Матем. Сборник
Том 182 (1991), № 1

Math. USSR Sbornik
Vol. 72 (1992), No. 1

SYMPLECTIC GEOMETRY AND NECESSARY CONDITIONS FOR OPTIMALITY

UDC 517.97

A. A. AGRACHEV AND R. V. GAMKRELIDZE

ABSTRACT. With the help of a symplectic technique the concept of a field of extremals in the classical calculus of variations is generalized to optimal control problems. This enables us to get new optimality conditions that are equally suitable for regular, bang-bang, and singular extremals. Special attention is given to systems of the form $\dot{x} = f_0(x) + uf_1(x)$ with a scalar control. New pointwise conditions for optimality and sufficient conditions for local controllability are obtained as a consequence of the general theory.

§1. AN INVARIANT FORMULATION OF THE MAXIMUM PRINCIPLE

We consider the time-optimal problem on a smooth manifold M for the system

$$\dot{x} = f(x, u), \qquad x \in M, \ u \in U \subset \mathbf{R}^r; \tag{1}$$

here M is a manifold of the class C^∞, $\dim M = n$, and $f(x, u) \in T_xM$ is an infinitely differentiable function of (x, u).

Consider the cotangent bundle

$$T^*M = \{(\psi, x) | x \in M, \ \psi \in T_x^*M\}$$

over M, and denote by dx the differential of the projection $(\psi, x) \mapsto x$ of the manifold T^*M onto M. Then the composition $\psi dx|_{(\psi, x)}$ is a cotangent vector to T^*M at the point (ψ, x). We define the differential 1-form Ω on T^*M by the equality $\Omega|_{(\psi, x)} = \psi dx|_{(\psi, x)}$ and let $\omega = -d\Omega$. The closed differential 2-form ω is the standard symplectic structure on T^*M.

To each function $\varphi \in C^\infty(T^*M)$ (Hamiltonian) there corresponds a Hamiltonian vector field $\boldsymbol{\varphi}$ on T^*M uniquely determined by the relation

$$d\varphi = i_{\boldsymbol{\varphi}}\omega,$$

where i_ξ is the interior multiplication by the field ξ, $i_\xi\omega(\eta) = \omega(\xi, \eta)$.

Let $\tilde{x}(t)$ and $\tilde{u}(t)$ be a solution of (1). We say that this solution satisfies the Pontryagin maximum principle if there exists a curve $\tilde{\psi}(t) \in T^*_{\tilde{x}(t)}M$, $\tilde{\psi}(t) \neq 0$, such that the pair $(\tilde{\psi}(t), \tilde{x}(t))$ satisfies the Hamiltonian system corresponding to the Hamiltonian $\widetilde{H}_t(\psi, x) = \psi f(x, \tilde{u}(t))$, along with the maximum condition

$$0 \leq \widetilde{H}_t(\tilde{\psi}(t), \tilde{x}(t)) = \max_{u \in U} \tilde{\psi}(t) f(\tilde{x}(t), u), \qquad t \geq 0.$$

1991 *Mathematics Subject Classification*. Primary 49K15, 58F05.

Such pairs $(\tilde{\psi}(t), \tilde{x}(t))$ are called extremals of the time-optimal problem for system (1).

§2. The perturbation equation for (1)

Let $P_t: M \to M$ be a flow (nonstationary) on M satisfying the equation

$$\frac{\partial}{\partial t} P_t(x) = f(P_t(x), \tilde{u}(t)), \qquad P_0 = \mathrm{id}.$$

Making the nonstationary change of variables $y = P_t^{-1}(x)$ in (1), we get the system

$$\dot{y} = g_t(y, u), \tag{2}$$

where

$$g_t(\cdot, u) = P_{t*}^{-1}(f(\cdot, u) - f(\cdot, \tilde{u}(t))).$$

We fix the initial condition $x_0 = \tilde{x}(0)$ we have that

$$P_t^{-1}(\overline{x}(t)) \equiv x_0, \qquad g_t(y, \tilde{u}(t)) \equiv 0 \quad \forall y \in M.$$

Note also that a solution $(x_0, \tilde{u}(t))$ of (2) satisfies the "simplified" maximum condition

$$\psi_0 g_t(x_0, u) \leq 0 \quad \forall u \in U, \quad \psi_0 = \tilde{\psi}(0).$$

Thus, the study of the behavior of system (1) near a given solution $(\tilde{x}(t), \tilde{u}(t))$ reduces to the study of the "perturbation equation" (2) near an equilibrium state.

§3. Regular extremals

The basic facts of the theory being developed have an especially simple appearance if we confine ourselves to the consideration of regular extremals.

Definition. An extremal $(\tilde{\psi}(t), \tilde{x}(t))$ is said to be *regular* if the corresponding control $\tilde{u}(t)$ is piecewise continuous and takes values in the interior of U, the $r \times r$ matrix

$$\frac{\partial^2}{\partial u^2} \psi_0 f_t(x_0, u)|_{\tilde{u}(t)} \tag{3}$$

is nonsingular for any t, and the equation

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

has a unique solution for $\tilde{u}(t)$.

It is easy to see that (3) is nonsingular if and only if the $r \times r$ matrix

$$\frac{\partial^2}{\partial u^2} \psi_0 g_t(x_0, u)|_{\tilde{u}(t)}$$

is nonsingular for any t. Thus, nondegeneracy of the extremal under consideration enables us to express u as a smooth function of (ψ, x) in the equation

$$\frac{\partial}{\partial u} \psi g_t(x, u) = 0$$

for all $(\psi, x) \in \mathscr{O}$, where $\mathscr{O}$ is some neighborhood of the point (ψ_0, x_0) in T^*M. Let $u_t = u_t(\psi, x)$ be the corresponding expression, and define the Hamiltonian

$$H_t(\psi, x) = \psi g_t(x, u_t(\psi, x)), \qquad H_t(\psi_0, x_0) = 0. \tag{4}$$

The solutions of the corresponding Hamiltonian system form a family of regular extremals of the control system (2). We consider in greater detail the collection of solutions satisfying the initial condition $x(0) = x_0$.

We define the *homogenization* of the cotangent bundle T^*M to be the quotient space $T^*M\backslash M$ with respect to the equivalence relation

$$(\alpha\psi, x) \sim (\psi, x) \quad \forall\alpha > 0.$$

Since the Hamiltonian (4) is homogeneous in ψ with degree of homogeneity 1, the domain of definition $\mathcal{O}$ of the Hamiltonian can be assumed to be conical:

$$(\psi, x) \in \mathcal{O} \Rightarrow (\alpha\psi, x) \in \mathcal{O} \quad \forall\alpha > 0.$$

Denote by Q_t the flow of our Hamiltonian system, and by $\mathcal{L}_0$ a conical neighborhood of the point (ψ_0, x_0) in $T^*_{x_0}M$ such that the family of submanifolds $\mathcal{L}_t = Q_t(\mathcal{L}_0)$ is defined for all t under consideration.

Since the fiber $T^*_{x_0}M$ is a Lagrangian submanifold of T^*M, it follows that the submanifolds $\mathcal{L}_t$ are also Lagrangian submanifolds. It is easy to see that all the submanifolds $\mathcal{L}_t$ are conical and that their homogenizations are well defined. If we denote the homogenization operator by $\mathcal{H}$, then

$$\mathcal{HL}_t = \mathcal{H}(Q_t\mathcal{L}_0) = Q_t(\mathcal{HL}_0).$$

The geometric properties of the projection $(\mathbf{R}_+\psi, x) \mapsto x$ of the manifold $\mathcal{HL}_t$ onto the base M turn out to be closely related to the property of local optimality for the trajectory $\tilde{x}(t)$.

We say that the regular extremal under consideration is *strongly regular* if the matrix

$$\int_{t_1}^{t_2} \frac{\partial g_t}{\partial u}\frac{\partial g_t^*}{\partial u}\,dt\Big|_{(x_0, \tilde{u})}$$

has rank $n-1$ (i.e., the maximal possible rank) for all $t_1 \neq t_2$. In this case ψ_0 is determined uniquely up to a positive factor by the condition

$$\psi_0 g_t(x_0, u) \leq 0 \quad \forall t, u.$$

The so-called regular case of the classical calculus of variations with $r = n-1$ obviously satisfies this condition.

If the extremal under consideration is strongly regular, then for sufficiently small $\tau > 0$ the projection $\mathcal{HL}_\tau \to M$ is an immersion at the point $(\mathbf{R}\psi_0, x_0)$, and there is a simple necessary condition for optimality on the given interval $0 \leq \tau \leq t$:

If the solution $\tilde{u}(\tau)$, $\tilde{x}(\tau)$, $0 \leq \tau \leq t$, is optimal, then the projection $\mathcal{HL}_\tau \to M$ is an immersion at the point $(\mathbf{R}\psi_0, x_0)$ $\forall\tau \in (0, t)$.

This assertion is a direct generalization of the field theory in the classical calculus of variations.

