum to $I(x, u)$ relative to pairs (x, u) of $M_{D_{\bar{u}}}$.

In 4. (§ 1) we started from the function S which, of course, determines Φ_0. Conversely, Φ_0 determines S for an additive constant. Our task is to show that when Φ_0 exists it is also completely determined by $\bar{u}$.

To this effect, we observe, by (2.2), that the difference $L - \Phi_0$ regarded as a function of u only attains its minimum equal to zero at $u = \bar{u}$. This requires that, if we assume $\bar{y}(t, x) = S_x(t, x)$, then

$$S_t = L - f S_x = L - f \bar{y} \tag{2.3}$$

at $u = \bar{u}$. If we put S_t of (2.3) in inequality (2.2), then the latter is known as the Hamilton-Jacobi-Bellman sufficient condition for global optimality (see [3]), and if we denote $-H(t, x, \bar{u}(t, x), S_x) = L(t, x, \bar{u}(t, x)) - f(t, x, \bar{u}(t, x))\, S_x(t, x)$, then (2.3) becomes

$$S_t = -H(t, x, \bar{u}(t, x), S_x), \tag{2.4}$$

i.e. the Hamilton-Jacobi-Bellman differential equation. Simultaneously, by the above

$$H(t, x, \bar{u}(t, x), \bar{y}(t, x)) = \max_u \{\bar{y}(t, x) f(t, x, u) - L(t, x, u)\} \quad (t, x) \in D_{\bar{u}} \tag{2.5}$$

or

$$H(t, x, \bar{u}(t, x), \bar{y}(t, x)) - H(t, x, u, \bar{y}(t, x)) \geqslant 0 \tag{2.6}$$

for all admissible (t, x, u). The left-hand side of (2.6) is known as the extension o the excess function of Weierstrass (see [4]).

We may also express (2.3) by saying that for rectifiable curves C lying in a simply connected subdomain of $D_{\bar{u}}$, the curvintegral

$$\int_C \Big(L(t, x, \bar{u}(t, x)) - f(t, x, \bar{u}(t, x))\, \bar{y}(t, x)\Big)\, dt + \bar{y}(t, x)\, dx \tag{2.7}$$

depends only on the endpoints of C. This expression is termed Hilbert's independent integral: its value, on account of (2.3), is clearly $S(B)$-$S(A)$ where A, B are the endpoints of C.

We shall term $D_{\bar{u}}$-levels the set of points $P \in D_{\bar{u}}$ on which $S(P)$ is constant. If we denote the left-hand side of (2.6), by E, then we have $L = \Phi_0 + E$ and therefore, for any pair (x, u) of M whose arc of the trajectory C joins in $D_{\bar{u}}$ two levels $S=S_1$ and $S=S_2$,

$$\int_C L(t, x, u)\, dt = S_2 - S_1 + \int_C E dt\,. \tag{2.8}$$

This is an extension of the famous formula of Weierstrass, whose equivalent in the classical calculus, according to Young [5], has revolutionized the variational calculus. In contrast, to date, formula (2.8) or its analogon, that is, the exact expression for $\Delta I = I(x, u) - S(1, e)$ (see [4]), do not appear to have played a central role in modern control theory, although the expressions ΔI do appear as tools in the derivation of optimization algorithms (see [3]).

3. The necessary exactness condition

The exactness condition that we have to study is that the expression

$$\bar{y}(t, x)\, dx + (L(t, x, u) - \bar{y}(t, x)\, \dot{x})\, dt \tag{3.1}$$

reduces to an exact differential dS in the variables (t, x) when we substitute $\dot{x} = f(t, x, u)$ and $u = \bar{u}(t, x)$; $\bar{y}$ is defined as in § 2. We wish to find out when this occurs.

To this effect, we shall enlarge our setting by using the transformation

$$(t, \sigma) \to (t, x)\,, \quad (t, x) \in D_{\bar{u}}\,, \tag{3.2}$$

in which the time t is unchanged. This means that x is replaced by a function $x(t, \sigma)$ where (t, σ) belongs to some open simply connected set $(0, 1) \times G$ of R^{m+1}. Next, we enlarge (3.2) to a map of triplets

$$(t, \sigma, \dot{\sigma}) \to (t, x, \dot{x}) \tag{3.2'}$$

by putting $\dot{x} = x_t + x_\sigma \dot{\sigma}$, i.e. we substitute for the variable $\dot{x}$ the total derivative of $x(t, \sigma)$. Evidently, now

$$f(t, x(t, \sigma), u) = x_t + x_\sigma \dot{\sigma}\,. \tag{3.3}$$

Further, the dot over a function of (t, σ) will always denote the total derivative. We suppose that the function $x(t, \sigma)$ defined in (3.2′) is C^2 in $(0, 1) \times G$, and $\bar{y}(t, \sigma) = \bar{y}(t, x(t, \sigma))$ is C^1 there.

Now, we study under what additional circumstances a family of curves of $(0, 1) \times G$ given by a differential equation

$$x_\sigma(t, \sigma)\, \dot{\sigma} = -x_t(t, \sigma) + f\big(t, x(t, \sigma), \bar{u}\,(t, x(t, \sigma))\big) \tag{3.3'}$$

is such that map (3.2′) turns expression (3.1) into an exact differential dS in the variables (t, σ). Then, the integral of (3.1) on an arbitrary curve of $(0, 1) \times G$ depends only on the endpoints. We call it the corresponding Hilbert independent integral.

