Journal of Mathematical Systems, Estimation, and Control
Vol. 5, No. 1, 1995, pp. 1–31

On Abnormal Extremals for Lagrange Variational Problems*

A.A. Agrachev† A.V. Sarychev†

Key words: Lagrange problem, abnormal extremals, theory of 2nd variation, minimality conditions, index and nullity theorems

AMS Subject Classifications: 49K15, 90C30, 93B06, 93B29

1 Introduction

In the paper we are going to provide a systematic way of treating *abnormal extrema* in variational problems.

Suppose a point $\hat{x}$ of a Banach space X is a point of extremum for a smooth functional $\mathcal{J} : X \longrightarrow R$ under equality constraints $F(x) = 0$, where $F : X \longrightarrow Y$ is a smooth mapping of X into a finite-dimensional vector space Y. The *'Lagrange multiplier rule'* claims the existence of a nonzero pair $(\lambda_0, \lambda^*) \in (R \times Y^*)$ of Lagrange multipliers, such that

$$\lambda_0 \mathcal{J}'(\hat{x}) + \lambda^* F'(\hat{x}) = 0. \tag{1.1}$$

Here $\mathcal{J}'(\hat{x}) \in X^*$, is the gradient and $\lambda^* F'(\hat{x}) = (F'(\hat{x}))^* \lambda^* \in X^*$, where $F'(\hat{x}) : X \longrightarrow Y$ is the differential of F at $\hat{x}$, and $(F'(\hat{x}))^* : Y^* \longrightarrow X^*$ is its adjoint.

This *first-order optimality condition* is hardly considered to be satisfactory unless so-called *normality condition* holds. This last is nonvanishing of λ_0. For the above mentioned problem this normality can be provided, for example, by so-called *regularity condition*, which is: $\mathrm{Im} F'(\hat{x}) = Y$. In nonlinear programming the normality condition is provided by Slater condition. In normal case one can renormalize the Lagrange multipliers and equation (1.1) in such a way that $\lambda_0 = 1$.

---

*Received March 23, 1993; received in final form July 12, 1994. Summary appeared in Volume 5, Number 1, 1995.

†This work was partially supported by the A. von Humboldt Foundation, Germany and by Deutschen Forschungsgemeinschaft (Schwerpunktprogramm 'Anwendungsbezogene Optimierung und Steuerung').

The points which meet the condition (1.1) but with vanishing Lagrange multiplier λ_0 are called *abnormal extremals*. One often tries to avoid them either looking for another set of Lagrange multipliers with nonvanishing λ_0 or simply by treating the abnormal case as a degeneration of constraints (see for example [13]) which has little to do with the extremality of $\hat{x}$.

Banal two-dimensional examples, such like

$$x_1 \longrightarrow \min, \ x_1^2 + \sin^2 x_2 = 0,$$

demonstrate a possibility for abnormal extremal points to be isolated (locally nonvariable) points of the set $\{x \mid F(x) = 0\}$. Traditional approach to optimization problems treats the functional and the constraints in different ways and hence the impossibility of varying the point $\hat{x}$ along the set $\{x \mid F(x) = 0\}$ is considered as a pathology which makes further analysis senseless. The same point of view has been adopted in calculus of variations (see [19], where these cases are named 'sad facts of life') and came inherently to optimal control.

Therefore the main activity in the field was directed to elimination of abnormal extremals; they either should not exist, or should not be optimal. Some examples of such activity can be found in sub-Riemannian geometry, which treats length functional along paths which are tangent to a completely nonintegrable (nonholonomic) vector distribution on a Riemannian manifold M. Preprint [16] of R. Montgomery lists several (given by different authors) false proofs of the fact, that minimizing sub-Riemannian geodesic should not be abnormal extremal. The preprint contains also an example of minimizing abnormal sub-Riemannian geodesic.

In this paper we investigate the phenomenon of abnormality from the point of view of geometric control theory. The main claim is: abnormal extremals exist, they can be optimal and optimality conditions for them are not worse, than in the normal case, although they have different meaning. Thus we show that 2-nd order sufficient optimality condition implies the local isolatedness of an abnormal extremal point $\hat{x}$.

We investigate firstly the problem of smooth minimization under equality constraints. Then we pass over to the abnormal extremals of the Lagrange problem of Calculus of Variations. Here we define the second variation along a corank 1 abnormal extremal and formulate second order necessary/sufficient conditions for weak optimality for abnormal extremals. Finally we present a method of computation for *Morse index* and *nullity* of abnormal extremals which play crucial role in the verification of the optimality conditions.

Our attention was attracted to the subject after a discussion on abnormal sub-Riemannian geodesics at the Conference 'Geometric Methods in Nonlinear Optimal Control' organized by IIASA in Sopron, Hungary in July 1991. One more source of inspiration was preprint [9] of B. Bonnard

and I. Kupka, which treats Legendre-Jacoby-type optimality conditions for time-optimal affine control problems. The authors are grateful to R.V. Gamkrelidze and H.W. Knobloch for their support. We are thankful to the anonymous referee who has pointed to an omission in the proof of the Theorem 3.2 and invested a lot of effort improving the style of the paper and correcting numerous grammatical errors. We also thank J.M. Sá Esteves for his help in LaTeX-drawing of the figures of Section 5.

2 Preliminaries

Below we use notation and technical tools of the chronological calculus developed by A.A. Agrachev and R.V. Gamkrelidze (see [4, 5]).

We will identify C^∞ diffeomorphisms $P : M \longrightarrow M$ with automorphisms of the algebra $C^\infty(M)$ of smooth functions on M: $\phi(\cdot) \longrightarrow P\phi = \phi(P(\cdot))$. The image of a point $q \in M$ under a diffeomorphism P will be denoted by $q \circ P$. C^∞ vector fields on M are 1st order differential operators on M or arbitrary derivations of the algebra $C^\infty(M)$, i.e. R-linear mappings $X : C^\infty(M) \longrightarrow C^\infty(M)$, satisfying the Leibnitz rule: $X(\alpha\beta) = (X\alpha)\beta + \alpha(X\beta)$. The value $X(q)$ of a vector field X at a point $q \in M$ lies in the tangent space T_qM to the manifold M at the point q. We denote by $[X^1, X^2]$ Lie bracket or commutator $X^1 \circ X^2 - X^2 \circ X^1$ of vector fields X^1, X^2. It is again a 1st order differential operator and if $X^1 = \sum_{i=1}^n X_i^1 \partial/\partial x_i$, $X^2 = \sum_{i=1}^n X_i^2 \partial/\partial x_i$ in local coordinates on M then the Lie bracket

$$[X^1, X^2] = \sum_{i=1}^n (\partial X_i^2/\partial x X^1 - \partial X_i^1/\partial x X^2)\partial/\partial x_i.$$

This operation introduces in the space of vector fields the structure of a Lie algebra denoted Vect M. For $X \in \mathrm{Vect}M$ the notation ad$X$ stands for the inner derivation of Vect M: $(\mathrm{ad}X)X' = [X, X'], \forall X' \in \mathrm{Vect}M$.

For a diffeomorphism P we use the notation AdP for the following inner automorphism of the Lie algebra Vect M: $\mathrm{Ad}PX = P \circ X \circ P^{-1} = P_\star^{-1}X$. The last notation stands for the result of translation of the vector field X by the differential of the diffeomorphism P^{-1}.

A flow on M is an absolutely continuous w.r.t. $\tau \in R$ curve $\tau \rightarrow P_\tau$ $(\tau \in R)$ in the group of diffeomorphisms Diff M, satisfying the condition $P_0 = I$ (where I is the identity diffeomorphism). We assume all time-dependent vector fields X_τ to be locally integrable with respect to τ. A time-dependent vector field X_τ defines an ordinary differential equation $\dot{q} = X_\tau(q(\tau)), q(0) = q^0$ on the manifold M; if solutions of this differential equation exist for all $q^0 \in M, \tau \in R$, then the vector field X_τ is called complete and defines a flow on M, being the unique solution of the (operator) differential equation:

$$dP_\tau/d\tau = P_\tau \circ X_\tau, P_0 = I. \tag{2.1}$$

The solution will be denoted by $P_t = \overrightarrow{\exp} \int_0^t X_\tau d\tau$, and is called (see [4]) a right chronological exponential of X_τ. If the vector field X_τ is time-independent ($X_\tau \equiv X$), then the corresponding flow is denoted by $P_t = e^{tX}$.