The assertion "$\mathcal{HL}_\tau \to M$ is an immersion" is equivalent to the Lagrange subspace $T_{(\psi_0, x_0)}\mathcal{L}_\tau$ intersecting the "vertical" Lagrangian subspace $\Pi_0 = T_{(\psi_0, x_0)}(T^*_{x_0}M)$ only in the line $T_{(\psi_0, x_0)}(\mathbf{R}\psi_0)$. We introduce the abbreviated notation $\Lambda(\tau) = T_{(\psi_0, x_0)}\mathcal{L}_\tau$ and $\Sigma = T_{(\psi_0, x_0)}(T^*M)$. If strong regularity is lacking, then the immersion condition formulated above is equivalent to the equality $\dim(\Pi_0 \cap \Lambda_\tau) = 1$, which is no longer necessary for the optimality of the control. Nevertheless, taking account of the arrangement of the family of subspaces $\Lambda(\tau)$ with respect to Π_0 in the symplectic space Σ enables us to get nice conditions for optimality also in the general regular case. In the strongly regular situation these conditions reduce to the immersion requirement already considered, and perhaps clarify the meaning of this immersion.

The optimality conditions are formulated in terms of symplectic geometry. We briefly describe the concepts used. Details can be found in [5] and [1].

The skew inner product in the symplectic space Σ is given by the form $\omega|_{(\psi_0, x_0)}$. To each subset $S \subset \Sigma$ corresponds its skew orthogonal complement $S^\perp = \{\xi \in \Sigma | \omega(S, \xi) = 0\}$. The symbol $L(\Sigma)$ denotes the Lagrangian Grassmannian—the collection of all Lagrangian subspaces of Σ, $L(\Sigma) = \{\Lambda \subset \Sigma | \Lambda^\perp = \Lambda\}$. The Lagrangian Grassmannian is a smooth connected $n(n+1)/2$-dimensional manifold. For an arbitrary $\Pi \in L(\Sigma)$ let

$$\mathscr{M}_\Pi = \{\Lambda \in L(\Sigma) | \Lambda \cap \Pi \neq 0\}.$$

The closed subset $\mathscr{M}_\Pi$ of $L(\Sigma)$ is not a smooth submanifold; however, it is a so-called pseudomanifold. More precisely, we have the following easily verified assertions:

1) For $r = 1, \ldots, n$ the subset $\mathscr{M}_n^r = \{\Lambda \in L(\Sigma) | \dim(\Lambda \cap \Pi) = r\}$ is a smooth imbedded submanifold of $L(\Sigma)$ of dimension $n(n+1)/2 - r(r+1)/2$.

2) The decomposition $\mathscr{M}_\Pi = \bigcup_{r=1}^n \mathscr{M}_\Pi^r$ is a Whitney stratification for $\mathscr{M}_\Pi$.

3) The hypersurface $\mathscr{M}_\Pi^1$ in $L(\Sigma)$ has a natural co-orientation:

Let Δ_ε, $\varepsilon \in \mathbf{R}$, be a smooth curve in $L(\Sigma)$ that intersects $\mathscr{M}_\Pi^1$ transversally at the point Δ_0, and let $\lambda_\varepsilon \in \Sigma$ be an arbitrary smooth curve in Σ satisfying the conditions $\lambda_\varepsilon \in \Delta_\varepsilon \ \forall \varepsilon$, $\lambda_0 \in \Delta_0 \cap \Pi$, and $\lambda_0 \neq 0$. One says that the curve Δ_ε has positive index of intersection (equal to $+1$) with $\mathscr{M}_\Pi$ at the point Δ_0 if $\omega(d\lambda/d\varepsilon|_{\varepsilon=0}, \lambda_0) > 0$, and negative index of intersection (equal to -1) with $\mathscr{M}_\Pi$ at the same point if $\omega(d\lambda/d\varepsilon|_{\varepsilon=0}, \lambda_0) < 0$. It is not hard to see this definition is unambiguous, i.e., does not depend on the choice of the curve λ_ε.

These properties enable us to correctly define the index of intersection of $\mathscr{M}_\Pi$ with an arbitrary curve $\Delta(\tau)$, $\alpha \leq \tau \leq \beta$, satisfying the condition $\Delta(\alpha) \cap \Pi = \Delta(\beta) \cap \Pi = 0$.

Indeed, it is possible by a small deformation to turn this curve into a smooth curve $\Delta'(\tau)$ that is transversal to any of the submanifolds M_Π^r, $r = 1, \ldots, n$. The curve $\Delta'(\tau)$ intersects M_Π^1 at finitely many points, and does not intersect $\mathscr{M}_\Pi^r$ at all for $r \geq 2$.

The index of intersection of $\Delta'(\tau)$ with $\mathscr{M}_\Pi$, denoted by $\mathrm{Ind}_\Pi \Delta'(\cdot)$ in what follows, is simply the sum of the indices described in property 3) relating to each point of the intersection of $\Delta'(\tau)$ with $\mathscr{M}_\Pi^1$. Actually, $\mathrm{Ind}_\Pi \Delta'(\tau)$ depends only on $\Delta(\cdot)$, and not on the method of deforming this curve, provided that in the deformation process the endpoints of the curve do not intersect $\mathscr{M}_\Pi$. The proof is based on the fact that the submanifolds $\mathscr{M}_\Pi^r$, $r \geq 2$, have codimension at least three in $L(\Sigma)$, and hence not only a curve in general position, but also a homotopy in general position does not intersect these submanifolds.

It remains to set

$$\mathrm{Ind}_\Pi \Delta(\cdot) = \mathrm{Ind}_\Pi \Delta'(\cdot).$$

The quantity $\mathrm{Ind}_\Pi \Delta(\cdot)$ is called the *Maslov index* of the curve $\Delta(\cdot)$ (with respect to Π).

The tangent spaces $T_\Delta L(\Sigma)$, $\Delta \in L(\Sigma)$, have a remarkable additional structure that was used in part in the definition of the co-orientation of the submanifold $\mathscr{M}_\Pi^1$. Let Δ_ε, $\varepsilon \in \mathbf{R}$, be a smooth curve in $L(\Sigma)$ and λ_ε a smooth curve in Σ, with $\lambda_\varepsilon \in \Delta_\varepsilon \ \forall \varepsilon \in \mathbf{R}$. It is not hard to show that the quantity $\frac{1}{2}\omega(d\lambda/d\varepsilon|_{\varepsilon=0}, \lambda_0)$ depends only on $\lambda_0 \in \Lambda_0$ and $d\Lambda/d\varepsilon|_{\varepsilon=0} \in T_{\Lambda_0} L(\Sigma)$, and not on the choice of the curve λ_ε.

Computing this quantity for each $\lambda_0 \in \Lambda_0$, we get a quadratic form on Λ_0:

$$\lambda_0 \mapsto \frac{1}{2}\omega\left(\left.\frac{d\lambda}{d\varepsilon}\right|_{\varepsilon=0}, \lambda_0\right), \qquad \lambda_0 \in \Lambda_0. \tag{5}$$

Thus, to each tangent vector $d\Delta/d\varepsilon \in T_{\Delta_0} L(\Sigma)$ there corresponds a quadratic form

(5) on Δ_0. This correspondence is a linear isomorphism of the space $T_{\Delta_0}L(\Sigma)$ onto the space of real quadratic forms on Δ_0. The isomorphism is natural, i.e., does not use any additional structures on Σ besides the symplectic structure, and below we identify without special mention the tangent vectors in $T_{\Delta_0}L(\Sigma)$ with the corresponding quadratic forms on Δ_0. The cone of nonnegative forms determines a partial order in the space of quadratic forms on Δ_0. Using the natural isomorphism, we get a partial order on $T_{\Delta_0}L(\Sigma)$: for $\xi \in T_{\Delta_0}L(\Sigma)$ the expression $\xi \geq 0$ means that the corresponding quadratic form on Δ_0 is nonnegative.

An absolutely continuous curve $\Delta(\tau)$, $\tau \in \mathbf{R}$, in $L(\Sigma)$ is said to be nondecreasing if $d\Delta_\tau/d\tau \geq 0$ at every point of differentiability of the curve Δ_τ. It follows from the definition of the natural co-orientation of the hypersurfaces $\mathscr{M}_\Pi^1$ that $\mathrm{Ind}_\Pi \Delta(\cdot) \geq 0$ for any nondecreasing curve $\Delta(\tau)$ in $L(\Sigma)$ and any $\Pi \in L(\Sigma)$ transversal to the endpoints of this curve.

Let $q_t \colon \Sigma \to \mathbf{R}$ be a nonstationary quadratic Hamiltonian on Σ (i.e., a family of quadratic forms on Σ dependent on $t \in \mathbf{R}$), and $Q_t \colon \Sigma \to \Sigma$ the corresponding linear Hamiltonian flow. Let $\Delta \in L(\Sigma)$. Then $Q_t(\Delta) \in L(\Sigma)\ \forall t$, and hence $(d/dt)Q_t(\Delta)$ is a quadratic form on $Q_t(\Delta)$. It can be established by direct verification that

$$dQ_t(\Delta)/dt = q_t|_{Q_t(\Delta)}.$$

In particular, if the quadratic Hamiltonian is nonnegative, $q_t \geq 0$, then all the trajectories of the flow $\Delta \mapsto Q_t(\Delta)$, $t \in \mathbf{R}$, on $L(\Sigma)$ are nondecreasing, and the Maslov index of any segment of any trajectory with respect to any $\Pi \in L(\Sigma)$ (transversal to the endpoints of the segment) is nonnegative.

All this has a direct relation to the curve $\Lambda(\tau) = T_{(\psi_0, x_0)}\mathscr{L}_\tau$ of interest to us, which will be called the *Jacobian curve of the extremal* $(\tilde{\psi}(\tau), \tilde{x}(\tau))$ in what follows.