We shall suppose that, for the functions $\tilde{L}(t,\sigma)=L\big(t,x(t,\sigma),\bar{u}(t,x(t,\sigma))\big)$, $\tilde{f}(t,\sigma)=f\big(t,x(t,\sigma),\bar{u}(t,x(t,\sigma))\big)$, there exist continuous derivatives $\tilde{L}_\sigma(t,\sigma)$, $\tilde{f}_\sigma(t,\sigma)$ in $(0,1)\times G$ and $\frac{\partial}{\partial\sigma}L\big(t,x,\bar{u}(t,x(t,\sigma))\big)$, $\frac{\partial}{\partial\sigma}f\big(t,x,\bar{u}(t,x(t,\sigma))\big)$, for each fixed $(t,x)\in D_{\bar{x}}$, in $(0,1)\times G$ and they satisfy at (t,x), $x=x(t,\sigma)$, the relations:

$$\begin{aligned}\frac{\partial\tilde{L}}{\partial\sigma}&=\frac{\partial}{\partial\sigma}L\big(t,x,\bar{u}(t,x(t,\sigma))\big)+L_x\big(t,x,\bar{u}(t,x(t,\sigma))\big)\,x_\sigma(t,\sigma)\,,\\ \frac{\partial\tilde{f}}{\partial\sigma}&=\frac{\partial}{\partial\sigma}f\big(t,x,\bar{u}(t,x(t,\sigma))\big)+f_x\big(t,x,\bar{u}(t,x(t,\sigma))\big)\,x_\sigma(t,\sigma)\,.\end{aligned}\tag{3.3''}$$

We note at once that the induced feedback control $\bar{u}(t,x(t,\sigma))$ expressed by map (3.2) satisfies, for $(t,\sigma)\in(0,1)\times G$, the induced maximum relation

$$\begin{aligned}&\bar{y}(t,x(t,\sigma))\,f\big(t,x(t,\sigma),\bar{u}(t,x(t,\sigma))\big)-L\big(t,x(t,\sigma),\bar{u}(t,x(t,\sigma))\big)=\\ &\qquad=\max_u\,\{\bar{y}(t,x(t,\sigma))\,f(t,x(t,\sigma),u)-L(t,x(t,\sigma),u)\}\,.\end{aligned}\tag{3.3'''}$$

It is clear that one of the necessary and sufficient conditions for the existence of $S(t,\sigma)$ is satisfaction of the relations

$$S_{\sigma_i\sigma_j}=S_{\sigma_j\sigma_i},\quad S_{\sigma_i t}=S_{t\sigma_i},\qquad i,j=1,\dots,m,\quad \sigma=(\sigma_1,\dots,\sigma_m)\,.\tag{3.4}$$

Let $1\leqslant i,j\leqslant m$ be arbitrarily fixed. After substitution of (3.2′) in (3.1) we easily get from the expression obtained that

$$S_{\sigma_i\sigma_j}-S_{\sigma_j\sigma_i}=\frac{\partial}{\partial\sigma_j}(\tilde{y}x_{\sigma_i})-\frac{\partial}{\partial\sigma_i}(\tilde{y}x_{\sigma_j})=x_{\sigma_i}\tilde{y}_{\sigma_j}-x_{\sigma_j}\tilde{y}_{\sigma_i}\,,\tag{3.5}$$

$$S_{\sigma_i t}-S_{t\sigma_i}=\frac{\partial}{\partial t}(\tilde{y}x_{\sigma_i})-\frac{\partial}{\partial\sigma_i}\big(\tilde{y}x_t+(\tilde{L}-\tilde{y}\dot{x})\big)\,.\tag{3.6}$$

We calculate exactly the right-hand sides of (3.5) and (3.6). Thus

$$\frac{\partial}{\partial t}(\tilde{y}x_{\sigma_i})=\tilde{y}x_{\sigma_i t}+x_{\sigma_i}\tilde{y}_t=x_{\sigma_i}\dot{\tilde{y}}+\tilde{y}x_{\sigma_i t}+x_{\sigma_i}(\tilde{y}_t-\dot{\tilde{y}})\,,\tag{3.7}$$

$$\frac{\partial}{\partial\sigma_i}(\tilde{y}x_t)=\tilde{y}_{\sigma_i}x_t+\tilde{y}x_{\sigma_i t}=\tilde{y}_{\sigma_i}\dot{x}+\tilde{y}x_{\sigma_i t}+\tilde{y}_{\sigma_i}(x_t-\dot{x})\,.\tag{3.8}$$

Note that, by (3.5),

$$\begin{aligned}x_{\sigma_i}(\tilde{y}-\dot{\tilde{y}}_t)-\tilde{y}_{\sigma_i}(x_t-\dot{x})&=-\sum_{j=1}^{m}(x_{\sigma_i}\tilde{y}_{\sigma_j}-\tilde{y}_{\sigma_i}x_{\sigma_j})\,\dot{\sigma}_j=\\ &=-\sum_{j=1}^{m}(S_{\sigma_i\sigma_j}-S_{\sigma_j\sigma_i})\,\dot{\sigma}_j\,.\end{aligned}\tag{3.9}$$

We also note that, since (2.5) is satisfied (comp. also (3.3‴)), remembering (3.3),

$$\frac{\partial}{\partial \sigma_i}(\tilde{y}\dot{x}-\tilde{L})=\frac{\partial}{\partial \sigma_i}(\tilde{y}\tilde{f}-\tilde{L})=\tilde{y}_{\sigma_i}\dot{x}+\tilde{y}f_x x_{\sigma_i}-L_x x_{\sigma_i}+\frac{\partial}{\partial \sigma_i}\{\tilde{y}(t,x)\times$$
$$\times f\big(t,x,\bar{u}\left(t,x(t,\sigma)\right)\big)-L\big(t,x,\bar{u}\left(t,x(t,\sigma)\right)\big)\}=\tilde{y}_{\sigma_i}\dot{x}+\tilde{y}f_x x_\sigma-L_x x_{\sigma_i} \qquad (3.10)$$

where the t, x in brackets are momentarily fixed points. By substituting (3.7), (3.8) and (3.10) in (3.6), we find that

$$S_{\sigma_i t}-S_{t\sigma_i}=x_{\sigma_i}[\dot{\tilde{y}}+\tilde{y}f_x-L_x]-\sum_{j=1}^{m}(S_{\sigma_i\sigma_j}-S_{\sigma_j\sigma_i})\,\dot{\sigma}_j\,.$$

Hence

$$x_{\sigma_i}[\dot{\tilde{y}}+\tilde{y}f_x-L_x]=(S_{\sigma_i t}-S_{t\sigma_i})+\sum_{j=1}^{m}(S_{\sigma_i\sigma_j}-S_{\sigma_j\sigma_i})\,\dot{\sigma}_j\,. \qquad (3.11)$$

Here, by (3.4), the summand on the right vanishes and, therefore, so do both sides of (3.11) for each i, $i=1, \ldots, m$. If we change the order of the two factors on the left, the resulting set of equations, obtained from the vanishing of the left-hand side for various i, may now be written as the single vector equation

$$[\dot{\tilde{y}}+\tilde{y}f_x-L_x]\,x_\sigma=0\,. \qquad (3.12)$$

This equation will be termed the induced conjugate differential equation.