We introduce also Volterra expansion (or Volterra series) for the chronological exponential. It is (see [4]):

$$\overrightarrow{\exp} \int_0^t X_\tau d\tau \asymp I + \sum_{i=1}^{\infty} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \dots \int_0^{\tau_{i-1}} d\tau_i (X_{\tau_i} \circ \cdots X_{\tau_1}).$$

We will exploit only the terms of zero-, first- and second-order in this expansion, which are

$$\overrightarrow{\exp} \int_0^t X_\tau d\tau \asymp I + \int_0^t X_\tau d\tau + \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 (X_{\tau_2} \circ X_{\tau_1}) + \cdots \tag{2.2}$$

For time-independent X one obtains

$$e^{tX} \asymp I + tX + (t^2/2) X \circ X + \cdots \tag{2.3}$$

One more tool from the chronological calculus is the 'generalized variational formula' (see [4, 5] for a drawing):

$$\overrightarrow{\exp} \int_0^t (\hat{X}_\tau + X_\tau) d\tau =$$
$$= \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau \circ \overrightarrow{\exp} \int_0^t \mathrm{Ad}(\overrightarrow{\exp} \int_t^\tau \hat{X}_\theta d\theta) X_\tau d\tau. \tag{2.4}$$

Applying the operator $\mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta)$ to a vector field Y and differentiating $\mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) Y = (\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) \circ Y \circ (\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta)^{-1}$ w.r.t. τ one comes to the equality (see [4, 5]):

$$\frac{d}{d\tau} \mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta Y) = \mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) \,\mathrm{ad}\, \hat{X}_\tau Y, \tag{2.5}$$

which is of the same form as (2.1). Therefore $\mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta)$ can be presented as an operator chronological exponential $\overrightarrow{\exp} \int_0^t \mathrm{ad}\, \hat{X}_\theta d\theta$ which for a time-independent vector field $\hat{X}_\tau \equiv \hat{X}$ can be written as $e^{t \,\mathrm{ad}\, \hat{X}}$.

We also have to introduce some notions of symplectic geometry (see [6, 10, 15] for more details). A *symplectic structure* in an even-dimensional linear space Σ is defined by a nondegenerate skew symmetric bilinear form $\sigma(\cdot,\cdot)$. Two vectors $\xi_1, \xi_2 \in \Sigma$ are called skew orthogonal, written $\xi_1 \flat \xi_2$, if $\sigma(\xi_1, \xi_2) = 0$. If N is a subspace of Σ, let us denote by $N^\flat$ its skew orthogonal complement: $N^\flat = \{\xi \in \Sigma \mid \sigma(\xi, \nu) = 0, \forall \nu \in N\}$. Evidently $\dim N + \dim N^\flat = \dim \Sigma$. A subspace $, \subseteq \Sigma$ is called *isotropic*, when $, \subseteq\ ,^\flat$, and *coisotropic*, when $, \supseteq\ ,^\flat$. A subspace $\Lambda \subset \Sigma$ is called *Lagrangian*, when $\Lambda^\flat = \Lambda$. Such subspaces have dimension $\frac{1}{2}\dim \Sigma$.

The symplectic group $Sp(\Sigma)$ is the group of those linear transformations of Σ which preserve the symplectic form:

$$Sp(\Sigma) = \{S \in GL(\Sigma) \mid \sigma(S\xi_1, S\xi_2) = \sigma(\xi_1, \xi_2)\ \forall \xi_1, \xi_2 \in \Sigma\}.$$

The elements of this group are called symplectic transformations. The Lie algebra of the symplectic group is:

$$sp(\Sigma) = \{A \in gl(\Sigma) \mid \sigma(A\xi_1, \xi_2) = \sigma(A\xi_2, \xi_1)\ \forall \xi_1, \xi_2 \in \Sigma\}.$$

Let H be a real quadratic form on Σ and $d_\xi H$ be the differential of H at a point $\xi \in \Sigma$. Then $d_\xi H$ is a linear form on Σ which depends linearly on ξ. For every $\xi \in \Sigma$ there exists a unique vector $\overrightarrow{H}(\xi) \in \Sigma$ which satisfies the equality $\sigma(\overrightarrow{H}(\xi), \cdot) = d_\xi H$. It is easy to show that the linear operator $\overrightarrow{H}\colon \Sigma \to \Sigma$ belongs to $sp(\Sigma)$, and the mapping $H \to \overrightarrow{H}$ maps the space of quadratic forms onto $sp(\Sigma)$ isomorphically. The differential equation $\dot{\xi} = \overrightarrow{H}(\xi)$ is called the linear Hamiltonian system corresponding to the quadratic Hamiltonian H.

Denote by $\mathcal{L}(\Sigma)$ the *Lagrangian Grassmanian*, i.e. the set of Lagrangian planes in Σ. This is a smooth compact submanifold of of the Grassmanian of all $(n+1)-$dimensional subspaces of Σ. Its dimension is $\frac{1}{8}\dim\Sigma(\dim\Sigma + 2)$. Symplectic transformations transform Lagrangian planes into Lagrangian ones, hence the symplectic group acts on $\mathcal{L}(\Sigma)$. It is easy to show that it acts transitively.

If Λ is a Lagrangian plane and $,$ is isotropic, then it is easy to prove, that $(\Lambda \cap\ ,^\flat) +\ , = (\Lambda +\ ,) \cap\ ,^\flat$ is a Lagrangian plane. We denote it by $\Lambda^{,}$. The mapping $\Lambda \to \Lambda^{,}$ is continuous on a subset of $\mathcal{L}(\Sigma)$ where $\dim(\Lambda \cap\ ,) = \mathrm{const}$.

Let us consider the tangent space $T_\Lambda \mathcal{L}(\Sigma)$ at $\Lambda \in \mathcal{L}(\Sigma)$. To every quadratic form h on Σ there corresponds a linear Hamiltonian vector field $\overrightarrow{h}$ and a one-parameter subgroup $t \to e^{t\overrightarrow{h}}$ in $Sp(\Sigma)$. Let us consider the linear mapping

$$h \longrightarrow d(e^{t\overrightarrow{h}}\Lambda)/dt\ |_{t=0}$$

of the space of quadratic forms to $T_\Lambda \mathcal{L}(\Sigma)$. This mapping is surjective and its kernel consists of all quadratic forms which vanish on Λ. Thus two different quadratic forms correspond to the same vector from $\mathcal{T}_\Lambda \mathcal{L}(\Sigma)$ if and only if the restrictions of these forms on Λ coincide. Hence we obtain a natural identification of the space $\mathcal{T}_\Lambda \mathcal{L}(\Sigma)$ with the space of quadratic forms on Λ.

A tangent vector $\eta \in \mathcal{T}_\Lambda \mathcal{L}(\Sigma)$ is called nonnegative if the corresponding quadratic form is nonnegative on Λ. An absolutely continuous curve Λ_τ ($\tau \in [0,T]$) in $\mathcal{L}(\Sigma)$ is called nondecreasing if the velocities $\dot{\Lambda}_\tau \in \mathcal{T}_{\Lambda_\tau}\mathcal{L}(\Sigma)$ are nonnegative for almost all $\tau \in [0,T]$.

Treating the action of symplectic group $Sp(\Sigma)$ on $\mathcal{L}(\Sigma)$ one can easily verify, that *pairs* of Lagrangian planes (Λ, Λ') have only one invariant w.r.t. this action, namely $\dim(\Lambda \cap \Lambda')$. For *triples* of Lagrangian planes, there are more invariants.

Let $\Lambda_1, \Lambda_2, \Lambda_3$ be Lagrangian planes. Let us present a vector $\lambda \in (\Lambda_1 + \Lambda_3) \cap \Lambda_2$ as a sum $\lambda = \lambda_1 + \lambda_3$ and consider on $(\Lambda_1 + \Lambda_3) \cap \Lambda_2$ a quadratic form $\beta(\lambda) = \sigma(\lambda_1, \lambda_3)$. The *Maslov index* of the triple $(\Lambda_1, \Lambda_2, \Lambda_3)$ is the signature of $\beta(\lambda)$. It is invariant under the action of symplectic group.

In [1] a slightly different invariant was exploited for computation of Morse indices of singular extremals.

Definition 2.1 *Consider the quadratic form* $\beta(\lambda) = \sigma(\lambda_1, \lambda_3)$, *which is properly defined on the space* $((\Lambda_1 + \Lambda_3) \cap \Lambda_2)/\bigcap_{i=1}^3 \Lambda_i$. *The sum* $\frac{1}{2}\dim\ker\beta + \mathrm{ind}\beta$, *where* $\mathrm{ind}\beta$ *is the negative inertia index of* β, *is called the modified Maslov index of the triple* $(\Lambda_1, \Lambda_2, \Lambda_3)$ *of Lagrangian planes and will be denoted by* $\mathrm{ind}_{\Lambda_2}(\Lambda_1, \Lambda_3)$. □

Let us note, that $\ker\beta = ((\Lambda_1 \cap \Lambda_2) + (\Lambda_2 \cap \Lambda_3))/\bigcap_{i=1}^3 \Lambda_i$. We refer to [1] for a simple formula connecting this modified Maslov index with Maslov index of the triple and for the proof of the following 'triangle inequality':

$$\mathrm{ind}_{\Lambda_0}(\Lambda_1, \Lambda_3) \leq \mathrm{ind}_{\Lambda_0}(\Lambda_1, \Lambda_2) + \mathrm{ind}_{\Lambda_0}(\Lambda_2, \Lambda_3).$$

It also follows directly from the definition, that

$$\mathrm{ind}_{\Lambda_1}(\Lambda_1, \Lambda_3) = \frac{1}{2}\dim\ker\beta = \frac{1}{2}(\dim\Lambda_1 - \dim(\Lambda_1 \cap \Lambda_3)). \qquad (2.6)$$

Definition 2.2 *A continuous curve* $\Lambda(\tau) \in \mathcal{L}(\Sigma)$, $0 \leq \tau \leq 1$, *is called simple if there exists* $\Delta \in \mathcal{L}(\Sigma)$ *such that* $\Lambda(\tau) \cap \Delta = 0\ \forall \tau \in [0,1]$. □

Lemma 2.1 *If* $\Lambda(\tau) \in \mathcal{L}(\Sigma)$ $0 \leq \tau \leq 1$, *is a simple nondecreasing curve in* $\mathcal{L}(\Sigma)$, *and* $\Pi \in \mathcal{L}(\Sigma)$, *then*

$$\mathrm{ind}_\Pi(\Lambda(0), \Lambda(1)) = \mathrm{ind}_\Pi(\Lambda(0), \Lambda(\tau)) + \mathrm{ind}_\Pi(\Lambda(\tau), \Lambda(1)),\ \forall \tau \in [0,1]. \square$$

Lemma 2.2 *Let* $\Lambda^0, \Lambda^1 \in \mathcal{L}(\Sigma)$. *There exist* $\Delta \in \mathcal{L}(\Sigma)$ *and neighborhoods* $V^0 \ni \Lambda^0$, $V^1 \ni \Lambda^1$ *in* $\mathcal{L}(\Sigma)$ *such that whenever* $\Lambda \in V^0, \Lambda' \in V^1$ *and* $\dim(\Lambda \cap \Lambda') = \dim(\Lambda^0 \cap \Lambda^1)$ *then there exists a simple nondecreasing curve* $\Lambda(\tau)$, $\tau \in [0,1]$ *such that* $\Lambda(0) = \Lambda, \Lambda(1) = \Lambda'$, $\Lambda(\tau) \cap \Delta = 0\ \forall \tau \in [0,1]$. □

Both lemmas are proved in [1].