Proposition 1. *The Jacobian curve* $\Lambda(\tau)$, $0 \leq \tau \leq t$, *is nondecreasing.*

Indeed, let $\mathscr{H}_\tau \colon \Sigma \to \Sigma$, $\tau \in [0, t]$, be the linear Hamiltonian flow on Σ determined by the quadratic Hamiltonian $d^2_{(\psi_0, x_0)}H_\tau$. Then $\Lambda_\tau = \mathscr{H}_\tau(\Pi_0)$, $0 \leq \tau \leq t$. At the same time,

$$d^2_{(\psi_0, x_0)}H_\tau \colon \xi \mapsto -u'(\xi)^* \left(\frac{\partial^2}{\partial u^2}\psi g_\tau|_{(x_0, \tilde{u}(\tau))}\right) u'(\xi), \qquad \xi \in \Sigma,$$

where $u' \colon \Sigma \to \mathbf{R}^2$ is the differential of the vector-valued function $u(\psi, x)$ at the point (ψ_0, x_0). ∎

We continue the consideration of our regular but perhaps not strongly regular extremal. Let

$$\Psi_t = \left\{\psi \in T^*_{x_0}M \mid \psi f(x_0, \tilde{u}(0)) \geq 0,\ \frac{\partial}{\partial u}\psi g_\tau|_{x_0, \tilde{u}(\tau)} = 0,\right.$$
$$\left.\frac{\partial^2}{\partial u^2}\psi g_\tau|_{(x_0, \tilde{u}(\tau))} < 0\right\}, \qquad 0 \leq \tau \leq t.$$

It is easy to see that Ψ_t is a convex cone, and

$$\dim \Psi_t = n - \operatorname{rank} \int_0^t \frac{\partial g_\tau}{\partial u}\frac{\partial g_\tau^*}{\partial u}\bigg|_{x_0, \tilde{u}(\tau)} d\tau.$$

Moreover, $(\psi, x_0) \in \mathscr{L}_\tau\ \forall \psi \in \Psi_t$, $0 \leq \tau \leq t$. Let $\widehat{\Psi}_t = \operatorname{span}\Psi_t$. Identifying the tangent space $T_{(\psi, x_0)}T^*_{x_0}M$ with $T^*_{x_0}M$ in the standard way, we get that $T_{(\psi, x_0)}\mathscr{L}_\tau \supset \widehat{\Psi}_t\ \forall \psi \in \Psi_t$, $\tau \in [0, t]$. Using the last inclusion, we introduce the subspaces

$$\Lambda_\psi(\tau) = T_{(\psi, x_0)}\mathscr{L}_\tau/\widehat{\Psi}_t, \qquad \psi \in \Psi_t,\ \tau \in [0, t],$$

which are Lagrangian subspaces of the symplectic space $\widehat{\Psi}_t^{\perp}/\widehat{\Psi}_t$. Finally, we write $\Pi_0 = (T^*_{x_0}M/\widehat{\Psi}_t)$ and note that $\Pi_0 = \Lambda_\psi(0)$, $\psi \in \Psi_t$.

Theorem 1. *If the control* $\tilde{u}(\tau)$, $0 \le \tau \le t$, *is locally optimal on the interval* $[0, t]$, *then*

$$\overline{\lim_{\Pi \to \Pi_0}} \min\{\mathrm{Ind}_\Pi \Lambda_\psi(\cdot) | \psi \in \Psi_t, \Pi_0, \Lambda_\psi(t) \notin \mathscr{M}_\Pi\} < n. \tag{6}$$

The proof is based on results announced in [2] and [3] and proved in [4] and [1]. The simplest situation is in the case $\dim \Psi_t = 1$. In this case Ψ_t is a half-line, inequality (6) simplifies sharply, and the local optimality condition is "almost sufficient". But if $\dim \Psi_t > 1$, then (6) suffices only for a so-called quasiextremal (see [2] and [4]), while the stronger local optimality conditions use more detailed information about the dependence of $\mathrm{Ind}_\Pi \Lambda_\psi(\cdot)$ on $\psi \in \Psi_t$ (see [1]), and a discussion of them goes beyond the framework of this paper.

It should also be noted that, although the definition of the Maslov index given above is geometrically clear and convenient for certain theoretical considerations, it is not very suitable for computations. For nondecreasing curves we later (in §6) give explicit formulas that require neither reducing to general position nor finding the points of intersection of the curves with $\mathscr{M}_\Pi$.

§4. Systems affine in the control

We considered regular extremals first of all in order to describe the necessary concepts in a relatively simple situation, and to specify the direction of the investigations. The main goal of this paper is nonregular extremals. If the extremal under consideration is nonregular, then the maximum principle

$$\psi g_t(x, u) = \max_{v \in U} \psi g_t(x, v)$$

does not enable us to express u as a smooth function of (ψ, x), not even for (ψ, x) close to (ψ_0, x_0). We also lack a smooth Hamiltonian $H_t(\psi, x)$, and in general the extremal singled out cannot be included in a Hamiltonian flow whose trajectories satisfy the maximum principle (i.e., are extremals). It turns out, however, that in many cases there is a family of extremals starting out on $T^*_{x_0}M$ that, although not includable in a flow (for example, it contains branchings), has a structure such that the values of the extremals of the family at time t fill some Lagrangian manifold $\mathscr{L}_t$ or at least can be well approximated by a Lagrangian manifold $\mathscr{L}_t$ near (ψ_0, x_0). At the same time, it is Lagrangian manifolds and not a Hamiltonian flow that are needed to get optimality conditions of the type formulated above. In this article we describe the corresponding constructions for systems of the form

$$\dot{x} = f_0(x) + u f_1(x), \qquad |u| \le 1, \ x(0) = x_0, \tag{7}$$

with scalar control, in the case when they are most transparent. Suppose, as above, that $\tilde{u}(\tau)$ and $\tilde{x}(\tau)$ form a solution satisfying the maximum principle. In general the control $\tilde{u}(\tau)$ takes values both on the boundary and in the interior of the interval $[-1, 1]$ (which contains both relay and singular parts). Let $B = \{\tau \ge 0 | |\tilde{u}(\tau)| = 1\}$. For $\tau \in B$ only one-sided variations of $\tilde{u}(\tau)$ are allowed. This inconvenience can be avoided by using "time change" variations.

A "time change" variation is defined to be the following universal construction, which is applicable to an arbitrary controllable system.

Let $\varphi: [0, t] \to [0, t]$ be an absolutely continuous monotonically increasing invertible mapping, let $U(\tau)$, $\tau \in [0, t]$, be an admissible control of the general system

(1), and let $x(\tau)$ be a solution of the equation $dx/d\tau = f(x, u(\varphi^{-1}(\tau)))$. Then

$$\frac{d}{dt}x(\varphi(\tau)) = \dot{\varphi} f(x, u(\tau)), \qquad \varphi(t) = t. \tag{8}$$

We can proceed in reverse order, considering the extended system

$$\begin{cases} \dot{x}^0 = \nu, \\ \dot{x} = \nu f(x, u), \qquad u \in U, \ \nu \geq \varepsilon > 0, \end{cases} \quad \begin{pmatrix} x^0(0) \\ x(0) \end{pmatrix} = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}, \tag{1'}$$

with controlling parameters u and ν.

It follows from (8) that a point $x \in M$ is attainable via trajectories of system (1) in time t if the point $\binom{t}{x}$ is attainable via trajectories of system (1′).

We turn anew to the system (7) with distinguished control $\tilde{u}(\tau)$. The extended system (7′) has two controlling parameters, but they are not independent in general. For our purposes it is expedient to use controls $(u(\tau), \nu(\tau))$ such that $u(\tau) = \tilde{u}(\tau)$ for $\tau \in B$ and $\nu(\tau) = 1$ for $\tau \notin B$ (the time change is switched on only at the points of the set B, and the usual variations apply outside B).

Let $\chi_B(\tau)$ be the characteristic function of B. Setting

$$v(\tau) = \chi_B(\tau)(\nu - 1) + (1 - \chi_B(\tau))(u - \tilde{u}(\tau)), \qquad \tilde{f}_\tau = \begin{pmatrix} 1 \\ f_0 + \tilde{u}(\tau) f_1 \end{pmatrix},$$

$$f_\tau = \chi_B(\tau)\tilde{f}_\tau + (1 - \chi_B(\tau))\begin{pmatrix} 0 \\ f_1 \end{pmatrix}, \qquad X = \begin{pmatrix} x^0 \\ x \end{pmatrix}, \qquad y^0 = \begin{pmatrix} 0 \\ x_0 \end{pmatrix},$$

we get the system

$$\dot{X} = \tilde{f}(X) + v f_\tau(X), \qquad X(0) = y_0. \tag{9}$$

It is essential that at each moment of time the controlling parameter v can take both positive and negative values. At the same time, if $\tilde{u}(\cdot)$ is an optimal control for system (7), then the zero control is optimal for system (9). The perturbation system for system (9) has the form

$$\dot{y} = v g_\tau(y), \qquad y(0) = y_0, \tag{10}$$

where

$$g_\tau = p_{\tau *}^{-1} f_\tau, \qquad \frac{\partial p_\tau}{\partial \tau}(X) = \tilde{f}_\tau(p_\tau(x)), \qquad p_0 = \mathrm{id}.$$

§5. The piecewise constant case

We begin our investigation of system (10) with the case when $g_\tau(y)$ depends in a piecewise constant manner on τ: $g_\tau = g_i$ for $t_i < \tau \leq t_{i+1}$, where $0 = t_0 < t_1 < \cdots < t_l < t_{l+1} = t$. For example, this is the case when $\tilde{u}(\tau)$ is the usual bang-bang control, i.e. $|\tilde{u}(\tau)| \equiv 1$ and $\tilde{u}(\tau)$ is constant on each half-open interval $(t_i, t_{i+1}]$, $i = 0, \ldots, l$. Indeed, in this case $\tilde{f}_\tau$ does not depend on τ for $t_i < \tau \leq t_{i+1}$, and

$$\frac{\partial}{\partial \tau} g_\tau = \frac{\partial}{\partial \tau}(p_{\tau *}^{-1} \tilde{f}_\tau) = p_{\tau *}^{-1}[\tilde{f}_\tau, \tilde{f}_\tau] + p_{\tau *}^{-1}\frac{\partial}{\partial \tau}\tilde{f}_\tau = 0 + 0.$$

Aside from its being of independent interest, the piecewise constant case will be used below for approximating the general case.