The induced triplet (σ, u^*, y^*), when σ, $u^*=\bar{u}(t, x(t,\sigma))$, $y^*=y^*(t,\sigma)$ satisfy (3.3′), (3.3‴), (3.12), respectively, will be referred to as satisfying the induced maximum principle.

We have thus proved the following

THEOREM 3.1. *Let transformation* (3.2) *be such that* $x(t,\sigma)$ *is* C^2 *in* $(0,1)\times G$, $\tilde{y}(t,\sigma)$ *is* C^1 *there, and there exist* $\tilde{L}_\sigma$, $\tilde{f}_\sigma$ *that satisfy relations* (3.3″). *Moreover,* (3.2) *turns the expression* (3.1) *into an exact differential* dS *in* (t,σ). *Then, for each induced trajectory* σ *under control* $\bar{u}(t,\sigma)=\bar{u}(t,x(t,\sigma))$, *i.e., satisfying* (3.3′), *there exists an induced conjugate function* $y(t,\sigma)$ *such that the triplet* $(\sigma,\bar{u},y)$ *satisfies the induced maximum principle. The triplets* $(\sigma,\bar{u}(t,\sigma),\tilde{y}(t,\sigma))$ *also satisfy the induced maximum principle.*

Let now (3.2) be the identity map, so that x_σ is the identity matrix and let the assumptions of theorem 3.1 be satisfied; setting there the variable x instead of σ and taking the open simply connected set $D_{\bar{u}}$ instead of $(0,1)\times G$, we easily conclude that:

THEOREM 3.1′. *If* $x^0(t)$, $t\in(0,1)$, *satisfies*

$$\dot{x}^0(t)=f\big(t,x^0(t),\bar{u}(t,x^0(t))\big)\,,$$

then the function $y^0(t)=\tilde{y}(t, x^0(t))=S_x(t, x^0(t))$, $t\in(0, 1)$, *satisfies the conjugate vector differential equation*

$$\dot{y}^0(t)=-y^0(t) f_x\left(t, x^0(t), \bar{u}\left(t, x^0(t)\right)\right)+L_x\left(t, x^0(t), \bar{u}\left(t, x^0(t)\right)\right).$$

4. The sufficient exactness condition

In this section we give conditions for the following functions defined in $(0, 1)\times G$ to vanish, where we assume $m=n$

$$\varphi_i(t, \sigma)=S_{\sigma_i t}(t, \sigma)-S_{t\sigma_i}(t, \sigma) \quad i=1, \ldots, n, \tag{4.1}$$

$$\psi_{ij}(t, \sigma)=S_{\sigma_i\sigma_j}(t, \sigma)-S_{\sigma_j\sigma_i}(t, \sigma), \quad i, j=1, \ldots, n. \tag{4.1'}$$

We put $\psi_i=(\psi_{i1}, \ldots, \psi_{in})$, $\psi_j=(\psi_{1j}, \ldots, \psi_{nj})$, $i, j=1, \ldots, n$.

(4.2) From now on, we can and shall suppose, besides the assumptions of § 3, that: the family of induced triplets N satisfies the induced maximum principle, $\det(x_\sigma)\neq 0$ and $\tilde{y}$ is C^2 in $(0, 1)\times G$, there exist $\tilde{f}_{\sigma\sigma}$ and $\tilde{L}_{\sigma\sigma}$.

In accordance with Lagrange we shall apply the following, easy to verify identities (see, e.g. [5], § 15)

$$\frac{\partial}{\partial\sigma}\psi_{ij}-\frac{\partial}{\partial\sigma_j}\psi_i-\frac{\partial}{\partial\sigma_i}\psi_j=0. \tag{4.3}$$

$$\frac{\partial}{\partial t}\psi_{ij}-\frac{\partial}{\partial\sigma_j}\varphi_i+\frac{\partial}{\partial\sigma_i}\varphi_j=0. \tag{4.4}$$

Thus, because of (3.11) and (4.1), (4.1′), we have

$$\varphi_i+\dot{\sigma}\psi_i=0. \tag{4.5}$$

The total derivatives of $\psi_{ij}(t, \sigma)$ have the form

$$\dot{\psi}_{ij}=\frac{\partial}{\partial t}\psi_{ij}+\dot{\sigma}\frac{\partial}{\partial\sigma}\psi_{ij}. \tag{4.6}$$

Applying successively (4.3), (4.4) and (4.5) to the right-hand side of (4.6), we find, remembering (3.3′), that

$$\dot{\psi}_{ij}=-\psi_i\frac{\partial\dot{\sigma}}{\partial\sigma_j}-\psi_j\frac{\partial\dot{\sigma}}{\partial\sigma_i}, \quad i, j=1, \ldots, n. \tag{4.7}$$

This is rather a classical result first obtained by Lagrange and we can encounter similar results in, for example, [5] or [1].

Note that (4.7) is a system of homogeneous linear differential equations for ψ_{ij}. Moreover, the φ_i are, by (4.5), homogeneous linear combinations of the ψ_{ij}. Thus we have ϑ.

THEOREM 4.1. *The functions* φ_i, ψ_{ij}, $i, j=1, \ldots, n$, *will vanish identically in* $(0, 1)\times G$ *if* ψ_{ij} *vanish at some point of each trajectory of the family* N *(i.e. the curve satisfying* (3.3′)).