Definition 2.3 *Let* $\Lambda(t)$, $0 \le t \le T$, *be a nondecreasing curve in* $\mathcal{L}(\Sigma)$ *and* $0 = t_0 < t_1 < \cdots < t_l = T$ *be a partition of* $[0,T]$ *such that the curves* $\Lambda(\cdot)\,|_{[t_i,t_{i+1}]}$, $i = 0, \ldots l-1$ *are simple and* $\Pi \in \mathcal{L}(\Sigma)$. *The expression*

$$\mathrm{ind}_\Pi \Lambda(\cdot) = \sum_{i=0}^{l-1} \mathrm{ind}_\Pi(\Lambda(t_i), \Lambda(t_{i+1})) \tag{2.7}$$

is called the Maslov index of the curve $\Lambda(t)$ *with respect to* Π. □

It follows from the Lemma 2.1 that (2.7) does not depend on a choice of $t_1 < \cdots < t_{l-1}$. If the curve $\Lambda(t)$ is closed ($\Lambda(0) = \Lambda(T)$), then $\mathrm{ind}_\Pi \Lambda(\cdot)$ does not depend on Π (cf. [1]).

3 Smooth Extremal Problem with Equality Constraints: Abnormal Case

Let $\mathcal{J} : X \longrightarrow Y$ be a continuously Frechet differentiable function on a Banach space X, while $F : X \longrightarrow Y$ be a continuously Frechet differentiable mapping of X into a finite-dimensional vector space Y. Let us consider the following extremal problem

$$\mathcal{J}(x) \longrightarrow \min, \tag{3.1}$$

on a set $S \subset X$ given by an equation

$$F(x) = 0. \tag{3.2}$$

The Lagrange multiplier rule says, that if $\hat{x} \in X$ supplies the minimum to problem (3.1)-(3.2), then there exists a nonzero element $(\lambda_0, \lambda^*) \in (R \times Y^*)$, such that $\hat{x}$ is a critical point of the Lagrangian $\mathcal{L} = \lambda_0 \mathcal{J}(x) + \lambda^* F(x)$. In other words

$$\lambda_0 \mathcal{J}'(\hat{x}) + \lambda^* F'(\hat{x}) = 0. \tag{3.3}$$

(If the point $\hat{x}$ is a *regular* point of F, i.e. $\mathrm{Im} F'(\hat{x}) = Y$, then $\lambda_0 \neq 0$ and dividing (3.3) by λ_0 we come to the canonical pair $(1, \lambda^*)$ of Lagrange multipliers.)

Second-order optimality conditions for the problem (3.1)-(3.2) usually preassume the regularity condition. If $\mathcal{J}, F$ are twice Frechet differentiable on X and (3.3) holds at a regular point $\hat{x}$, with $\lambda_0 = 1$, then the nonnegativeness of the quadratic form $\mathcal{L}_{xx}(\hat{x}, 1, \lambda^*)$ on $\ker F'(\hat{x}) = \{\xi \in X \mid F'(\hat{x})\xi = 0\}$ is necessary for optimality of $\hat{x}$, while positive definiteness ($\mathcal{L}_{xx}(\hat{x}, 1, \lambda^*)(\xi, \xi) \ge \gamma \|\xi\|^2$, $\forall \xi \in \ker F'(\hat{x})$, and some $\gamma > 0$) is sufficient for local optimality of $\hat{x}$.

Let us investigate what happens, when $\hat{x}$ does not satisfy the normality assumption. Namely we assume, that λ_0 may vanish in (3.3). Still the above mentioned sufficient optimality condition is true also in this case.

Theorem 3.1 (Sufficient Optimality Condition) *Let $\mathcal{J}, F$ be twice Frechet differentiable on X. If for some $(\lambda_0, \lambda^*) \in (R_+ \times Y^*)$ the point $\hat{x}$ is a critical point of the Lagrangian $\mathcal{L}(x) = \lambda_0 \mathcal{J}(x) + \lambda^* F(x)$ and the quadratic form $\mathcal{L}''_{xx}(\hat{x})(\xi, \xi)$ is positive definite on* $\ker F'(\hat{x})$, *then $\hat{x}$ supplies a local minimum to the problem (3.1)-(3.2). If in addition $\lambda_0 = 0$, then $\hat{x}$ is an isolated point of the set $S = \{x \mid F(x) = 0\}$.* □

Remark. For $\lambda_0 \neq 0$ this is a standard fact (see [11]). It is definitely known also for $\lambda_0 = 0$, but we could not find a corresponding source for a reference. □

We will weaken now the requirement of positive definiteness, assuming that Banach space X is densely embedded into a separable Hilbert space $H: X \hookrightarrow H$ and $\mathcal{L}''_{xx}(\hat{x})$ is positive definite w.r.t. the norm of H. The relevant example is $L^r_\infty[0,T]$, trivially densely embedded into $L^r_2[0,T]$.

Theorem 3.2 (Modified Sufficient Optimality Condition) *Let a Banach space X be densely embedded into a separable Hilbert space $H: X \hookrightarrow H$. Let $\hat{x} \in X$ be such a point, that: i) $F(\hat{x}) = 0$, ii) $\mathcal{J}, F$ are Frechet differentiable at $\hat{x}$, and iii) for some $(\lambda_0, \lambda^*) \in (R_+ \times Y^*)$ the differential of Lagrange function $\mathcal{L}(x) = \lambda_0 \mathcal{J}(x) + \lambda^* F(x)$ vanishes at $\hat{x}$ ($\hat{x}$ is critical point of $\mathcal{L}(x)$). Let*

$$\|F(\hat{x} + x) - F(\hat{x}) - F'(\hat{x})x\| = O(1)\|x\|^2_H, \ \text{as } \|x\|_X \to 0, \tag{3.4}$$

and the Lagrange function $\mathcal{L}(x)$ admit Taylor expansion at $\hat{x}$ of the form:

$$|\mathcal{L}(\hat{x} + x) - \mathcal{L}(\hat{x}) - \frac{1}{2}\mathcal{L}''_{xx}(\hat{x})(x,x)| = o(1)\|x\|^2_H, \ \text{as } \|x\|_X \to 0, \tag{3.5}$$

with the quadratic form $\mathcal{L}''_{xx}(\hat{x})(x,x)$ continuously extendable onto H.

If the quadratic form $\mathcal{L}_{xx}(\hat{x})(\xi, \xi)$ is $H-$positive definite on $\ker F'(\hat{x})$, *i.e.:*

$$\text{for some } \gamma > 0, \ \mathcal{L}''_{xx}(\hat{x})(\xi, \xi) \geq 2\gamma\|\xi\|^2_H, \ \forall \xi \in \ker F'(\hat{x}), \tag{3.6}$$

then $\hat{x}$ supplies strict local minimum to the problem (3.1)-(3.2). If, in addition, $\lambda_0 = 0$, then $\hat{x}$ is an isolated point of the set $S = \{x \mid F(x) = 0\}$.
□

Proof: We start with the abnormal case: $\lambda_0 = 0$, that implies $\mathcal{L}(\hat{x}) = \lambda^* F(\hat{x}) = 0$. Without loss of generality we may assume, that $\hat{x}$ coincides with the origin of X.

We are going to establish (for some $b > 0$) an estimate $\|F(x)\| \geq b\|x\|_H^2$ for all x from some small (in X) neighborhood of the origin.