A solution $(v(\tau), y(\tau))$ of (10) with $v(\tau)$ sufficiently close to zero satisfies the maximum principle if and only if there exists a nonzero $\psi(\tau) \in T^*_{y(\tau)}(\mathbf{R} \times M)$ such that $(\psi(\tau), y(\tau))$ satisfies the Hamiltonian system corresponding to the Hamiltonian $(\psi, y) \mapsto \psi g_\tau(y) v(\tau)$, and $\psi(\tau) g_\tau(y(\tau)) \equiv 0$.

Curves $(\psi(\tau), y(\tau))$ in $T^*(\mathbf{R}\times M)$ satisfying these conditions are called *extremals* in what follows. We assume that the solution $v(\tau) \equiv 0$, $y(\tau) \equiv y_0$ satisfies the maximum principle; hence $\psi_0 g_\tau(y_0) \equiv 0$ for some ψ_0.

Let $h_i(\psi, y) = \psi g_i(y)$; then $h_i = \psi g_\tau$, $t_i < \tau \leq t_{i+1}$. Let $\mathbf{h}_i$ be the Hamiltonian vector field on $T^*(\mathbf{R}\times M)$ corresponding to the Hamiltonian h_i, and let $e^{\tau \mathbf{h}_i}$: $T^*(\mathbf{R}\times M) \to T^*(\mathbf{R}\times M)$, $\tau \in \mathbf{R}$, be the Hamiltonian flow determined by the field $\mathbf{h}_i$. Denote by $K_i = \{(\psi, y) \in T^*(\mathbf{R}\times M) | h_i(\psi, y) = 0\}$ the zero level set of the Hamiltonian h_i.

Let $\mathscr{L}_\tau$ be the set of values at time τ of all the extremals beginning on the set $\mathscr{L}_0 = T^*_{y_0}(\mathbf{R}\times M) \cap \mathscr{O}$, where $\mathscr{O}$ is some neighborhood of (ψ_0, y_0) in $T^*(\mathbf{R}\times M)$ whose size is not specified beforehand. Then $\mathscr{L}_\tau$ is piecewise constant (constant on each half-open interval $(t_i, t_{i+1}]$). We have that

$$\mathscr{L}_\tau = \{e^{v\mathbf{h}_i}(\psi, y) | (\psi, y) \in \mathscr{L}_{t_i} \cap K_i, \ v \in R\} \quad \text{for } \tau \in (t_i, t_{i+1}], \ i = 0, 1, \dots, l.$$

Furthermore, $(\psi_0, y_0) \in \mathscr{L}_\tau \ \forall \tau$.

For our purposes the sets $\mathscr{L}_\tau$ are too large, since arbitrarily large values of the control $v(\tau)$ are used in their construction. For every $\varepsilon > 0$ we define $\mathscr{L}_\tau(\varepsilon)$ by setting $\mathscr{L}_0(\varepsilon) = \mathscr{L}_0$ and

$$\mathscr{L}_\tau(\varepsilon) = \{e^{v\mathbf{h}(\tau-\tau_i)}(\psi, y) |_{(\psi, y)} \in \mathscr{L}_{t_i}(\varepsilon) \cap K_i, \ |v| < \varepsilon\} \quad \text{for } \tau \in (t_i, t_{i+1}].$$

Proposition 2. *Assume that* $g_0(y_0) \neq 0$ *and*

$$\psi_0[g_i, g_{i+1}](y_0) \stackrel{\text{def}}{=} \{h_i, h_{i+1}\}(\psi_0, y_0) \neq 0, \qquad i = 0, 1, \dots, l-1.$$

Then for all sufficiently small ε *and all* $\tau \in [0, t]$ *the sets* $\mathscr{L}_\tau(\varepsilon)$ *are smooth Lagrangian manifolds.*

This proposition is an immediate consequence of the following assertion.

Lemma. *Suppose that* $\mathscr{L}$ *is a Lagrangian manifold,* h *is a Hamiltonian, and* $(\psi_0, y_0) \in \mathscr{L}$. *Assume that* $h(\psi_0, y_0) = 0$ *and* $\hat{\omega}(\xi, \mathbf{h}(\psi_0, y_0)) \neq 0$ *for some* $\xi \in T_{(\psi_0, y_0)}\mathscr{L}$, *where* $\hat{\omega}$ *is a skew inner product in the symplectic space* $\hat{\Sigma} = T_{(\varphi_0, y_0)}(T^*(\mathbf{R}\times M))$. *Then there exists a neighborhood* W *of the point* (ψ_0, y_0) *in* $\mathscr{L}$ *and an* $\varepsilon > 0$ *such that the set*

$$\{e^{s\mathbf{h}}(\psi, y) | (\psi, y) \in W, \ h(\psi, y) = 0, \ |s| < \varepsilon\} \tag{11}$$

is a Lagrangian manifold.

Proof of the lemma. We have that

$$\langle d_{(\psi_0, y_0)} h, \xi \rangle = \hat{\omega}(\mathbf{h}(\psi_0, y_0), \xi) \neq 0.$$

Consequently, the hyperplane $h^{-1}(0)$ is transversal to the submanifold $\mathscr{L}$ at the point (ψ_0, y_0) (and hence at all nearby points). Hence, $h^{-1}(0) \cap W$ is a smooth $(n-1)$-dimensional submanifold for any sufficiently small neighborhood W of the point (ψ_0, y_0) in $\mathscr{L}$. Further, the vector field $\mathbf{h}$ is tangent to $h^{-1}(0)$ and is transversal to the submanifold $\mathscr{L}$ near (ψ_0, y_0) (otherwise the vector $\mathbf{h}(\psi_0, y_0)$ would be skew orthogonal to $T_{(\psi_0, y_0)}\mathscr{L}$). Consequently, $\mathbf{h}$ is transversal to $h^{-1}(0) \cap W$, so that the set (11) is a smooth n-dimensional submanifold. It remains to see that it is a Lagrangian manifold. Since the flow $e^{s\mathbf{h}}$ preserves the symplectic structure, it suffices to verify the Lagrangian property at the points $(\psi, y) \in h^{-1}(0) \cap W$. We have that

$$\hat{\omega}(\mathbf{h}(\psi_0, y_0), \eta) = \langle d_{(\psi_0, y_0)} h, \eta \rangle = 0$$
$$\forall \eta \in T_{(\psi, y)}\mathscr{L} \cap T_{(\psi, y)} h^{-1}(0) = T_{(\psi, y)}(W \cap h^{-1}(0)). \quad \blacksquare$$

Proposition 2 is easy to derive from the lemma by using induction on i and the relations

$$\mathbf{h}_i(\psi_0, y_0) \in \mathscr{L}_\tau(\varepsilon) \quad \text{for } t_i < \tau \leq t_{i+1},$$
$$\omega(\mathbf{h}_{i+1}(\psi_0, y_0), \ \mathbf{h}_i(\psi_0, y_0)) = \{h_i, h_{i+1}\}(\psi_0, y_0) \neq 0.$$

It should be noted that Proposition 1 gives a sufficient but far from necessary condition for the $\mathscr{L}_\tau(\varepsilon)$ to be Lagrangian manifolds. Broader (but less explicit) conditions are easy to get by analyzing the proof of the proposition.

The tangent space $\Lambda_\tau = T_{(\psi_0, y_0)}\mathscr{L}_\tau(\varepsilon)$ is defined by the following recursion relations (we make the identification $T^*_{y_0}(\mathbf{R} \times M) = T_{(\varphi_0, y_0)}\,(T^*_{y_0}(\mathbf{R} \times M)) \subset \widehat{\Sigma}$)

$$\text{(A)} \qquad \Lambda(\tau) = (\Lambda_{t_i} + \mathbf{R}\mathbf{h}_i(\psi_0, y_0)) \cap \{\mathbf{h}_i(\psi_0, y_0)\}^\perp, \qquad t_i < \tau < t_{i+1}, \quad i = 0, 1, \dots, l.$$

The relations (A) determine a piecewise constant curve $\Lambda_{(\tau)}$ in the Lagrangian Grassmannian $L(\widehat{\Sigma})$. To find this curve it is by no means necessary to know the manifold $\mathscr{L}_\tau(\varepsilon)$ itself.