Proof. It is a direct consequence of the property of solutions of a system of homogeneous linear differential equations.

In the particular case of $\dot{\sigma}=0$, i.e. if $\sigma=$const along each curve satisfying (3.3′), then along each such curve the functions $\psi_{ij}=x_{\sigma_i}\tilde{y}_{\sigma_j}-x_{\sigma_j}\tilde{y}_{\sigma_i}$ are constant.

Now, we give sufficient condition for the vanishing of the φ_i, ψ_{ij}.

To this effect, let $t=t(\sigma)$, $\sigma\in G$, denote some C^1-locus which cuts each trajectory σ of N but is not tangent to it (i.e. $t_\sigma\dot{\sigma}\neq 1$) and let

$$Y(t, \sigma)\, d\sigma - K(t, \sigma)\, dt \tag{4.8}$$

where $Y=(Y_1, \ldots, Y_n)=\tilde{y}x_\sigma$, $-K=\tilde{y}x_t+(L-\tilde{y}\dot{x})$, denote expression (3.1) in map (3.2′).

THEOREM 4.2. *The functions* φ_i, ψ_{ij}, $i, j=1, \ldots, n$, *will vanish in* $(0, 1)\times G$ *if expression* (4.8) *is an exact differential in the variable* σ *on the locus* $t=t(\sigma)$, $\sigma\in G$.

Proof. We first observe that, by the definition of (4.1′), $\psi_{ii}=0$, $\psi_{ij}=-\psi_{ji}$, $i, j=1, \ldots, n$. Hence $\sum_{i,j}\dot{\sigma}_i\psi_{ij}\dot{\sigma}_j=0$ and, by (4.5), we have the identity

$$\varphi\dot{\sigma}=0\,. \tag{4.9}$$

where $\varphi=(\varphi_1, \ldots, \varphi_n)$.

Next, we set $\zeta=t_\sigma\dot{\sigma}$, $\bar{Y}(\sigma)=(\bar{Y}_1(\sigma), \ldots, \bar{Y}_n(\sigma))=Y(t(\sigma), \sigma)$, $\bar{K}(\sigma)=K(t(\sigma), \sigma)$. At the locus $t=t(\sigma)$, $\sigma\in G$, (4.8) takes the form

$$(\bar{Y}-\bar{K}t_\sigma)\, d\sigma=\sum_{i=1}^{n}(\bar{Y}_i-\bar{K}t_{\sigma_i})\, d\sigma_i\,. \tag{4.10}$$

Since (4.10) is exact by hypothesis, we must have

$$\frac{\partial}{\partial\sigma_j}(\bar{Y}_i-t_{\sigma_i}\bar{K})-\frac{\partial}{\partial\sigma_i}(\bar{Y}_j-t_{\sigma_j}\bar{K})=0\,,\quad i, j=1, \ldots, n\,. \tag{4.11}$$

Here the left-hand side is the value of the expression

$$\left(\frac{\partial}{\partial\sigma_j}+t_{\sigma_j}\frac{\partial}{\partial t}\right)(Y_i-t_{\sigma_i}K)-\left(\frac{\partial}{\partial\sigma_i}+t_{\sigma_i}\frac{\partial}{\partial t}\right)(Y_j-t_{\sigma_j}K)$$

for $t=t(\sigma)$, which is equivalent to

$$\left(\frac{\partial Y_i}{\partial\sigma_j}-\frac{\partial Y_j}{\partial\sigma_i}\right)+t_{\sigma i}\left(\frac{\partial Y_i}{\partial t}+\frac{\partial K}{\partial\sigma_i}\right)-t_{\sigma i}\left(\frac{\partial Y_i}{\partial t}+\frac{\partial K}{\partial\sigma_j}\right) \tag{4.11′}$$

and may be written in terms of (4.1), (4.1′) (comp. (3.5), (3.6)) in the form $\psi_{ij}+t_{\sigma_j}\varphi_i-t_{\sigma_i}\varphi_j$. By (4.11), this vanishes for $t=t(\sigma)$, $i,j=1,\dots,n$. Hence we obtain the vector equation

$$\psi_i+t_\sigma\varphi_i=t_{\sigma_i}\varphi\,. \tag{4.12}$$

By multiplying scalarwise by $\dot\sigma$ and using (4.5), the definition of ζ and identity (4.9), we find that

$$(\zeta-1)\,\varphi_i=0\,,\ i=1,\dots,n\,,$$

whence $\varphi_i=0$, $i=1,\dots,n$, and thus $\varphi=0$, so that (4.12) reduces to $\psi_i=0$. This shows that, for $t=t(\sigma)$, all ψ_{ij} vanish and so, the assumptions of theorem 4.1 are satisfied. This completes the proof. ∎

NOTE 1. In the particular case of $\dot\sigma=0$, the nontangency condition is always satisfied.

NOTE 2. If the functions φ_i, ψ_{ij}, $i,j=1,\dots,n$, vanish at the locus $t=t(\sigma)$, $\sigma\in G$, then expression (4.8) is an exact differential in σ there. This is a direct consequence of the relations

$$\frac{\partial}{\partial\sigma_j}(\bar Y_i-t_{\sigma_i}\bar K)-\frac{\partial}{\partial\sigma_i}(\bar Y_j-t_{\sigma_j}\bar K)=\psi_{ij}+t_{\sigma j}\varphi_i-t_{\sigma i}\varphi_j\,.$$

COROLLARY 4.1. *If the assumptions of theorem 4.2 are satisfied, the function* $S(\bar t,\bar\sigma)$, $(\bar t,\bar\sigma)\in(0,1)\times G$, *exists and may be defined by*

$$S(\bar t,\bar\sigma)=-\int_{\bar t}^{t(\sigma)} L\big(t,x(t,\sigma),\tilde u(t,\sigma)\big)\,dt$$

where the integral is taken along the trajectory σ *of* N *passing through the point* $(\bar t,\bar\sigma)$ *and cutting the locus* $t=t(\sigma)$, $\sigma\in G$.