Let us present X as a direct sum of $\ker F'(0)$ and a (finite-dimensional) complement Z, which is mapped isomorphically by $F'(0)$ onto the image $F'(0)X$. Any $x \in X$ can be split uniquely into the sum $x_1 + x_0$, where $x_1 \in Z$, $x_0 \in \ker F'(0)$. Obviously

$$\text{for some } \mu > 0: \ \|x\|_H^2 \leq \mu(\|x_0\|_H^2 + \|x\|_H^2), \ \forall x \in X, \tag{3.7}$$

and

$$\text{for some } c > 0: \ \|F'(0)x\| = \|F'(0)x_1\| \geq c\|x_1\|_H, \ \forall x \in X. \tag{3.8}$$

Denote $N = \{y \in Y | \lambda^* y = 0\}$ and choose a vector $\eta \in Y$ such, that $\lambda^*\eta = 1$. Evidently $Y = R\eta \oplus N$ and $\mathrm{Im}F'(0) \subseteq N$. The value of the mapping $F(x)$ can be split into two addends:

$$F(x) = (\lambda^* F(x))\eta + R_N(x)$$

with

$$R_N(x) = F'|_0 x_1 + B_N(x, x),$$

taking its values in N. Evidently, for some $a > 0$:

$$\|F(x)\| \geq a(|\lambda^* F(x)| + \|R_N(x)\|), \ \forall x \in X.$$

In virtue of (3.5)

$$\begin{aligned}\lambda^* F(x) &= \mathcal{L}(x) - \mathcal{L}(\hat{x}) \\ &= \frac{1}{2}\mathcal{L}''_{xx}(\hat{x})(x, x) + o(1)(\|x_0\|_H^2 + \|x_1\|_H^2), \ \text{as } \|x\|_X \to 0,\end{aligned}$$

while the continuity of $\mathcal{L}''_{xx}(\hat{x})(x, x)$ in the norm of H implies

$$|\mathcal{L}''_{xx}(\hat{x})(x, x) - \mathcal{L}''_{xx}(\hat{x})(x_0, x_0)| = O(1)\|x\|_H\|x_1\|_H,$$

providing us with an estimate:

$$\begin{aligned}\lambda^* F(x) = \frac{1}{2}\mathcal{L}''_{xx}(\hat{x})(x_0, x_0) + o(1)(\|x_0\|_H^2 + \|x_1\|_H^2) + O(1)\|x\|_H\|x_1\|_H = \\ = \frac{1}{2}\mathcal{L}''_{xx}(\hat{x})(x_0, x_0) + o(1)\|x_0\|_H^2 + O(1)\|x\|_H\|x_1\|_H, \text{as } \|x\|_X \to 0.\end{aligned}$$

Choosing $\beta > 0$ and a neighborhood V of the origin in X, where the rest terms admit upper estimates $\delta\|x_0\|_H^2$ (with $\delta \leq \gamma/2$) and $\beta\|x\|_H\|x_1\|_H$ correspondingly, we derive from (3.6) an estimate:

$$\mathcal{L}(x) - \mathcal{L}(\hat{x}) = \lambda^* F(x) \geq \max(0, \frac{\gamma}{2}\|x_0\|_H^2 - \beta\|x\|_H\|x_1\|_H), \ \forall x \in V. \tag{3.9}$$

In virtue of (3.4) we may assume for some $\alpha' > 0$ an estimate $\|B_N(x,x)\| \leq \alpha' \|x\|_H^2$ to hold for any $x \in V$. Together with (3.8) this provides for $\alpha = \alpha'\mu$:

$$\|R(x)\| \geq \max(0, c\|x_1\|_H - \alpha'\|x\|_H^2) \geq \max(0, c\|x_1\| - \alpha(\|x_0\|_H^2 + \|x_1\|_H^2)).$$

Without loss of generality we may assume that $\forall x \in V$:

$$\|x\|_H \leq \epsilon,\ \|x_0\|_H \leq \epsilon,\ \|x_1\|_H \leq \epsilon,\ \alpha\epsilon \leq c/2,\ 8\alpha\beta\epsilon/c \leq \gamma/4,\ c/4\alpha\epsilon \geq 1.$$

Then

$$\|R(x)\| \geq \max(0, (c - \alpha\epsilon)\|x_1\|_H - \alpha\|x_0\|_H^2) \geq \\ \geq \max(0, \frac{c}{2}\|x_1\|_H - \alpha\|x_0\|_H^2). \tag{3.10}$$

The two estimates (3.10) and (3.9) imply for any $x \in V$

$$\|F(x)\| \geq a(\max(0, \frac{c}{2}\|x_1\|_H - \alpha\|x_0\|_H^2) + \max(0, \frac{\gamma}{2}\|x_0\|_H^2 - \beta\|x\|_H\|x_1\|_H).$$

Assume firstly that $c\|x_1\|_H \geq 8\alpha\|x_0\|_H^2$. Then

$$\|F(x)\| \geq a(\frac{c}{2}\|x_1\|_H - \alpha\|x_0\|_H^2) \geq$$

$$a(\frac{c}{4}\|x_1\|_H + \alpha\|x_0\|_H^2) \geq a(\frac{c}{4\epsilon}\|x_1\|_H^2 + \alpha\|x_0\|_H^2).$$

If, on the contrary, $c\|x_1\|_H \leq 8\alpha\|x_0\|_H^2$, then:

$$\|F(x)\| \geq a\max(0, \frac{\gamma}{2}\|x_0\|_H^2 - \beta\|x\|_H\|x_1\|_H) \geq \\ \geq a\max(0, \frac{\gamma}{2}\|x_0\|_H^2 - \beta\epsilon\frac{8\alpha}{c}\|x_0\|_H^2) \geq a\frac{\gamma}{4}\|x_0\|_H^2 \geq \\ \geq a\frac{\gamma}{8}(\|x_0\|_H^2 + \frac{c}{8\alpha\epsilon}\|x_1\|_H^2).$$

For the normal case $\lambda_0 \neq 0$ ($\lambda_0 = 1$) the proof can be proceeded along the same line. To this end let us consider an extended mapping $\Phi = (\mathcal{J}(x) - \mathcal{J}(\hat{x}), F(x))$ in place of $F(x)$ and $(1, \lambda^*)$ in place of λ^*. Repeating just presented reasoning we conclude that, under the conditions of the theorem, $\hat{x}$ is an isolated (in X) point of the set $\Phi^{-1}(0)$, and for some $b > 0$ $\|\Phi(x)\| \geq b\|x - \hat{x}\|_H^2$ in a small enough neighborhood of $\hat{x}$ in X. This last estimate implies obviously $|\mathcal{J}(x) - \mathcal{J}(\hat{x})| \geq b\|x - \hat{x}\|_H^2$, whenever $F(x) = 0$. Given, in addition (see (3.9)), $\mathcal{J}(x) - \mathcal{J}(\hat{x}) = \mathcal{L}(x) - \mathcal{L}(\hat{x}) \geq 0$, we come to the conclusion of the theorem. □

When setting 2nd-order *necessary* optimality conditions, one should distinguish in both, normal and abnormal, situations the cases, when the range of the differential $(\mathcal{J}'(\hat{x}), F'(\hat{x}))$ at a critical point $\hat{x}$ of the mapping $(\mathcal{J}, F)$ has codimension 1, and when this range has codimension $\nu > 1$.

In the second case the investigation of optimality amounts essentially to investigation of the images of R^ν-valued quadratic forms (see [2] for treatment of the subject); for $\nu > 2$ the images of these forms in R^ν *can be nonconvex* and one can not use standard separation arguments for setting optimality conditions. We restrict ourselves to the *abnormal corank 1 case;* namely we assume

$$\text{codim Im}(\mathcal{J}'(\hat{x}), F'(\hat{x})) = 1 \text{ in } R \times Y.$$

Actually this means, that there is unique (up to a constant multiplier) pair (λ_0, λ^*) of Lagrange multipliers. The following proposition is almost evident.

Proposition 3.3 *Let $\hat{x}$ be a local minimizer for the problem (3.1)-(3.2) and $\hat{x}$ is twice Frechet differentiable at $\hat{x}$. Then for some nonzero pair of Lagrange multipliers $(\lambda_0, \lambda^*) \in R_+ \times Y^*$ the point $\hat{x}$ is a critical point of the Lagrangian $\mathcal{L} = \lambda_0 \mathcal{J}(x) + \lambda^* F(x)$, and if such a pair is (up to a constant multiplier) unique, then*

$$\mathcal{L}''_{xx}(\hat{x}, \lambda_0, \lambda^*)(\xi, \xi) \geq 0$$

on

$$\ker(\mathcal{J}'(\hat{x}), F'(\hat{x})) = \{\xi \in X | \mathcal{J}'(\hat{x})\xi = 0, F'(\hat{x})\xi = 0\}.$$

□

Proposition 3.3 is valid whether λ_0 vanishes or not. Actually if the quadratic form $\mathcal{L}''_{xx}(\hat{x}, \lambda_0, \lambda^*)$ is indefinite on $\ker(\mathcal{J}'(\hat{x}), F'(\hat{x}))$ in this corank 1 case, then $(\mathcal{J}, F)$ maps a neighborhood of $\hat{x}$ onto some neighborhood of $(\mathcal{J}(\hat{x}), 0)$ in $R \times Y$ and hence $\hat{x}$ can not be point of extremum.

Note an essential gap between the sufficient condition, given by Theorem 3.1, and the necessary condition of Proposition 3.3: the domains of the quadratic forms in these two statements differ by dimension 1. The following theorem removes this gap, giving true necessary optimality condition for an abnormal extremum.