What is more, the relations (A) always determine a piecewise constant family of Lagrangian subspaces, without any preliminary conditions on h_i and independently of whether or not the $\mathscr{L}_\tau(\varepsilon)$ are smooth manifolds. In degenerate situations the space $\Lambda_{(\tau)}$ can be regarded as the generalized tangent space to $\mathscr{L}_\tau(\varepsilon)$ at the point (ψ_0, y_0).

It is possible to define the index of an arbitrary piecewise constant curve in the Lagrangian Grassmannian with respect to a fixed Lagrangian subspace Π. This index is defined in a purely algebraic manner, but turns out to be closely connected with the Maslov index of the continuous curve considered in §3.

Let $\Pi, \Delta_0, \Delta_1 \in L(\widehat{\Sigma})$. We define a real quadratic form q on the space $(\Delta_0 + \Delta_1) \cap \Pi$ by setting $q(\delta) = \hat{\omega}(\delta_0, \delta_1)$ if $\delta_0 + \delta_1 = \delta \in (\Delta_0 + \Delta_1) \cap \Pi$, $\delta_0 \in \Delta_0$, $\delta_1 \in \Delta_1$. It is easy to see that $q(\delta)$ is well defined even when δ does not have a unique representation as a sum of elements in Δ_0 and Δ_1. As usual, $\operatorname{ind} q$ denotes the maximal dimension of a subspace of $(\Delta_0 + \Delta_1) \cap \Pi$ on which the form q is negative definite.

Let

$$\operatorname{ind}_\Pi(\Delta_0, \Delta_1) = \operatorname{ind} q + \tfrac{1}{2}(\dim(\Pi \cap \Delta_0) + \dim(\Pi \cap \Delta_1)) - \dim(\Pi \cap \Delta \cap \Delta_1),$$

a symplectic invariant of this triple of Lagrangian subspaces that takes nonnegative half-integer values. It can be shown (see [1]) that the triangle inequality holds:

$$(12) \qquad \operatorname{ind}_\Pi(\Delta_0, \Delta_2) \leq \operatorname{ind}_\Pi(\Delta_0, \Delta_1) + \operatorname{ind}_\Pi(\Delta_1, \Delta_2) \quad \forall \Pi, \Delta_0, \Delta_1, \Delta_2 \in L(\widehat{\Sigma}).$$

Let $\Delta(\tau) = \Delta_i \in L(\widetilde{\Sigma})$ for $t_{i-1} < \tau < t_i$, $i = 0, 1, \dots, l+1$.

The index of the curve $\Delta(\cdot)$ with respect to Π is defined to be the quantity

$$\operatorname{ind}_\Pi \Delta(\cdot) = \sum_{i=0}^{l} \operatorname{ind}_\Pi(\Delta_i, \Delta_{i+1}).$$

The quantity

$$\overline{\operatorname{ind}}\, \Delta(\cdot) \stackrel{\text{def}}{=} \operatorname{ind}_\Pi \Delta(\cdot) + \operatorname{ind}_\Pi(\Delta_{l+1}, \Delta_0)$$

is called the *upper index* of the curve $\Delta(\cdot)$. It can be shown that the upper index does not depend on Π, and that

$$\overline{\operatorname{ind}}\,\Delta(\cdot) = \max_{\Pi\in L(\widehat{\Sigma})} \operatorname{ind}_\Pi \Delta(\cdot).$$

Using this index, we can formulate a necessary condition for optimality that is analogous to Theorem 1. Let

$$\Phi_t = \{\psi \in T^*_{y_0}(\mathbf{R}\times M) = (\mathbf{R}\oplus T_{x_0}M)^* | \psi \neq 0,\ \psi(1\oplus 0) \leq 0, \\ h_i(\psi, y_0) = 0,\ i = 0, 1, \dots, t\},$$

a half-space in $T^*_{y_0}(\mathbf{R}\times M)$, with

$$\dim \Phi_t = n + 1 - \operatorname{rank}\left(\sum_{i=0}^{l} g_i(y_0) g_i(y_0)^*\right).$$

Let $\widehat{\Phi}_t = \operatorname{span}\Phi_t$, an isotropic subspace of $\widehat{\Sigma}$. For any $\psi \in \Phi_t$ we define the piecewise constant curve $\Lambda_\psi(\tau)$, $0 \leq \tau \leq t$, in the Lagrangian Grassmannian $L(\widehat{\Phi}_t^{\perp}/\widehat{\Phi}_t)$ by setting

$$\Lambda_\psi(0) = T^*_{y_0}(\mathbf{R}\times M)/\widehat{\Phi}_t, \qquad \Lambda_\psi(\tau) = (\Lambda_\psi(t_i) + \mathbf{R}\mathbf{h}_i(\psi, y_0)) \cap (\mathbf{h}_i(\psi, y_0))^{\perp}, \\ t_i < \tau \leq t_{i+1},\ i = 0, 1, \dots, l.$$

In particular, $\Lambda_{\psi_0}(\tau) = \Lambda(\tau)/\Phi_t$.

Theorem 2. *If the control* $\tilde{u}(\tau)$, $0 \leq \tau \leq t$, *is optimal for system* (7), *then*

$$\min_{\psi\in\Phi_t} \overline{\operatorname{ind}}\,\Lambda_\psi(\cdot) \leq n.$$

Remark. In Theorem 2, in contrast to Theorem 1, the operation of taking the limit superior with respect to Π is not present. It is possible to simplify the formulation because in the situation under consideration the operations $\overline{\lim}$ and min can be transposed.

§6. Curves of bounded monotone variation on the Lagrangian Grassmannian

In order to proceed further it is necessary to go somewhat deeper into the geometry of the Lagrangian Grassmannian and, in particular, to clarify the connection between the indices considered in §§3 and 5.

Let Σ be a symplectic space with skew inner product ω, and let $\dim\Sigma = 2n$. A curve $\Delta(\tau)$, $\alpha \leq \tau \leq \beta$, on $L(\Sigma)$ is said to be *simple* if there exists a $\Lambda \in L(\Sigma)$ such that $\Delta(\tau)\cap\Lambda = 0\ \forall\tau\in[0, t]$.

Proposition 3. 1) *For any* $\Delta_0, \Delta_1 \in L(\Sigma)$ *there exists a simple smooth nondecreasing curve* $\Delta(\tau)$, $0 \leq \tau \leq 1$, *such that* $\Delta(0) = \Delta_0$ *and* $\Delta(1) = \Delta_1$.

2) *For any simple nondecreasing curve* $\Delta(\tau)$, $\alpha \leq \tau \leq \beta$, *and for all* $\Pi \in L(\Sigma)$ *with* $\Pi\cap\Delta(\alpha) = \Pi\cap\Delta(\beta) = 0$,

$$\operatorname{Ind}_\Pi\Delta(\cdot) = \operatorname{ind}_\Pi(\Delta(\alpha), \Delta(\beta)).$$

Corollary. *If* $\Delta(\tau)$, $\alpha \leq \tau \leq \beta$, *is a simple nondecreasing curve, then*

$$\operatorname{ind}_\Pi(\Delta(\alpha), \Delta(\beta)) = \operatorname{ind}_\Pi(\Delta(\alpha), \Delta(\tau)) \\ + \operatorname{ind}_\Pi(\Delta(\tau), \Delta(\beta)) \quad \forall\tau\in[\alpha, \beta],\ \Pi\in L(\Sigma).$$

See [1] for a proof.

Proposition 3 connects the indices of piecewise constant curves with the Maslov indices of continuous curves. The index of a piecewise constant curve with respect to Π introduced in §5 is equal to the Maslov index of the continuous curve (also with respect to Π) obtained if the points of the piecewise constant curve are joined successively by simple nondecreasing curves. The upper index of a piecewise constant curve turns out to be equal to the Maslov index of the correspondingly constructed closed nondecreasing curve.

Conversely, if $\Delta(\tau)$, $0 \le \tau \le t$, is a continuous nondecreasing curve and $0 = t_0 < t_1 < \cdots < t_l < t_{l+1} = t$ are such that the pieces $\Delta(\cdot)|_{[t_i, t_{i+1}]}$ of the curve are simple, $i = 0, \ldots, l$, then

$$\operatorname{ind}_\Pi \Delta(\cdot) = \sum_{i=0}^{l} \operatorname{ind}_\Pi(\Delta(\tau_i), \Delta(\tau_{i+1}))$$

for all Π such that $\Pi \cap \Lambda(0) = \Pi \cap \Lambda(t) = 0$.