This is a direct consequence of the integration of (4.8) along σ in the interval $[\bar t,t(\sigma)]$.

5. The practical method of constructing an optimal feedback control

In the preceding sections we gave necessary and sufficient conditions for the independence of the Hilbert integral

$$\int \bar y(t,x)\,dx+\big(L\,(t,x,\bar u(t,x))-\bar y\,(t,x)\,f\,(t,x,\bar u(t,x))\big)\,dt \tag{5.1}$$

of the (t,σ) path of integration for given ends. They are as follows: each induced triplet (σ,u^*,y^*) satisfies the induced maximum principle and at some point of this triplet the functions ψ_{ij}, φ_i vanish. All that was true by the assumption that there exist a feedback control $\bar{u}(t, x)$, $(t, x)\in D_{\bar{u}}$ and a function $\bar{y}(t, x)$, $(t, x)\in D_{\bar{u}}$.

In this section we give the method of a practical construction of both functions $\bar{u}$ and $\bar{y}$.

To this effect, we shall describe a special family of induced triplets. Thus, let us define on an open simply connected set $G\subset R^n$ a pair of C^1-functions $t^-(\sigma)$, $t^+(\sigma)$, $t^-(\sigma)<t^+(\sigma)$, with values in the interval $[0, 1]$ and such that, for some $\sigma^0\in G$, $t^-(\sigma^0)=0$, $t^+(\sigma^0)=1$.

Denote by Z^-, Z, Z^+ the sets of pairs (t, σ) where $\sigma\in G$ and t satisfies, respectively, the conditions

$$0\leqslant t^-(\sigma)=t,\ t^-(\sigma)<t<t^+(\sigma),\ t=t^+(\sigma)\leqslant 1\,. \tag{5.2}$$

The notation $[Z]$ will be used for the union of the sets Z^-, Z, Z^+.

Now, we consider a family Σ of pairs (x, u) depending on a parameter σ, which satisfy the maximum principle (comp. 3. § 1), given by the functions

$$x(t, \sigma)\,,\ u(t, \sigma)\qquad (t, \sigma)\in Z\,. \tag{5.3}$$

Here σ is the parameter which distinguishes a member of the family, i.e. σ remains constant on each member of Σ, and this member is then defined in the open interval $t^-(\sigma)<t<t^+(\sigma)$.

Further, we consider the set $\tilde{G}\subset R^{2n}$ of (σ, ρ) and suppose that the set G is a projection of the set $\tilde{G}$ in the following sense:

$$\left\{\begin{array}{l}\text{Given any point } (\sigma^1, \rho^1)\in\tilde{G} \text{ and any sufficiently small open neigh-}\\ \text{bourhood } Q\subset G \text{ of } \sigma^1, \text{ there exists in } Q \text{ a } C^2\text{-function } \rho(\sigma) \text{ such}\\ \text{that } \rho(\sigma^1)=\rho^1 \text{ and that all points of the form } [\sigma, \rho(\sigma)] \text{ for } \sigma\in Q\\ \text{lie in } \tilde{G}.\end{array}\right. \tag{5.4}$$

Similarly as above, we denote by Z^{*-}, Z^*, Z^{*+} the sets of (t, σ, ρ) for which t is subject to respective conditions (5.2), and $(\sigma, \rho)\in\tilde{G}$. We write $[Z^*]$ for the union of the three sets. We shall further denote by Σ^* a family of canonical triplets (x, u, y) which correspond to a member of Σ and which are obtained by giving, with functions (5.3), a further conjugate vector function

$$y(t, \sigma, \rho)\qquad (t, \sigma, \rho)\in Z^*\,. \tag{5.5}$$

The parameter ρ, occurring in (5.5), distinguishes the corresponding canonical triplets. It appears here since, for a conjugate function which satisfies differential equation (1.4) (i), we have no additional boundary conditions.

The definitions of the functions $x(t, \sigma)$, $y(t, \sigma, \rho)$ will be supposed extended to the sets $[Z]$, $[Z^*]$. This means defining them for $t=t^+(\sigma)$ and $t=t^-(\sigma)$, where the values of x, y correspond to the end points of our members. The sets of pairs (t, x), where $x=x(t, \sigma)$ with (t, σ) belonging to Z^-, Z, Z^+, $[Z]$, will be denoted by D^-, D, D^+, $[D]$, respectively.

Moreover, we suppose the following conditions satisfied:

$$\left\{\begin{array}{ll} \text{(i)} & \text{For the functions } \tilde{L}(t,\sigma)=L\left(t, x(t,\sigma), u(t,\sigma)\right),\ \tilde{f}(t,\sigma)= \\ & =f\left(t, x(t,\sigma), u(t,\sigma)\right), \text{ there exist continuous derivatives } \tilde{L}_\sigma, \\ & \tilde{L}_{\sigma\sigma}, \tilde{f}_\sigma, \tilde{f}_{\sigma\sigma} \text{ in } [Z] \text{ and } \dfrac{\partial}{\partial\sigma} L\left(t, x, u(t,\sigma)\right), \dfrac{\partial}{\partial\sigma} f\left(t, x, u(t\ \sigma)\right) \\ & \text{for each fixed } (t,x)\in D, \text{ in } Z \text{ and they satisfy at } (t,x),\ x=x(t,\sigma), \\ & \text{the relations: } \dfrac{\partial\tilde{L}}{\partial\sigma}=\dfrac{\partial}{\partial\sigma} L\left(t, x, u(t,\sigma)\right)+L_x\left(t, x, u(t,\sigma)\right)\times \\ & \times x_\sigma(t,\sigma),\ \dfrac{\partial\tilde{f}}{\partial\sigma}=\dfrac{\partial}{\partial\sigma} f\left(t, x, u(t,\sigma)\right)+f_x\left(t, x, u(t,\sigma)\right) x_\sigma(t,\sigma). \\ \text{(ii)} & \text{The function } \tilde{y}(t,\sigma)=y(t,\sigma,\rho(\sigma)) \text{ is } C^2 \text{ in } [Z]. \\ \text{(iii)} & \text{The function } x(t,\sigma) \text{ is } C^2 \text{ in } [Z]. \\ \text{(iv)} & \det(x_\sigma)\neq 0 \text{ in } Z^-\cup Z \text{ and through each point of } D^-\cup D \text{ there} \\ & \text{passes one and only one trajectory } x \text{ of } \Sigma. \end{array}\right. \tag{5.6}$$