Theorem 3.4 (Necessary Optimality Condition) *Let $\hat{x}$ be a local minimizer for the problem (3.1)-(3.2) and F be twice Frechet differentiable at $\hat{x}$. Then for some nonzero pair of Lagrange multipliers $(\lambda_0, \lambda^*) \in R_+ \times Y^*$, $\hat{x}$ is a critical point of Lagrangian $\mathcal{L} = \lambda_0 \mathcal{J}(x) + \lambda^* F(x)$ and if such pair is (up to a constant multiplier) unique, then*

$$\mathcal{L}''_{xx}(\hat{x}, \lambda_0, \lambda^*)(\xi, \xi) \geq 0,$$

on

$$\ker F'(\hat{x}) = \{\xi \in X | F'(\hat{x})\xi = 0\}. \quad ^1\Box$$

Proof: For $\lambda_0 \neq 0$ the theorem is standard.

When proving it for $\lambda_0 = 0$, one may assume without loss of generality, that $\hat{x}$ coincides with the origin of X. If $\lambda_0 = 0$, then $\dim \mathrm{Im} F'(0) = \dim Y - 1$, and λ^* is an annihilator of $\mathrm{Im} F'(0)$.

Assume, that the quadratic form $\mathcal{L}''_{xx}(0,0,\lambda^*) = \lambda^* F''|_0(\xi,\xi)$ is indefinite on $\ker F'(0)$, namely there are two such vectors $\xi_1, \xi_2 \in \ker F'(0)$, that

$$\lambda^* F''|_0(\xi_1,\xi_1) = -\lambda^* F''|_0(\xi_2,\xi_2) = 1.$$

It follows from the uniqueness of the multiplier that $\mathrm{rank}(\mathcal{J}'(0), F'(0)) = \dim Y$, rank $F'(0) = \dim Y - 1$, and $\mathcal{J}'(0)$ does not vanish identically on $\ker F'(0)$, i.e. there exists such a vector $\xi_0 \in \ker F'(0)$, that $\mathcal{J}'(0)\xi_0 \neq 0$. If $\epsilon > 0$ is small enough, then the vectors $\xi_1 + \epsilon\xi_0, \xi_2 + \epsilon\xi_0$ span a two-dimensional subspace $X^0 \subset \ker F'(0)$, on which the quadratic form $\lambda^* F''|_0(\xi,\xi)$ is indefinite, while $\mathcal{J}'(0)$ does not vanish identically on X^0.

Let us take again a subspace $Z \subset X$, which $F'(0)$ maps onto $ImF'(0)$ isomorphically. Let us put $\bar{Z} = Z + X^0$. We are going to prove, that 0 is not a point of local minimum for $\mathcal{J}|_{\bar{Z}\cap F^{-1}(0)}$.

In order to investigate the set $\bar{Z} \cap F^{-1}(0)$ let us note that corank $F'(0)|_{\bar{Z}} = 1$, and the Hessian of $F|_{\bar{Z}}$ is a nondegenerate indefinite quadratic form which coincides with the restriction of $\lambda^* F''|_0$ on X^0. It follows from the 'Parametric Morse Lemma' (see [7, pp.163-165]), that by proper coordinate changes in $\bar{Z}$ and Y one can transform $F|_{\bar{Z}}$ into the mapping

$$(z_1, \ldots z_{n-1}, x_1, x_2) \longrightarrow (z_1, \ldots z_{n-1}, x_1^2 - x_2^2),$$

(here x_1, x_2 coordinatize X^0). In these coordinates the intersection of $\bar{Z}$ with $F^{-1}(0)$ is given locally by the equation: $x_1 = \pm x_2$. Since $\mathcal{J}'(0)$ does not vanish identically on X^0, then the restriction of $\mathcal{J}$ on one of the two curves

$$\gamma^-(s) = (\underbrace{0,\ldots,0}_{n-1}, s, -s), \ \gamma^+(s) = (\underbrace{0,\ldots,0}_{n-1}, s, s),$$

has nonzero derivative at $s = 0$. Both curves lie in $X^0 \cap F^{-1}(0)$, and hence $\mathcal{J}|_{\bar{Z}\cap F^{-1}(0)}$ has no extremum at the origin of $\bar{Z}$. $\Box$

[1] In [8] it was proved (for weaker type of abnormal extremum), that negative index of $\mathcal{L}_{xx}(\hat{x},\lambda_0,\lambda^*)(\xi,\xi)$ at optimal point should not exceed corank$F'(\hat{x})$, which in our case is 1. Actually in this case the index must vanish.

4 Extremals for Lagrange Problem of Calculus of Variations

Let us consider Lagrange problem of calculus of variations

$$\mathcal{J}(T, u(\cdot)) = \int_0^T \varphi(q(\tau), u(\tau))d\tau \longrightarrow \min, \tag{4.1}$$

for a nonlinear control system

$$\dot{q} = f(q, u), q(0) = q^0, \tag{4.2}$$

with end-point condition

$$q(T) = q^1, \tag{4.3}$$

and *free final time* T. Here for given $u \in R^r$ $f(\cdot, u)$ is a C^∞ vector field on the $n-$dimensional manifold M, $f(q, u)$ is C^3 w.r.t. u, admissible controls $u(\tau) \in L^r_\infty[0, T]$.

We investigate whether a given time T, an admissible control $\hat{u}(\cdot)$ and corresponding trajectory $\hat{q}(\cdot)$, meeting the conditions (4.2)- (4.3), supply L_∞-local or *weak* minimum for the problem (4.1)-(4.3).

We assume $\hat{u}(\cdot)$ to be continuous at $T-0$. The prolongation of $\hat{u}(\cdot)$ from $[0, T]$ onto $[0, T+\delta]$ by the constant $\hat{u}(T)$ will be denoted also by $\hat{u}(\cdot)$. We assume that the solution $\hat{q}(\cdot)$ of the equation $\dot{q} = f(q, \hat{u}(t))$, $q(0) = q^0$ exists on $[0, T + \delta]$. The weak minimality of the pair $(T, \hat{u}(\cdot))$ for the problem (4.1)-(4.3) means existence of a $\delta-$neighborhood $\mathcal{U}_\delta$ of $\hat{u}(\cdot)$ in $L^r_\infty[0, T+\delta]$ such that $\mathcal{J}(T', u(\cdot)) \geq \mathcal{J}(T, \hat{u}(\cdot))$ for any control $u(\cdot) \in \mathcal{U}_\delta$ which steers the system (4.2) from q^0 to q^1 in a time $T' \in (T - \delta, T + \delta)$.

We will introduce now classical 1st-order optimality condition (Euler-Lagrange equation) in Hamiltonian form

Theorem 4.1 *If $\hat{u}(\cdot)$ supplies a weak minimum for the problem (4.1)-(4.3), then there exists a nonzero pair $(\hat{\psi}_0, \hat{\psi}(\cdot))$ where $\hat{\psi}_0 \geq 0$ is a constant and $\hat{\psi}(\tau)$ is an absolutely continuous covector-function with domain $[0, T]$) such that the 5-tuple $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot), T)$:*

1) satisfies Hamiltonian system

$$\dot{q} = \partial H/\partial\psi, q(0) = q^0, \ q(T) = q^1, \tag{4.4}$$

$$\dot{\psi} = -\partial H/\partial q, \tag{4.5}$$

with a Hamiltonian

$$H = \hat{\psi}_0\varphi(q, u) + \langle\psi, f(q, u)\rangle; \tag{4.6}$$

2) meets the stationarity condition

$$\frac{\partial H}{\partial u}(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)) = 0, \text{ a.e. on } [0, T], \tag{4.7}$$

and transversality condition

$$H(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)) = 0, \; a.e. \; on \; [0, T]. \tag{4.8}$$

□

Definition 4.1 *The 5-tuple $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot), T)$ is called an extremal for the Lagrange problem (4.1)-(4.3). It is called a corank 1 extremal if the corresponding pair $(\hat{\psi}_0, \hat{\psi}(\cdot))$ is uniquely defined up to a scalar multiplier. The control $\hat{u}(\cdot)$ is called an extremal control.* □

It follows from Theorem 4.1 that for any extremal $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot), T)$ its restriction $(\hat{u}(\cdot)|_{[0,t]}, \hat{q}(\cdot)|_{[0,t]}, \hat{\psi}_0, \hat{\psi}(\cdot)|_{[0,t]}, t)$ to an interval $[0, t]$, $(0 < t \leq T)$ is also an extremal for the Lagrange problem (4.1)-(4.3).

Definition 4.2 *An extremal $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot), T)$ is called normal, if $\hat{\psi}_0 \neq 0$, and abnormal, if $\hat{\psi}_0 = 0$.* □

Remark: For an abnormal extremal the functional (4.1) does not enter the first-order optimality conditions.

Normal corank 1 extremals of problem (4.1)- (4.3) were intensively studied, and necessary and sufficient optimality conditions for them were established. Most of the results were obtained in the scope of *Theory of Second Variation,* developed by Legendre, Jacobi and M. Morse.

For the *Simplest Problem of the Calculus of Variations*

$$\mathcal{J} = \int_0^T \mathcal{F}(t, x(t), \dot{x}(t)) \longrightarrow \min, x(0) = x_0, x(T) = x_1,$$

(which has only corank 1 extremals) the scheme is as follows. Nonnegativeness and positive definiteness of the second variation are the necessary and sufficient optimality conditions correspondingly. These conditions can be formulated in terms of the *Morse Index* and *Nullity* (see the monography [17] of M. Morse) of an extremal, which are correspondingly the negative inertia index and the dimension of kernel of the second variation along the extremal.