We now turn to arbitrary discontinuous curves. Let $D = \{\tau_1, \ldots, \tau_k\}$ be some finite subset of $[0, t]$ with $\tau_1 < \cdots < \tau_k$, and let $\Delta(\tau)$, $\tau \in [0, t]$, be a curve on $L(\Sigma)$. The symbol $\Delta^D(\tau)$ denotes the piecewise constant curve defined as follows: $\Delta^D(0) = \Delta(0)$, and $\Delta^D(\tau) = \Delta(\tau_i)$ for $\tau_{i-1} < \tau < \tau_i$, $i = 1, \ldots, k+1$, where $\tau_0 = 0$ and $\tau_{k+1} = t$. Then for all $\Pi \in L(\Sigma)$

$$\operatorname{ind}_\Pi \Delta^D(\cdot) = \sum_{i=0}^{k} \operatorname{ind}_\Pi(\Delta_{\tau_i}, \Delta_{\tau_{i+1}}),$$

$$\overline{\operatorname{ind}}\, \Delta^D(\cdot) = \operatorname{ind}_\Pi \Delta^D(\cdot) + \operatorname{ind}_\Pi(\Delta_t, \Delta_0) \le \operatorname{ind}_\Pi \Delta^D(\cdot) + n.$$

Moreover,

$$\overline{\operatorname{ind}}^D(\cdot) = \max_{\Pi \in L(\Sigma)} \operatorname{ind}_\Pi \Delta^D(\cdot).$$

Denote by $\mathscr{D}$ the collection of all finite subsets of $[0, t]$. For $D_1, D_2 \in \mathscr{D}$ with $D_1 \subset D_2$ the triangle inequality (12) gives us that

$$\operatorname{ind}_\Pi \Delta^{D_1}(\cdot) \le \operatorname{ind}_\Pi \Delta^{D_2}(\cdot) \quad \forall \Pi \in L(\Sigma).$$

The *index of the curve* $\Delta(\tau)$, $0 \le \tau \le t$, with respect to $\Pi \in L(\Sigma)$ is defined to be the quantity

$$\operatorname{ind}_\Pi \Delta(\cdot) = \sup_{D \in \mathscr{D}} \operatorname{ind}_\Pi \Delta^D(\cdot),$$

which is equal to a nonnegative half-integer or to $+\infty$, and the *upper index of the curve* $\Delta(\cdot)$ is defined to be the quantity

$$\overline{\operatorname{ind}}\, \Delta(\cdot) = \sup_{D \in \mathscr{D}} \overline{\operatorname{ind}}\, \Delta^D(\cdot) = \sup_{D \in \mathscr{D}} \sup_{\Pi \in L(\Sigma)} \operatorname{ind}_\Pi \Delta^D(\cdot).$$

which is equal to a positive half-integer or to $+\infty$.

We have

$$\overline{\operatorname{ind}}\, \Delta(\cdot) = \operatorname{ind}_\Pi \Delta(\cdot) + \operatorname{ind}_\Pi(\Delta(t), \Delta(0)) \le \operatorname{ind}_\Pi \Delta(\cdot) + n.$$

In particular, $\overline{\operatorname{ind}}\, \Lambda(\cdot)$ is finite if and only if every $\operatorname{ind}_\Pi \Delta(\cdot)$ is finite.

A curve $\Delta(\tau)$, $0 \le \tau \le t$, in $L(\Sigma)$ is called a *curve of bounded monotone variation* if $\overline{\operatorname{ind}}\, \Delta(\cdot) < +\infty$.

Curves of bounded monotone variation have properties similar to the properties of bounded monotonically nondecreasing real functions.

Theorem 3. *Suppose that $\Delta(\tau)$, $0 \le \tau \le t$, is a curve of bounded monotone variation in $L(\Sigma)$. Then:*

1) *The limits $\Delta(\tau \pm 0) = \lim_{\theta \to \tau \pm 0} \Delta(\theta)$ exist for any $\tau \in [0, t]$.*

2) *The curve $\Delta(\cdot)$ has at most countably many points of discontinuity.*

3) *$\Delta(\tau)$ is differentiable almost everywhere on $[0, t]$ and at any points of differentiability the derivative is nonnegative, $d\Delta/d\tau \ge 0$.*

The proofs of this and the next theorem on compactness are based on the connection between the indices of a smooth nondecreasing curve and its length. We only briefly explain this connection; it is described in greater detail in [1]. Assume that an inner product $(\cdot|\cdot)$ is given on Σ that is compatible with the symplectic structure ω; compatibility means that $\omega(\lambda_1, \lambda_2) = (A\lambda_1|\lambda_2)$, where A: $\Sigma \to \Sigma$ is an orthogonal operator with respect to the given inner product.

Let $\Delta \in L(\Sigma)$. Recall that the space $T_\Delta L(\Sigma)$ is identified with the space of quadratic forms on Δ. To each such form $q \in T_\Delta L(\Sigma)$ there corresponds a symmetric linear operator $s(q)$, where $q(\lambda) = (s(q)\lambda|\lambda)$ $\forall \lambda \in \Delta$. We specify a Riemannian metric on $L(\Sigma)$ by setting the inner product of a pair of tangent vectors $q_1, q_2 \in T_\Delta L(\Sigma)$ equal to $\mathrm{tr}(s(q_1)s(q_2))$. The length of an absolutely continuous curve $\Delta(\tau)$, $\alpha \le \tau \le \beta$, is defined by

$$\mathrm{length}(\Delta(\cdot)) = \int_\alpha^\beta \sqrt{\mathrm{tr}(s(d\Delta/d\tau)^2)}\, d\tau.$$

Further, if the curve $\Delta(\tau)$ is closed, $\Delta(\alpha) = \Delta(\beta)$, then its Maslov index $\mathrm{Ind}_\Pi \Delta(\cdot)$ does not depend on Π and has the integral representation

$$\mathrm{Ind}_\Pi \Delta(\cdot) = \frac{2}{\pi} \int_\alpha^\beta \mathrm{tr}\left(s\left(\frac{d\Delta}{d\tau}\right)\right) d\tau.$$

If in addition the curve $\Delta(\tau)$ is nondecreasing, then

(a) $\mathrm{ind}_\Pi \Delta(\cdot) = \overline{\mathrm{ind}}\, \Delta(\cdot)$ and

(b) the eigenvalues of the symmetric operators $s(d\Delta/d\tau)$ are nonnegative, and

$$\sqrt{\mathrm{tr}\left(s\left(\frac{d\Delta}{d\tau}\right)^2\right)} \le \mathrm{tr}\, s\left(\frac{d\Delta}{d\tau}\right) \le \sqrt{n\,\mathrm{tr}\left(s\left(\frac{d\Delta}{d\tau}\right)^2\right)}.$$

Therefore, a closed nondecreasing curve satisfies the inequality

$$\mathrm{length}(\Delta(\cdot)) \le \overline{\mathrm{ind}}\, \Delta(\cdot) \le \sqrt{n}\, \mathrm{length}(\Delta(\cdot)).$$

According to Proposition 3, any nondecreasing curve can be completed to a closed curve by adding some simple nondecreasing curve. This does not increase the upper index; hence the inequality

$$\mathrm{length}(\Delta(\cdot)) \le \overline{\mathrm{ind}}\, \Delta(\cdot)$$

is valid for any nondecreasing curves, not necessarily closed.

The following theorem is an analogue of Helly's principle for real functions of bounded variation.

Theorem 4 (Compactness Principle). *Let $c \ge 0$. The set of all curves $\Delta(\tau)$ of bounded monotone variation in $L(\Sigma)$ defined on the interval $0 \le \tau \le t$ and satisfying the condition $\overline{\mathrm{ind}}\, \Delta(\cdot) \le c$ is compact in the topology of pointwise convergence.*

Let $\mathscr{A}$ be a directed set (i.e., a partially ordered set such that for any pair of elements there is an element majorizing both of them) and let $\Delta_\alpha(\cdot)$: $[0, t] \to L(\Sigma)$,

$\alpha \in \mathscr{A}$, be a generalized sequence of curves on the Lagrangian Grassmannian. If the upper indices of the curves are uniformly bounded, then the sequence contains a pointwise convergent subsequence. By considering not only the upper indices but also the indices with respect to different subspaces $\Pi \in L(\Sigma)$ it is possible to get a useful criterion for convergence of the whole sequence $\Delta_\alpha(\cdot)$, $\alpha \in \mathscr{A}$.

Proposition 4. *Suppose that $\Delta(\tau)$, $0 \le \tau \le t$, is a nonconstant curve of bounded monotone variation on $L(\Sigma)$, and let $\Delta_\alpha(\cdot)$: $[0, t] \to L(\Sigma)$ be a generalized sequence of curves. Assume that for any $\tau \in [0, t]$ there exists a dense subset $\mathscr{P}_\tau$ of $L(\Sigma)$ such that*

$$\mathscr{A}\text{-}\lim \operatorname{ind}_\Pi(\Delta_\alpha(\cdot)|_{[0,t]}) = \operatorname{ind}_\Pi(\Delta(\cdot)|_{[0,t]}) \quad \forall \Pi \in \mathscr{P}_\tau, \ \tau \in [0, t].$$

Then

$$\mathscr{A}\text{-}\lim \Delta_\alpha(\tau) = \Delta(\tau) \quad \forall \tau \subset [0, t].$$

Remark. The above sufficient condition for pointwise convergence is not necessary in general. It becomes necessary if instead of curves on $L(\Sigma)$ we consider liftings of them to the universal covering of this manifold—the noncompact manifold $\widehat{L}(\Sigma)$. Continuous curves can be lifted to the universal covering "by definition". Curves of bounded monotone variation can also be lifted. Indeed, if we glue together the discontinuities of a monotonically bounded curve by means of simple nondecreasing curves, then we get a continuous rectifiable curve that can be lifted, and then we can remove the liftings of the attached parts. The result does not depend on how the gluing is done. The definition of the index of a curve with respect to a Lagrangian subspace extends to curves on $\widehat{L}(\Sigma)$. There is a compactness principle that is a strengthening of Theorem 4. In general, the whole theory (including the optimality condition formulated in the next section) acquires a more complete form if we move from $L(\Sigma)$ to $\widehat{L}(\Sigma)$. However, this material goes beyond the framework of the present article and will be given in a more detailed publication.