By assumptions (5.6) (iii) and (iv), the mapping

$$(t,\sigma)\to\left(t, x(t,\sigma)\right): Z^-\cup Z\to D^-\cup D \tag{5.7}$$

is a C^2-diff of $Z^-\cup Z$ onto $D^-\cup D$ with the inverse C^2-diff

$$\theta: (t,x)\to(t,\sigma(t,x)): D^-\cup D\to Z^-\cup Z\,. \tag{5.8}$$

For $(t,x)\in D^-\cup D$, let us set

$$\bar{u}(t,x)=u\left(t,\sigma(t,x)\right), \tag{5.9}$$

$$\bar{y}(t,x)=y\left(t,\sigma(t,x),\rho(\sigma(t,x))\right)=\tilde{y}\left(t,\sigma(t,x)\right). \tag{5.10}$$

Of course, (5.10) is defined only locally. We extended $\bar{u}(t,x)$ and $\bar{y}(t,x)$ to the set D^+ taking there, for $\bar{u}(t,x)$ and $\bar{y}(t,x)$, any value of $u(t^+(\sigma),\sigma)$ and $y(t^+(\sigma),\sigma,\rho(\sigma))$, respectively, for that σ for which $t=t^+(\sigma)$, $x=x(t^+(\sigma),\sigma)$. It is clear that $\bar{u}(t,x)$, $\bar{y}(t,x)$ satisfy (2.5) in $[D]$.

Now, we note that map (5.7) may be considered as transformation (3.2) and that the $x(t,\sigma)$, $\bar{y}\left(t,x(t,\sigma)\right)=y\left(t,\sigma,\rho(\sigma)\right)=\tilde{y}(t,\sigma)$ defined here satisfy the assumptions about them made in § 4, i.e. (4.2). Moreover, in (3.2′) we now set $\dot{\sigma}=0$, so that (3.3′) takes the form

$$x_t(t,\sigma)=f\left(t,x(t,\sigma),u(t,\sigma)\right), \tag{5.11}$$

and (3.12), since $\det(x_\sigma)\neq 0$, takes the form

$$y_t\left(t,\sigma,\rho(\sigma)\right)=-y\left(t,\sigma,\rho(\sigma)\right)f_x\left(t,x(t,\sigma),u(t,\sigma)\right)+ \\ +L_x\left(t,x(t,\sigma),u(t,\sigma)\right). \tag{5.12}$$

Hence we conclude that the induced triplets (σ,u^*,y^*), where $\sigma=$const, $u^*=\bar{u}\left(t,x(t,\sigma)\right)=u(t,\sigma)$, $y^*=\bar{y}\left(t,x(t,\sigma)\right)=y(t,\sigma,\rho(\sigma))$, satisfy the induced maximum principle.

Suppose further that the Hilbert differential $\bar{y}dx+(L-\bar{y}f)\,dt$ is an exact differential in the variable σ on the locus $t=t^+(\sigma)$, $\sigma\in G$, i.e. in Z^+. By theorem 4.2 and note 1 following it, the fuctnions ψ_{ij}, φ_i vanish in $[Z]$. Thus, Hilbert's integral (5.1) is dependent of the (t, σ) path of integration lying in $[Z]$ for given ends (see note 2 § 4).

Families of pairs (x, u) satisfying the maximum principle (having form (5.3) they satisfy conditions (5.6)), for which Hilbert's integral (5.1) is independent of the (t, x) path joining in $D^- \cup D$ two points of $D^- \cup D$, will be called sprays of extremal pairs, and the corresponding families of cannonical triplets (x, u, y) — canonical sprays of extremal triplets. We state an existence theorem for canonical sprays.

THEOREM 5.1. *If the family Σ^*, described above, satisfies conditions* (5.6), *then Hilbert's integral* (5.1) *with $\bar{u}, \bar{y}$ defined as in* (5.9), (5.10) *is independent of a rectifiable path in $D^- \cup D$ joining two points of $D^- \cup D$ if and only if the line integral*

$$\int L\big(t, x(t, \sigma), u(t, \sigma)\big)\,dt+y\big(t, \sigma, \rho(\sigma)\big)\,x_\sigma\,d\sigma \tag{5.13}$$

is independent of a rectifiable path in $Z^- \cup Z$ joining two points in $Z^- \cup Z$.

Proof. We shall represent integral (5.1) as a line integral (5.13). By virtue of the diff θ (5.8), of $D^- \cup D$ onto $Z^- \cup Z$, the Hilbert integral over a rectifiable curve γ in $D^- \cup D$ equals a line integral of form (5.13) over the arc $\Gamma=\theta(\gamma)$ in $Z^- \cup Z$ provided one sets $x=x(t, \sigma)$ in the coefficients of dx and dt in (5.1) and

$$dx=x_t(t, \sigma)\,dt+x_\sigma(t, \sigma)\,d\sigma\,. \tag{5.14}$$

Subject to the diff θ (5.8) and relations (5.11), (5.14).,

$$J(\gamma)=\int_\gamma \bar{y}(t, x)\,dx+\Big(L\big(t, x, \bar{u}(t, x)\big)-\bar{y}\big(t, x\big)f\big(t, x, \bar{u}(t, x)\big)\Big)\,dt=$$

$$=\int_\Gamma L\big(t, x(t, \sigma), u(t, \sigma)\big)\,dt+y\big(t, \sigma, \rho(\sigma)\big)\,x_\sigma\,d\sigma\,. \tag{5.15}$$

We have to remember that $\rho(\sigma)$ is defined only locally. Thus, each curve Γ must be divided into a finite number of arcs on which $\rho(\sigma)$ exists and so, in (5.15) we have such sums on both sides.