The *Legendre Condition* $\mathcal{F}_{\dot{x}\dot{x}} \geq 0$ and the *Strong Legendre Condition* $\mathcal{F}_{\dot{x}\dot{x}}(\xi, \xi) \geq k||\xi||^2$, are correspondingly necessary and sufficient conditions for *finiteness* of the Morse Index.

When restricting the Second Variation to subspaces $C^1[0,t] \subset C^1[0,T]$ ($t \in [0,T]$) one obtains a family of quadratic forms depending on t. A point t_* is called a *conjugate point of multiplicity* ν for the extremal if nullity of the corresponding quadratic form equals ν. For the 'Simplest Problem of the Calculus of Variations' under Strong Legendre Condition conjugate points of an extremal are isolated, and the Morse Index equals to sum of the multiplicities of the conjugate points located on $(0,T)$ (Morse Index Theorem; see [17]). Finally, if the Morse Index and the Nullity of an extremal on $[0,T]$ both vanish, then the second variation along it is positive definite, and the extremal is optimal (Jacobi Sufficient Optimality Condition). Vice versa, if an extremal is optimal, then its Morse Index must vanish and hence there must be no conjugate points on $(0,T)$.

For the *Lagrange Problem* of the Calculus of Variations the situation becomes more complicated. On one hand the Legendre Condition may degenerate along a subarc of an extremal (this happens, for example, if the problem is affine w.r.t. control). On another hand the conjugate points of extremal can be nonisolated and even fill whole subintervals of the time axis (see [12]). This phenomenon is connected with violation of some regularity condition along the extremal (see [18]).

Finiteness of the Morse Index, which is necessary for optimality, can be guaranteed in this case, by the so-called *Generalized Legendre Conditions* (see [3, 14, 1] for their invariant setting). An algorithm for the computation of the Morse Index, when intervals of conjugate points present, was proposed in [18]. General computations which withstand various degenerations of Legendre and regularity conditions were presented in [1]. They are based on techniques of symplectic geometry.

Using these techniques one can compute the Index and the Nullity for normal extremals of the Lagrange Problem. Then the Strong Legendre Condition together with vanishing of both the Index and the Nullity for the given extremal guarantees the *positive definiteness of the second variation* which is sufficient for *weak optimality* of the extremal if the extremal *is normal*.

For *abnormal* extremals new effects appear. One of them is that an abnormal extremal can be *locally nonvariable*, namely, as for our problem (4.1)-(4.3), no admissible control different from $\hat{u}(\cdot)$ and in some small L_∞-neighborhood of $\hat{u}(\cdot)$ can steer the system (4.1) from point q^0 to q^1 in a time close to T. This possibility, looking disappointing, has inspired attempts to eliminate abnormal extremals from consideration by proving, that they either do not exist, or are not optimal, or 'are not better' than normal ones, i.e. that for any abnormal extremal of some Lagrange problem there exist normal extremal with the same value of cost.

What we are suggesting is on the contrary a systematic approach to investigation of the abnormal extremals of a Lagrange problem. The first step will be reduction of this problem to the one, which was treated in the previous section.

Let us introduce firstly *(time $\times$ input)/state mapping* (see [5]) for the system (4.2). Namely, let us consider the mapping $F : R_+ \times L^r_\infty[0,T] \longrightarrow M$, which maps a pair $(t, u(\cdot))$ into the point $x(t)$, where $x(\cdot)$ is the trajectory of the system (4.2), produced by the control $u(\cdot)$. Obviously, if t is fixed, then the image of $F(t,\cdot)$ is the *attainable set* of the system (4.2) in time t. Also $q^1 = F(T, \hat{u}(\cdot))$.

A well-known fact is that for $(T, \hat{u}(\cdot))$ to be optimal the point $(T, \hat{u}(\cdot))$ of $R \times L^r_\infty$ must be a critical point of the mapping $(\mathcal{J}, F)(t, u(\cdot))$, which maps $R \times L^r_\infty$ into $R \times M$. Indeed otherwise the system of equations

$$\mathcal{J}(t, \hat{u}(\cdot) + u(\cdot)) = \mathcal{J}(T, \hat{u}(\cdot)) - \epsilon, \;\; F(t, u(\cdot)) = q^1,$$

is solvable for any sufficiently small $\epsilon \geq 0$, and hence the system (4.2) can be steered from q^0 to q^1 with the value of the functional $\mathcal{J}$ equal to $\mathcal{J}(T, \hat{u}(\cdot)) - \epsilon < \mathcal{J}(T, \hat{u}(\cdot))$.

For $(T, \hat{u}(\cdot))$ to be a critical point for the above mentioned mapping is equivalent to the fact that $(T, \hat{u}(\cdot))$ is part of an extremal $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot), T)$. If $\hat{\psi}_0 = 0$, then the functional $\mathcal{J}$ does not enter the extremality condition and it follows that the pair $(T, \hat{u}(\cdot))$ is part of an abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ if and only if it is a critical point of the mapping F.

Let us put

$$\hat{f}_\tau(q) = f(q, \hat{u}(\tau)), \;\; g_\tau(q,u) = f(q,u) - \hat{f}_\tau(q).$$

Further we often write $\hat{f}_\tau$ and $g_\tau(u)$ instead of $\hat{f}_\tau(q)$ and $g_\tau(q,u)$ correspondingly. Then

$$F(t, u(\cdot)) = q^0 \circ \overrightarrow{\exp} \int_0^t (\hat{f}_\tau + g_\tau(u(\tau)))d\tau,$$

or in virtue of the generalized variational formula (2.4)

$$\begin{aligned} F(t, u(\cdot)) &= q^0 \circ \overrightarrow{\exp} \int_0^t \hat{f}_\tau d\tau \circ \overrightarrow{\exp} \int_0^t Y_{t,\tau}(u(\tau))d\tau = \\ &= \hat{q}(t) \circ \overrightarrow{\exp} \int_0^t Y_{t,\tau}(u(\tau))d\tau, \end{aligned} \tag{4.9}$$

where

$$Y_{t,\tau}(q,u) = \mathrm{Ad}\, \overrightarrow{\exp} \int_t^\tau \hat{f}_\theta d\theta g_\tau(q,u). \tag{4.10}$$

From the formula (2.5) it follows that

$$dY_{t,\tau}/dt = -\,\mathrm{ad}\, \hat{f}_t Y_{t,\tau}. \tag{4.11}$$

We will need the first and the second differentials of F at the point $(T, \hat{u}(\cdot)) \in R \times L_\infty^r$. Putting $Y_\tau = Y_{T,\tau}$ and taking the Taylor expansion of $Y_\tau(u)$ at the point $\hat{u}(\tau) \in R^r$:

$$Y_\tau(u) = Y_\tau^1 u + \frac{1}{2} Y_\tau^2(u,u) + \cdots, \tag{4.12}$$

where

$$Y_\tau^1 = \mathrm{Ad}\,\overrightarrow{\exp} \int_T^\tau \hat{f}_\theta d\theta \frac{\partial f}{\partial u}|_{\hat{u}(\tau)},\ Y_\tau^2 = \mathrm{Ad}\,\overrightarrow{\exp} \int_T^\tau \hat{f}_\theta d\theta \frac{\partial^2 f}{\partial^2 u}|_{\hat{u}(\tau)},$$

one obtains for the first differential of F at the point $(T, \hat{u}(\cdot))$:

$$F'|_{(T,\hat{u}(\cdot))}(\delta\theta, u(\cdot)) = \hat{f}_T(q^1)\delta\theta + \int_0^T Y_\tau^1(q^1) u(\tau) d\tau. \tag{4.13}$$

The pair $(T, \hat{u}(\cdot))$ is part of an abnormal extremal if and only if $\mathrm{Im} F' \neq \mathcal{T}_{q^1}M$. In this case there exists a nonzero covector $\psi_T \in T_{q^1}^* M$ such that:

$$\langle \psi_T, \hat{f}_T(q^1) \rangle = 0, \tag{4.14}$$

and

$$\forall u(\cdot) \in L_\infty^r[0,T] : \langle \psi_T, \int_0^T Y_\tau^1(q^1) u(\tau) d\tau \rangle = 0,$$

or in virtue of Dubois-Raymond Lemma:

$$\langle \psi_T, Y_\tau^1(q^1) \rangle = 0 \text{ a.e. on } [0,T]. \tag{4.15}$$

These conditions can be transformed in a standard way to the stationarity and transversality conditions (4.7)-(4.8) of Theorem 4.1 with 'abnormal' ($\lambda_0 = 0$!) Hamiltonian $H = \langle \psi, f(q,u) \rangle$; the covector ψ_T entering (4.14) and (4.15) corresponds to the end-point value $\hat{\psi}(T)$ of the solution of the adjoint equation (4.5). The equality (4.8) can be transformed into

$$\langle \hat{\psi}(\tau), \hat{f}_\tau(\hat{q}(\tau)) \rangle = \langle \hat{\psi}(\tau), f(\hat{q}(\tau), \hat{u}(\tau)) \rangle = 0, \quad \text{a.e. on } [0,T]. \tag{4.16}$$

For the abnormal case we set

Definition 4.3 *The first differential $F' : R \times L_\infty^r \longrightarrow \mathcal{T}_{q^1}M$, given by the formula (4.13) is called first variation along the abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$.* □

We now define the second variation along an abnormal extremal. It is the Hessian or quadratic differential of F at the critical point $(T, \hat{u}(\cdot)) \in R \times L_\infty^r$ (see [7]). Let us choose a function $\chi : M \longrightarrow R$, such that

$d\chi|_{q^1} = \psi_T$, and consider the function $\chi(F(t,u(\cdot)))$. In virtue of (4.13 - 4.15) the point $(T,\hat{u}(\cdot))$ is a critical point for this function.