§7. THE MAIN THEOREM

We now return to the investigation of system (10) (in order to get conditions for the optimality of the control $\tilde{u}(\cdot)$ for system (7)), and we consider the case of an arbitrary dependence of $g_\tau(y)$ on τ, rather than just a piecewise constant dependence as in §5.

Let $h_\tau(\psi, y) = \psi g_\tau(y)$, where $y \in \mathbf{R} \times M$ and $\psi \in T_y^*(\mathbf{R} \times M)$, and let $\mathbf{h}_\tau(\psi, y)$ be the Hamiltonian vector field on $T^*(\mathbf{R} \times M)$ corresponding to the Hamiltonian h_τ.

It is assumed that the control $\tilde{u}(\tau)$, $0 \le \tau \le t$, satisfies the maximum principle. This is equivalent to the nonemptiness of the set

$$\Phi_t = \{\psi \in T_{y_0}^*(\mathbf{R} \times M) = (\mathbf{R} \oplus T_{x_0}M)^*|_{\psi \ne 0}, \ \psi(1 \oplus 0) \le 0, \\ h_\tau(\psi, y_0) = 0, \ \text{for a.e. } \tau \in [0, t]\}.$$

Let $\widehat{\Phi}_t = \operatorname{span} \Phi_t$, a subspace of codimension

$$\operatorname{rank} \int_0^t g_\tau(y_0) g_\tau(y_0)^* \, d\tau \quad \text{in } T_{y_0}^*(\mathbf{R} \times M).$$

Choose some $\psi \in \Phi_t$. The space $T_{y_0}^*(\mathbf{R} \times M)$ is identified in the usual way with the "vertical" Lagrangian subspace in the symplectic space $T_{(\varphi_0, y_0)}(T^*(\mathbf{R} \times M))$. Thus, we can regard $\widehat{\Phi}$ as an isotropic subspace of $T_{(\psi_0, y_0)}(T^*(\mathbf{R} \times M))$ and consider the symplectic space $\widehat{\Phi}_t^\perp / \widehat{\Phi}_t$. Let $D = \{\tau_1, \ldots, \tau_k\}$ be a finite subset of $[0, t]$ with

$\tau_1 < \cdots < \tau_k$. We define a piecewise constant curve $\Lambda_\psi^D(\tau)$, $0 \le \tau \le t$, in the Lagrangian Grassmannian $L(\widehat{\Phi}_t^\perp/\widehat{\Phi}_t)$ by setting

$$\Lambda_\psi^D(0) = T_{y_0}^*(\mathbf{R} \times M)/\widehat{\Phi}_t,$$
$$\Lambda_\psi^D(\tau) = (\Lambda_\psi^D(\tau_i) + \mathbf{R}\mathbf{h}_{\tau_i}(\psi, y_0)) \cap (\mathbf{h}_{\tau_i}(\psi, y_0))^\perp$$

for $\tau_i < \tau \le \tau_{i+1}$, $i = 0, 1, \dots, k$, where $\tau_0 = 0$ and $\tau_{k+1} = t$.

Proposition 5. *Suppose that D_1 and D_2 are finite subsets of $[0, t]$ with $D_1 \subset D_2$. Then*

$$\operatorname{ind}_\Pi \Lambda_\psi^{D_1}(\cdot) \le \operatorname{ind}_\Pi \Lambda_\psi^{D_2}(\cdot) \quad \forall \Pi \in L(\widehat{\Phi}_t^\perp/\widehat{\Phi}_t),$$

so that $\Lambda^{D_1}(t) \cap \Pi = 0$.

The scheme of the proof is as follows: with each finite set $D = \{\tau_1, \dots, \tau_k\} \in [0, t]$ we associate the linear space

$$V_\Pi^D = \left\{(\lambda, v_1, \dots, v_k) | \lambda \in \Lambda_\psi^D(0),\ v_i \in R,\ i = 1, \dots, k,\ \lambda + \sum_{i=1}^k v_i \mathbf{h}_{\tau_i}(\psi, y_0) \in \Pi\right\}$$

and the quadratic form $Q_\Pi^D \colon V_\Pi^D \to \mathbf{R}$ defined by

$$Q_\Pi^D(\lambda, v_1, \dots, v_k) = \sum_{i=0}^k \omega\left(\lambda + \sum_{j=0}^{i-1} v_j \mathbf{h}_{\tau_j}(\psi, y_0),\ v_i \mathbf{h}_{\tau_i}(\psi, y_0)\right), \tag{13}$$

where $\omega(\cdot, \cdot)$ is the symplectic form on $\widehat{\Phi}_t^\perp/\widehat{\Phi}_t$.

It can be shown that

$$\operatorname{ind}_\Pi \Lambda_\psi^D(\cdot) = \operatorname{ind} Q_\Pi^D + \frac{1}{2}\left(\dim(\Lambda_\psi^D(0) \cap \Pi) + \dim(\Lambda_\psi^D(t) \cap \Pi) - \dim\left(\bigcap_{i=0}^{k+1} \Lambda_\psi^D(\tau_i) \cap \Pi\right)\right). \tag{14}$$

If $D_1 \subset D_2$, then $V_\Pi^{D_1} \subset V_\Pi^{D_2}$ and $Q_\Pi^{D_1} = Q_\Pi^{D_2}|_{V_\Pi^{D_1}}$, which implies the required inequality. ■

We remark that

$$\dim \bigcap_{i=0}^{k+1} \Lambda_\psi^D(\tau_i) = \operatorname{rank} \int_0^t g_\tau(g_0) g_\tau(y_0)^* \, d\tau - \operatorname{rank}\left(\sum_{i=0}^k g_{\tau_i}(y_0) g_{\tau_i}(y_0)^*\right).$$

This quantity can always be made zero by adding some points to D.

Let T be the collection of all points in $[0, t]$ that are points of density of the measurable vector-valued functions $\tau \mapsto \mathbf{h}_\tau(\psi, y_0)$ for any $\psi \in \widehat{\Phi}_t$. The set $[0, t] \backslash \mathrm{T}$ has zero measure (recall that $\mathbf{h}_\tau$ depends affinely on ψ). Denote by $\mathscr{D}$ the directed set whose elements are the finite subsets of T, partially ordered by inclusion.

Theorem 5. *If the control $\tilde{u}(\tau)$, $0 \le \tau \le t$, is optimal for the system* (7), *then*

1) $\min_{\psi \in \Phi_t} \sup_{D \in \mathscr{D}} \overline{\operatorname{ind}}\, \Lambda_\psi^D(\cdot) \le n$; *and*

2) *for any* $\psi \in \Phi_t$ *such that the quantity* $I(\psi) = \sup_{D\in\mathscr{D}} \overline{\operatorname{ind}}\, \Lambda_\psi^D(\cdot)$ *is finite and any* $\tau \in [0, t]$ *the limit* $\mathscr{D}\text{-}\lim \Lambda_\psi^D(\tau) = \Lambda_\psi(\tau)$ *exists, the curve* $\tau \mapsto \Lambda_\psi(\tau)$ *in* $L(\widehat{\Phi}^\perp/\widehat{\Phi})$ *has bounded monotone variation, and*

$$\overline{\operatorname{ind}}\, \Lambda_\psi(\cdot) \le I(\psi).$$

The proof of assertion 1) in Theorem 5, like the proofs of Theorems 1 and 2, is based on results in [2] and [3] and on a study of the second variation of the control system (10) along the zero control. Formula (14) connects $\operatorname{ind}_\Pi \Lambda_\psi^D(\cdot)$ with the index of the quadratic form (13), and the form (13) with $\Pi = \Lambda_\psi^D(0)$ coincides in essence with the second variation of system (10), regarded on the special finite-dimensional space of "spike" variations of the zero control.

Assertion 2) follows from Propositions 4 and 5 and Theorem 4. The Lagrangian space

$$(\Lambda_\psi(\tau) + \widehat{\Phi}_t) \subset T_{(\psi, y_0)}(T^*(\mathbf{R} \times M))$$

can be interpreted geometrically as the tangent space at the point (ψ, y_0) to the set of values at time τ of the extremals of system (10) that begin on $T^*_{y_0}(\mathbf{R} \times M)$.

§8. Pointwise conditions for optimality

Theorem 5 yields the following fact.

Corollary. *If the control* $\tilde{u}(\tau)$, $0 \le \tau \le t$, *is optimal for system* (7), *then for some* $\psi \in \Phi_t$ *there is a curve of bounded monotone variation* $\Lambda(\tau)$, $\tau \in [0, t]$, *in the Lagrangian Grassmannian* $L(T_{(\psi, y_0)}(T^*(\mathbf{R} \times M)))$ *such that*

$$\Lambda(0) = T^*_{y_0}(\mathbf{R} \times M) \quad \text{and} \quad \mathbf{h}_\tau(\psi, y_0) \in \Lambda(\tau + 0) \quad \forall \tau \in \mathrm{T}. \tag{15}$$

It turns out that simply from the existence of such a curve $\Lambda(\tau)$ (without constructing it explicitly) we can derive all the generalized Legendre conditions for optimality of singular controls, and get much additional information.