Theorem 5.1 is an immediate consequence of relation (5.15). ∎

Relation (5.15) has another important consequence.

LEMMA 5.1. *If γ is a subarc of an extremal trajectory of the spray Σ, then $J(\gamma)=\int_\gamma L(t, x, u_\gamma)\,dt$ where $J(\gamma)$ is defined on the left-hand side of (5.15) and u_γ is the control corresponding to the extremal trajectory containing γ.*

Proof. In the coordinates t, σ the arc γ has a representation $x=x(t, \sigma)$ with σ constant. For this arc γ, one must set $d\sigma=0$ in the right member of (5.15). If $t\to\gamma(t)=x$ is a t-parametrization of γ with $t\in[t_1, t_2]\subset[0, 1]$, then (5.15) shows that

$$J(\gamma)=\int_{t_1}^{t_2} L\big(t, x(t, \sigma), u(t, \sigma)\big)\, dt\,.$$

Setting $u_\gamma=u(t, \sigma)$, we have $J(\gamma)=\int_{t_1}^{t_2} L\big(t, \gamma(t), u_\gamma(t)\big)\, dt=\int_\gamma L(t, x, u)\, dt$. ∎

NOTE 1. It is clear that if $t^+(\sigma)=1$, $x(t^+(\sigma), \sigma)=e$ or $y\big(t^+(\sigma), \sigma, \rho\big)=y^0$ for all $\sigma\in G$ or $(\sigma, \rho)\in\tilde{G}$, then Σ^* is the canonical spray.

NOTE 2. If Σ^* is the canonical spray, then the $\bar{u}(t, x)$, $(t, x)\in[D]$, of (5.9) is the optimal feedback control and $\bar{y}(t, x)$, $(t, x)\in[D]$ of (5.10) is the function $S_x(t, x)$ of § 2.

NOTE 3. If the canonical spray Σ^* exists, then the function $S(t, x)$, $(t, x)\in[D]$, (considered in §§ 1 and 2) exists and may be defined in $[D]$ by

$$S(t, x)=-\int_t^{t^+(\sigma^*)} L\big(\tau, x(\tau, \sigma^*), \bar{u}\big(\tau, x(\tau, \sigma^*)\big)\big)\, d\tau$$

where $x(\tau, \sigma^*)$ is the trajectory of Σ passing through the point $(t, x)\in[D]$ defined in $[t, t^+(\sigma^*)]$ (comp. Corollary 4.1), and its exact derivative is equal to the integrand of (5.1).

NOTE 4. We do not make any assumptions about either continuity or even measurability of the $u(t, \sigma)$ and, in consequence, of the feedback $\bar{u}(t, x)$. It is essential in practice.

The following theorem is an extension of the Weierstrass sufficiency theorem from the calculus of variations.

THEOREM 5.2. (*Sufficiency Theorem*). *Suppose that there exists a canonical spray Σ^* with the optimal feedback control $\bar{u}(t, x)$, $(t, x)\in[D]$, the function $\bar{y}(t, x)$ in $[D]$ and the pair (x^0, u^0) of M being a member of the spray Σ. Then the pair (x^0, u^0) affords a minimum to $I(x, u)$ (see (1.2)) relative to those pairs (x, u) of M whose graphs of trajectories of x lie in $[D]$ (i.e. relative to $M_{D\bar{u}}$).*

Proof. Denote by γ^0, γ the curves corresponding to the pairs (x^0, u^0) and (x, u) of M, where the graph of x lies in $[D]$, and set $\bar{I}(\gamma)=\bar{I}(x, u)$. Let $J(\gamma)$ be the Hilbert integral defined on the left-hand side of (5.15). According to lemma 5.1, $J(\gamma^0)=$

$=\bar{I}(\gamma^0)$. Since Σ^* is a canonical spray, $J(\gamma)=J(\gamma^0)$, so that $\bar{I}(\gamma^0)=J(\gamma)$. Thus $\bar{I}(\gamma^0)$ is the integral (5.1) taken along γ. Explicitly,

$$\bar{I}(\gamma^0)=\int_0^1 \Big\{L\big(t, x(t), \bar{u}(t, x(t))\big)+\bar{y}(t, x(t))\,\big(f(t, x(t), u(t))-$$
$$-f\big(t, x(t), \bar{u}(t, x(t))\big)\big)\Big\}\, dt\,.$$

Since $\bar{I}(\gamma)=\int_0^1 L(t, x(t), u(t))\, dt$, we are led to the extension of the Weierstrass formula

$$\bar{I}(\gamma)-\bar{I}(\gamma^0)=\int_0^1 \Big\{H\big(t, x(t), \bar{u}(t, x(t)), \bar{y}(t, x(t))\big)-$$
$$-H\big(t, x(t), u(t), \bar{y}(t, x(t))\big)\Big\}\, dt$$

where $H(t, x, u, \bar{y})$ is defined in § 2. By virtue of the definition of a feedback control (def. 2.1.), compare also (2.6),

$$\bar{I}(\gamma)-\bar{I}(\gamma^0)\geqslant 0\,. \tag{5.16}$$

This completes the proof. ∎

NOTE 5. If in (1.1) $f(t, x, u)=u$ and $U=R^n$, then problem (1.2) becomes the standard problem from the calculus of variations and $\bar{u}(t, x)$ is a geodesic slope or field slope, $\bar{y}(t, x)=L_u(t, x, \bar{u}(t, x))$, (5.16) is the integral from the Weierstrass E-condition

$$E=L(t, x, u)-L(t, x, \bar{u})-(u-\bar{u})\, L_u(t, x, \bar{u})\geqslant 0\,, \tag{5.17}$$

and the spray Σ is a geodesic family or a field of extremals.