Let us compute the quadratic term of the Taylor expansion of $\chi(F(t,u(\cdot)))$ at $(T,\hat{u}(\cdot))$. Appealing to the Taylor expansion (4.12) and to the Volterra expansion (2.2) for the right chronological exponential, we derive

$$((\frac{1}{2}\int_0^T Y_\tau^2(u(\tau),u(\tau))d\tau - \int_0^T [\hat{f}_T\delta\theta, Y_\tau^1 u(\tau)]d\tau + (\hat{f}_T\circ\hat{f}_T)\frac{\delta\theta^2}{2} + \\ + \int_0^T\int_0^\tau Y_\xi^1 u(\xi)d\xi \circ Y_\tau^1 u(\tau)d\tau + \hat{f}_T\delta\theta\circ\int_0^T Y_\tau^1 u(\tau)d\tau)\chi)(q^1). \tag{4.17}$$

(When proceeding with the computation one should take into account the equalities (4.11), (4.14) and (4.15)).

When restricting the quadratic form (4.17) to the kernel of F', we are able to subtract from (4.17) the vanishing quantity:

$$\frac{1}{2}((\hat{f}_T\delta\theta + \int_0^T Y_\tau^1 u(\tau)d\tau)\circ(\hat{f}_T\delta\theta + \int_0^T Y_\tau^1 u(\tau)d\tau)\chi)(q^1),$$

to obtain

$$\frac{1}{2}((\int_0^T Y_\tau^2(u(\tau),u(\tau))d\tau + \int_0^T [-\hat{f}_T\delta\theta + \int_0^\tau Y_\xi^1 u(\xi)d\xi, Y_\tau^1 u(\tau)]d\tau)\chi)(q^1).$$

The last expression does not depend on choice of χ but only on $\psi_T = d\chi|_{q^1}$. Hence we have

$$2F''|_{(T,\hat{u}(\cdot))}[\psi_T](\delta\theta,u(\cdot)) = \langle\psi_T, (\int_0^T Y_\tau^2(u(\tau),u(\tau))d\tau + \\ + \int_0^T [-\hat{f}_T\delta\theta + \int_0^\tau Y_\xi^1 u(\xi)d\xi, Y_\tau^1 u(\tau)]d\tau)(q^1)\rangle, \tag{4.18}$$

where $(\delta\theta, u(\cdot))$ satisfy an equation

$$\hat{f}_T(q^1)\delta\theta + \int_0^T Y_\tau^1(q^1)u(\tau)d\tau = 0. \tag{4.19}$$

Definition 4.4 *The quadratic form $F''|_{(T,\hat{u}(\cdot))}[\psi_T]$, defined by (4.18) - (4.19), is called the second variation along the abnormal extremal* $(\hat{u}(\cdot),\hat{q}(\cdot),0,\hat{\psi}(\cdot),T)$ □

Now we are prepared to formulate necessary-sufficient optimality conditions for abnormal extremals. These conditions are similar to the ones for normal extremals.

Theorem 4.2 (Necessary Optimality Condition) *If a corank 1 abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ supplies weak minimum to the Lagrange problem (4.1)-(4.3), then the second variation (4.18)- (4.19) along it is nonnegative.* □

Definition 4.5 *The second variation along an abnormal extremal is called positive definite, if on its domain there holds an inequality of the form:*

$$F''|_{(T,\hat{u}(\cdot))}[\psi_T](\delta\theta, u(\cdot)) \geq \alpha(\delta\theta^2 + \|u(\cdot)\|^2_{L_2}).$$

□

Theorem 4.3 (Sufficient Optimality Condition) *If the second variation along an abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ is positive definite, then the pair $(\hat{u}(\cdot), T)$ supplies weak minimum to the problem (4.1)-(4.3). Moreover for some $\epsilon > 0$ no other control in the ϵ- neighborhood of $\hat{u}(\cdot)$ in $L^r_\infty[0,T]$ is able to steer the system (4.2) from q^0 to q^1 in time $T' \in [T-\epsilon, T+\epsilon]$.* □

Proof of Theorem 4.2 is a direct corollary of the proof of Theorem 3.4. To realize it let us present the Lagrange problem (4.1)-(4.3) in the following form:

$$\mathcal{J}(t, u(\cdot)) \longrightarrow \min, \tag{4.20}$$

$$F(t, u(\cdot)) = q^1, \tag{4.21}$$

where $\mathcal{J}(t, u(\cdot))$ is given by (4.1) and F is the defined above mapping of the Banach space $R \times L^r_\infty[0,T]$ into M. Since our consideration is local, M can be identified with R^n and q^1 with the origin of R^n.

The (time×input)/state mapping is *not* smooth w.r.t. t, but its restriction onto the space of $C^\ell-$smooth controls $u(\cdot)$ *is* $C^\ell-$smooth. The Hessian of this restriction coincides with the restriction of the 2nd variation (4.18)-(4.19) onto $R \times C^\ell$. Corank of our abnormal extremal, which is by the definition corank of the differential $(\mathcal{J}', F')|_{(T,\hat{u}(\cdot))}$, is equal to 1. Therefore, by virtue of Theorem 3.4, the 2nd variation must be nonnegative on $R \times C^\ell$ and, by virtue of continuity, on $R \times L_2$ as well. □

Proof of Theorem 4.3 follows from the proof of Theorem 3.2 in the same way as the previous theorem follows from Theorem 3.4. Considering again the extremal problem (4.20)-(4.21) and applying to it Theorem 3.2, where the pair $X \hookrightarrow H$ is $R \times L^r_\infty \hookrightarrow R \times L^r_2$, we conclude that whenever the second variation (4.18) is L^r_2-positive definite on the kernel of the first variation (4.19), the point $(T, \hat{u}(\cdot))$ is an isolated point of the set $F^{-1}(q^1)$. □

5 Index and Nullity Theorems for Abnormal Extremals

In this section we are going to provide a method for computing Index and Nullity (see [17]) of the second variation (4.18)-(4.19). This will give a possibility of verifying necessary and sufficient optimality conditions for abnormal extremals established in the previous section (theorems 4.2 and 4.3).

If we put $\delta\theta = 0$ in formulae (4.18)-(4.19) for the second variation along a corank 1 abnormal extremal then the resulting quadratic form

$$2F_r''|_{(T,\hat{u}(\cdot))}[\psi_T](0,u(\cdot)) = \langle \psi_T, (\int_0^T Y_\tau^2(u(\tau),u(\tau))d\tau + \\ + \int_0^T [\int_0^\tau Y_\xi^1 u(\xi)d\xi, Y_\tau^1 u(\tau)]d\tau)(q^1)\rangle, \quad (5.1)$$

with domain consisting of those pairs $(0, u(\cdot))$, which satisfy

$$\int_0^T Y_\tau^1(q^1)u(\tau)d\tau = 0, \quad (5.2)$$

coincides with the Hessian of *input/state mapping* (see [5]) $u(\cdot) \to F(T, u(\cdot))$.

We will call (5.1)-(5.2) the *reduced second variation.* Its domain has codimension 1 or 0 in the domain of the second variation (4.18)-(4.19). Its index is not larger and differs at most by 1 from the index of the second variation. Starting with a formula for the index of the reduced second variation we derive from it an expression for the index of the second variation (4.18)-(4.19).

Let us start with the conditions of *finiteness of index*, which are evidently the same for the second variation and the reduced second variation.

It is known, for the index to be finite it is necessary, that for almost all $\tau \in [0,T]$ the quadratic forms $\langle \psi_T, Y_\tau^2(q^1)(u,u)\rangle$ on the space R^r of control parameters which enter both (4.18) and (5.1), must be nonnegative. This is the so-called *Legendre Condition* for extremal of Lagrange Problem. It is also known that if for some $\beta > 0$

$$\gamma_\tau(u,u) = \langle \psi_T, Y_\tau^2(q^1)(u,u)\rangle \geq \beta\|u\|^2 \text{ a.e. on } [0,\mathrm{T}], \quad (5.3)$$

then the indices of the reduced second variation and hence of the second variation are finite, and these variations are positive definite on some subspaces of finite codimension of their domains. The condition (5.3) is called a *Strong Legendre Condition* for an extremal.

If the Legendre Condition *degenerates*, i.e. $\langle\psi_T, Y^2_\tau(q^1)(u,u)\rangle = 0$, then one can proceed further and derive a series of Generalized Legendre Conditions which provide finiteness of index. Each condition of the series becomes effective, if all the previous conditions of the series degenerate. Here we treat only the nondegenerate abnormal case assuming that the Strong Legendre Condition (5.3) holds.

In what follows we use an approach developed in [1]. For computation of the Index and the Nullity we introduce a *symplectic representation* of the second variation (4.18)-(4.19). Let us put

$$V = \mathrm{span}\{\{\hat{f}_T(q^1)\} \cup \{Y^1_\tau(q^1) | \tau \in [0,T]\}\};$$

evidently $V \subset \mathcal{T}_{q^1}M$ coincides with the image $\mathrm{Im}F'$ of the first variation (4.13). It follows from (4.15)-(4.16), that ψ_T annihilates V.