Recall that $\hat{\omega}$ denotes the closed differential 2-form determining the standard symplectic structure on $\mathbf{R} \times M$.

Theorem 6. *Assume that* $\Lambda(\tau)$, $\tau \in [0, t]$, *is a curve of bounded monotone variation in the Lagrangian Grassmannian and satisfies condition* (15). *Then for any* $\tau \in [0, t]$ *there exists an* $\varepsilon > 0$ *such that either* $h_\theta \in \Lambda_{\tau+0}$ $\forall \theta \in (\tau, \tau+\varepsilon) \cap \mathrm{T}$, *or there exists a* 1*-form* Ω_τ *defined in some neighborhood of the point* (ψ, y_0) *in* $T^*(\mathbf{R} \times M)$ *such that* $d\Omega_\tau = \hat{\omega}$, *the zeros of the form* Ω_τ *constitute a Lagrangian manifold with* $\Lambda_{\tau+0}$ *as its tangent space at* (ψ, y_0), *and the following conditions are satisfied*:

(a) $(i_{\mathbf{h}_\theta} d i_{\mathbf{h}_\theta} \Omega_\tau)(\psi_0, y_0) > 0$ $\forall \theta \in (\tau, \tau+\varepsilon) \cap \mathrm{T}$;

(b) $\hat{\omega}(\lambda, \mathbf{h}_\theta(\psi, y_0))^2 / (i_{\mathbf{h}_\theta} d i_{\mathbf{h}_\theta} \Omega_\tau)(\psi, y_0) \to 0$ *as* $\theta \downarrow \tau$ $(\theta \in \mathrm{T})$ $\forall \lambda \in \Lambda_{\tau+0}$.

The proof of Theorem 6 is based on the use on the Lagrangian Grassmannian of natural local coordinates (see [5]) in which each Lagrangian subspace close to $\Lambda_{\tau+0}$ can be represented by a quadratic form on $\Lambda_{\tau+0}$. For example, the subspace $\Lambda_{\tau+0}$ itself can be represented by the zero form, and the quadratic form in (b) represents the subspace $(\Lambda_{\tau+0} + \mathbf{R}\mathbf{h}_\theta(\psi_0, y_0)) \cap \Lambda_{\tau+0}^\perp$. ■

The formulation of conditions (a) and (b) can be simplified at points where h_τ depends smoothly on τ: if the first nonzero derivative of the quantity

$$\alpha_\tau(\theta) = i_{\mathbf{h}_\theta}(d i_{\mathbf{h}_\theta} \Omega)(\psi_0, y_0)$$

with respect to θ has order m at $\theta = \tau$, then condition (b) is equivalent to the inclusion

$$\mathbf{h}_\tau^{(k)} \in \Lambda_{\tau+0}, \qquad 0 \le k \le \frac{m}{2}, \quad \text{where } \mathbf{h}_\tau^{(k)} = \frac{d^k}{d\theta^k} \mathbf{h}_\theta \big|_{\theta=\tau}, \tag{16}$$

and condition (a) is equivalent to the inequality $\alpha_\tau^{(m)} \stackrel{\text{def}}{=} (d^m/d\theta^m)\alpha_\tau(\theta)|_{\theta=\tau} > 0$. At the same time, if (16) holds, then $\alpha_\tau^{(m)}$ does not depend on the choice of Ω_τ and can be computed by the formula

$$\alpha_\tau^{(m)} = \sum_{k=0}^{[n/2]} \binom{m}{k} \hat{\omega}(\mathbf{h}_\tau^{(m-k)}, \mathbf{h}_\tau^{(k)}).$$

Recall that

$$\hat{\omega}(\mathbf{h}_\tau^{(i)}, \mathbf{h}_\tau^{(j)}) = \{h_\tau^{(j)}, h_\tau^{(i)}\}(\psi, y_0) = \psi[g_\tau^{(j)}, g_\tau^{(i)}](y_0),$$

where $g_\tau^{(k)} = (d^k/d\tau^k)g_\theta|_{\theta=\tau}$.

Corollary. *Let $\tilde{u}(\tau)$, $0 \le \tau \le t$, be an optimal control for system (7). Then there exists a $\psi \in \Phi_t$ such that the following condition holds for any integer $m > 0$ and any point $\tau \in [0, t)$ for which the vector field $g_\tau^{(m)}$ is defined.*

If

$$\sum_{j=0}^{[k/2]} \binom{k}{j} \psi[g_\tau^{(j)}, g_\tau^{(k-j)}]y_0 = 0, \qquad 0 \le k \le m, \tag{17}$$

then $\psi[g_\tau^{(i)}, g_\tau^{(j)}](y_0) = 0$ for $0 \le i, j \le m/2$, and

$$\sum_{j=0}^{[m/2]} \binom{m}{j} \psi[g_\tau^{(j)}, g_\tau^{(m-j)}](y_0) \ge 0.$$

Remark. It can be shown that if the equalities (17) hold identically for all τ' close to τ, then the corollary reduces to the generalized Legendre conditions (see [2]); however, in contrast to the generalized Legendre conditions, it is still of substance even when the equalities do not hold identically.

§9. A CONDITION FOR LOCAL CONTROLLABILITY

To every necessary condition for optimality there usually corresponds a sufficient condition for local controllability. Here we present a condition for local controllability corresponding to the optimality conditions in §8.

Consider the control system

$$\dot{x} = f_0(x) + u f_1(x), \qquad x \in M, \; x(0) = x_0, \; u \in \mathbf{R}, \tag{18}$$

where f_0 and f_1 are vector fields of class C^∞ on M.

Let x_t, $t \ge 0$, be the trajectory corresponding to the zero control.

Let $\varepsilon > 0$ be real and $N > 0$ an integer. Denote by $\mathfrak{A}_t(\varepsilon, N)$ the subset of M consisting of the points that can be reached in precisely the time t by moving from the point x_0 along trajectories of system (18) and using only piecewise constant controls $u(\tau)$, $0 \le \tau \le t$, having at most N points of discontinuity and satisfying the condition $|u(\tau)| \le \varepsilon$, $0 \le \tau \le t$.

We say that system (18) is *strongly locally controllable* along the trajectory x_t if for any $t > 0$ there is an N_t such that

$$x_t \in \operatorname{int} \mathfrak{A}_t(\varepsilon, N_t) \quad \forall \varepsilon > 0.$$

Let

$$S_0 = \operatorname{span}\{ad^i f_0 f_1(x_0), \; 0 \le i < +\infty\},$$
$$S_k = S_0 + \operatorname{span}\{[ad^i f_0 f_1, ad^j f_0 f_1](x_0), \; 0 \le i, j \le k\}, \qquad k = 1, 2, \dots,$$

which is a nondecreasing sequence of subspaces of $T_{x_0}M$, and let

$$\xi_k = \sum_{j=0}^{[k/2]} \binom{k}{j} [ad^j f_0 f_1, ad^{(k-j)} f_0 f_1](x_0), \qquad k = 1, 2, \dots,$$

which is a sequence of vectors in $T_{x_0}M$.

Proposition 6. *If $\xi_k \in S_{[k/2]}$ for any $k > 0$ and $S_m = T_{x_0}M$ for some m, then system* (18) *is strongly locally controllable along the trajectory x_t.*

Remark. The condition $\xi_1 \in S_0$ is the well-known necessary condition $[f_1, [f_0, f_1]] \in S_0$ for local controllability; if it holds, then $S_1 = S_0$, but S_2 now can differ from S_0. For example, the system in $\mathbf{R}^3$ given by

$$\begin{cases} \dot{x}_1 = 1 + u(x_1 x_2 + x_3), \\ \dot{x}_2 = -u x_1^2, \\ \dot{x}_3 = u(x_1^3 + c), \end{cases} \qquad x_i(0) = 0, \ i = 1, 2, 3,$$

where c is an arbitrary constant, satisfies the conditions of Proposition 6. For this system

$$S_1 = S_0 = \{(0, \alpha, \beta) | \alpha, \beta \in \mathbf{R}\}, \quad S_2 = \mathbf{R}^3, \quad \xi_1 = \xi_2 = \xi_3 = 0.$$

Bibliography

1. A. A. Agrachëv, *Quadratic mappings in geometric control theory*, Itogi Nauki i Tekhniki: Problemy Geometrii, vol. 20, VINITI, Moscow, 1988, pp. 111–205; English transl. in J. Soviet Math. **51** (1990), no. 6.
2. A. A. Agrachëv and R. V. Gamkrelidze, *The index of extremality and quasiextremal controls*, Dokl. Akad. Nauk SSSR **284** (1985), 777–781; English transl. in Soviet Math. Dokl. **32** (1985).
3. ——, *The Morse index and the Maslov index for the extremals of control systems*, Dokl. Akad. Nauk SSSR **287** (1986), 521–524; English transl. in Soviet Math. Dokl. **33** (1986).
4. ——, *Quasiextremality for control systems*, Itogi Nauki i Tekhniki: Sovremennye Problemy Mat.: Noveĭshie Dostizheniya, vol. 35, VINITI, Moscow, 1989, pp. 109–134; English transl. in J. Soviet Math. **55** (1991), no. 4.
5. Victor Guillemin and Shlomo Sternberg, *Geometric asymptotics*, Amer. Math. Soc., Providence, R.I., 1977.

Moscow

Received 5/MAR/90

Translated by H. H. McFADEN