NOTE 6. In many books on the calculus of variations a field of extremals is defined as the one which induces only a line independent Hilbert integral (5.1), but then curves that define the field satisfy only the Euler equation and so, in a sufficiency theorem there must occur E-condition (5.17). In optimization theory members of Σ satisfy the maximum principle and thus, also a suitable inequality (5.17). Hence the last inequality does not appear in our sufficiency theorem.

Now, we give a simple example to explain the above theory.

EXAMPLE: Let $U=[-1, 1]$; admissible controls are measurable functions $u\colon [0, 1]\to$ $\to[-1, 1]$; admissible trajectories are absolutely continuous functions $x\colon [0, 1]\to R$ satisfying

$$\dot{x}(t)=t\,\big(u(t)\big)^2\,. \tag{5.17$'$}$$

We find a minimum of the integral

$$I(x, u)=\int_0^1 \big(x(t)-\big(u(t)\big)^2\big)\, dt \tag{5.18}$$

from among all admissible pairs (x, u) whose trajectories satisfy $x(0)=0$, $x(1)=1/2$.

First, we calculate triplets (x, u, y) which satisfy the maximum principle. To this effect, we set $\tilde{H}(t, x, u, y) = -L(t, x, u) + yf(t, x, u) = -x + u^2 + ytu^2$ and $H(t, x, y) = \max_{-1 \leqslant u \leqslant 1} \tilde{H}(t, x, u, y)$. We know (comp. 3. of § 1) that $\dot{y} = -H_x$. In our case this gives $\dot{y} = 1$. Hence $y(t, \alpha) = t + \alpha$. Of course, $y(t, \alpha)$ are independent of x and u, so we can take $\alpha = -1$, i.e. $y(t) = t - 1$, $t \in [0, 1]$. We easily check that $\tilde{H}$ attains its, maximum at $u = 1$ or $u = -1$ for each $t \in [0, 1]$, $x \in R$. Thus $x(t, \sigma) = \frac{1}{2} t^2 + \sigma$, $t \in [0, 1]$, $\sigma \in R$ and, for $\sigma = 0$, we have the suspected trajectory $x^0(t) = x(t, 0) = \frac{1}{2} t^2$.

We can now define the families Σ and Σ^*. The set G considered above is now equal to R. The functions $t^-(\sigma)$, $t^+(\sigma)$ are constant and equal to 0 and 1, respectively. $[Z] = Z^- \cup Z \cup Z^+ = \{(0, \sigma) : \sigma \in R\} \cup \{(t, \sigma) : 0 < t < 1, \sigma \in R\} \cup \{(1, \sigma) : \sigma \in R\}$. The family Σ is given by the functions

$$x(t, \sigma) = \frac{1}{2} t^2 + \sigma, \quad u(t, \sigma) = 1 \quad (t, \sigma) \in [Z]. \tag{5.19}$$

Similarly, $[Z^*] = Z^{*-} \cup Z^* \cup Z^{+*} = \{(0, \sigma, \rho) : (\sigma, \rho) \in \tilde{G}\} \cup \{(t, \sigma, \rho) : 0 < t < 1, (\sigma, \rho) \in \tilde{G}\} \cup \{(1, \sigma, \rho) : (\sigma, \rho) \in \tilde{G}\}$ where $\tilde{G} = G \times R = R^2$ and Σ^* is obtained by adjoining to functions (5.19) the conjugate function

$$y(t, \sigma, \rho) = t - 1 \quad (t, \sigma, \rho) \in [Z^*]. \tag{5.20}$$

Of course, for the function $\rho(\sigma)$ defined in (5.4) we may take $\rho(\sigma) = \sigma$. It is easy to check that all assumptions (5.6) are satisfied here. Hence the diff θ of (5.8) has the form $\theta : (t, x) \to (t, x - \frac{1}{2} t^2)$, and

$$\begin{aligned} \bar{u}(t, x) &= 1 \quad \text{in } [D], \\ \bar{y}(t, x) &= t - 1 \quad \text{in } [D], \end{aligned} \tag{5.21}$$

where $[D] = \{(t, x) : x = \frac{1}{2} t^2 + \sigma, (t, \sigma) \in [Z]\} = \{(t, x) : 0 \leqslant t \leqslant 1, x \in R\}$. The sets D^-, D, D^+ are defined analogously.

Finally, by (5.20), we obtain that integral (5.13) is equal to zero for all rectifiable paths in Z^+, so that it is independent of them there.

Thus, all assumptions of the definition of a canonical spray are satisfied, hence our family Σ^* is the canonical spray, and, by the sufficiency theorem, the pair $(x^0, \mathring{u}$ where $x^0(t) = 1/2\, t^2$, $u^0(t) = 1$, $t \in [0, 1]$, gives the global minimum to (5.18). It is evident that the $\bar{u}(t, x)$ of (5.21) is the optimal feedback control.

References

[1] Caratheodory C. Calculus of Variations and Partial Differential Equations of First Order. San Francisco, Holden-Day, 1965.

[2] CESARI L. Optimization Theory and Applications. Springer-Verlag, 1983.

[3] MAYNE D.Q. Differential Dynamic Programming: A Unified Approach to Optimal Control. Control and Dynamic Systems. Vol. 10. Edited by C.T. Leondes. New York, Academic Press, 1973.

[4] MAYNE D.Q. Sufficient Conditions for a Control to Be a Strong Minimum. *Journal of Optimization Theory and Applications*, **21** (1977).

[5] YOUNG L.C. Calculus of Variations and Optimal Control Theory. Philadelphia, Saunders, 1969.

Received, April 1983.

Warunki konieczne wyznaczania minimum w klasycznym zadaniu sterowania optymalnego

Stosując klasyczny rachunek wariacyjny podano praktyczną metodę obliczania optymalnego sterowania ze sprzężeniem zwrotnym. W rezultacie otrzymano warunki dostateczne Weierstrassa wyznaczania minimum funkcjonału.

Необходимые условия определения минимума в классической задаче оптимального управления

Используя классическое вариационное исчисление, дается практический метод вычисления оптимального управления с обратной связью. В результате получены достаточные условия Веерштрасса для определения минимума функционала.