Consider skew symmetric bilinear form on $\mathcal{E}_V$, the space of vector fields, whose values at q^1 lie in V :

$$\langle\psi_T, [X, X'](q^1)\rangle, \forall X, X' \in \mathcal{E}_V. \tag{5.4}$$

This form has a kernel of finite codimension in $\mathcal{E}_V$, which is defined by the following equalities:

$$X(q^1) = 0; \langle\psi_T, (\partial X/\partial\xi)(q^1)\rangle = 0, \forall \xi \in V.$$

Taking the quotient of $\mathcal{E}_V$ w.r.t. this kernel, one obtains finite-dimensional quotient space Σ with *nondegenerate* skew symmetric bilinear form $\sigma(\cdot,\cdot)$ induced from (5.4), that is a symplectic structure on Σ. Direct calculation gives us $\dim\Sigma = 2\dim V = 2(n-1)$. We denote by $\underline{X}$ the image of an $X \in \mathcal{E}_V$ under the canonical projection $\mathcal{E}_V \longrightarrow \Sigma$.

Choose local coordinates $(x_1, \ldots x_n) : \mathcal{O} \longrightarrow R^n$ on some neighborhood $\mathcal{O}$ of q^1 in M, such that $x_i(q^1) = 0$, $(i = 1, \ldots n)$, the subspace V is defined by the equality $x_n = 0$ and $\psi_T = (0, \ldots 0, \psi_n)$. Then the canonical projection $X \rightarrow \underline{X}$ is:

$$X = \sum_{i=1}^{n} X_i(x)\partial/\partial x_i \rightarrow \underline{X} =$$
$$= (X_1(0), \ldots X_{n-1}(0), \partial(\psi_n X_n)/\partial x_1(0), \ldots \partial(\psi_n X_n)/\partial x_{n-1}(0)).$$

The symplectic form is then:

$$\sigma(\underline{X}, \underline{Y}) = \sum_{j=1}^{n-1}(X_j(0)\partial(\psi_n Y_n)/\partial x_j(0) - Y_j(0)\partial(\psi_n X_n)/\partial x_j(0)).$$

Let Π be the image under the canonical projection of the space of those vector fields which vanish at q^1. Direct calculation proves, that Π is a Lagrangian subspace given in the coordinates by $\underline{X}_j(0) = 0$, $j = 1, \ldots n-1$.

Instead of notations $\underline{Y}_\tau^1$ and $\underline{\hat{f}_T}$ for the images of the vector fields Y_τ^1 and $\hat{f}_T$ under the canonical projection $\mathcal{E}_V \to \Sigma$ we use Υ_τ and $\underline{f}$. By the definitions of $\sigma(\cdot,\cdot)$, Π and γ_τ the second variation (4.18)-(4.19) is:

$$2F''|_{(T,\hat{u}(\cdot))}[\psi_T](\delta\theta, u(\cdot)) = \int_0^T \gamma_\tau(u(\tau), u(\tau))d\tau + \\ + \int_0^T \sigma(-\underline{f}\delta\theta + \int_0^\tau \Upsilon_\xi u(\xi)d\xi, \Upsilon_\tau u(\tau))d\tau, \tag{5.5}$$

and its domain is:

$$\{(\delta\theta, u(\cdot)) | \underline{f}\delta\theta + \int_0^T \Upsilon_\tau u(\tau)d\tau \in \Pi\}. \tag{5.6}$$

If we put $\delta\theta = 0$ in (5.5)-(5.6), then we obtain the symplectic presentation for the reduced second variation (5.1)-(5.2):

$$2F_r''|_{(T,\hat{u}(\cdot))}[\psi_T](0, u(\cdot)) = \int_0^T \gamma_\tau(u(\tau), u(\tau))d\tau + \\ + \int_0^T \sigma(\int_0^\tau \Upsilon_\xi u(\xi)d\xi, \Upsilon_\tau u(\tau))d\tau, \tag{5.7}$$

with domain

$$\{(0, u(\cdot)) | \int_0^T \Upsilon_\tau u(\tau)d\tau \in \Pi\}. \tag{5.8}$$

Certainly the domain (5.8) either is a codimension 1 subspace of the domain (5.6) or coincides with it.

We now introduce a Hamiltonian form of the Jacobi Equation for abnormal extremals of the Lagrange Problem (4.1)-(4.3) (see [1] for more details). Let $\bar{\gamma}_\tau$ be the nonsingular selfadjoint operator $\bar{\gamma}_\tau : R^r \longrightarrow R^{r^*}$, which induces the positive definite form $\gamma_\tau(u,u)$ on $R^r : \gamma_\tau(u,v) = \langle \bar{\gamma}_\tau u, v\rangle$, $\forall u, v \in R^r$. Define a bilinear form γ_τ^{-1} on R^{r^*} by $\gamma_\tau^{-1}(u^*, v^*) = \langle \bar{\gamma}_\tau^{-1} u^*, v^*\rangle$, $\forall u^*, v^* \in R^{r^*}$. Apparently for any $x \in \Sigma$ the mapping $u \longrightarrow \sigma(\Upsilon_\tau \cdot, x)$ defines a linear form on R^r, i.e. an element of R^{r^*}, which depends *linearly* on $x \in \Sigma$. This means, that the correspondence

$$x \longrightarrow \frac{1}{2}\gamma_\tau^{-1}(\sigma(\Upsilon_\tau \cdot, x))$$

defines a quadratic form on Σ.

Treating this quadratic form as a time-dependent Hamiltonian on Σ, one obtains on Σ the time-dependent linear Hamiltonian system:

$$\dot{x} = \Upsilon_\tau \bar{\gamma}_\tau^{-1}(\sigma(\Upsilon_\tau \cdot, x)), \tag{5.9}$$

which will be called the *Jacobi equation* for the abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ of the Lagrange Problem (4.1)- (4.3).

If for any $\tau \in [0, T]$ the vectors $u_1(\tau), \ldots u_r(\tau)$ form a basis of R^r such that $\gamma_\tau(u_i(\tau), u_j(\tau)) = \delta_{ij}$, $(i, j = 1, \ldots r)$ then the equation can be presented as

$$\dot{x} = \sum_{i=1}^{r} \sigma(\Upsilon_\tau u_i(\tau), x) \Upsilon_\tau u_i(\tau).$$

Since a Hamiltonian flow preserves the symplectic structure of Σ, the Jacobi equation transforms Lagrangian planes into Lagrangian ones. Therefore one may consider the Hamiltonian flow as a flow on the Lagrangian Grassmanian $\mathcal{L}(\Sigma)$. It is generated by the following time-dependent Hamiltonian system on $\mathcal{L}(\Sigma)$:

$$\dot{\Lambda} = \frac{1}{2} \gamma_\tau^{-1}(\sigma(\Upsilon_\tau \cdot, x))|_\Lambda$$

(see Section 2 for details).

Definition 5.1 *The Jacobi curve corresponding to the reduced second variation (5.7)-(5.8) is the trajectory $\tau \longrightarrow \Lambda_\tau$ of the Jacobi equation on the Lagrangian Grassmanian, with the initial condition $\Lambda_\tau|_{\tau=0} = \Pi$.* □

The following proposition, gives a formula for the index of the reduced second variation, i.e. of the quadratic form (5.7)-(5.8). It is a corollary of Theorem 1 in [1].

Proposition 5.1 *Let $\tau \longrightarrow \Lambda_\tau$ be the Jacobi curve corresponding to the reduced second variation (5.7)-(5.8) along a corank* 1 *abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ of the Lagrange problem (4.1)-(4.3). Then for any subdivision $\zeta_{m+1} = 0 = \zeta_0 < \zeta_1 < \cdots < \zeta_m = T$ of $[0, T]$, such that all arcs $\Lambda|_{[\zeta_i, \zeta_{i+1}]}$, $(i = 0, \ldots m-1)$ are simple (see the Definition 2.2), the negative index of the reduced second variation (5.7)-(5.8) along the extremal equals*

$$\sum_{i=0}^{m} \mathrm{ind}_\Pi(\Lambda_{\zeta_i}, \Lambda_{\zeta_{i+1}}) + \dim \cap_{\tau \in [0,T]} \Lambda_\tau - (n-1). \ \Box \qquad (5.10)$$

Starting from the formula (5.10) we now compute negative index of the second variation (5.5)-(5.6). To this end we use one more technical result (see [1]).

Proposition 5.2 *Assume, that a quadratic form $Q(\cdot, \cdot)$ is defined on a Hilbert space and is positive definite on a subspace of finite codimension. Let $Q_\mathcal{N}$ be the restriction of Q on a closed subspace $\mathcal{N}$ of the Hilbert space and $\mathcal{N}_Q^\perp$ be the Q-orthogonal complement to $\mathcal{N}$: $\mathcal{N}_Q^\perp = \{y | \mathcal{B}(x, y) =$*

$0, \forall x \in \mathcal{N}\}$, *where* $\mathcal{B}$ *is corresponding to* Q *symmetric bilinear form (i.e.* $Q(x) = \mathcal{B}(x,x)$*). Then:*

$$\operatorname{ind} Q = \operatorname{ind} Q_{\mathcal{N}} + \operatorname{ind} Q|_{\mathcal{N}_Q^{\perp}} + \dim(\mathcal{N} \cap \mathcal{N}_Q^{\perp}) - \dim(\mathcal{N} \cap \ker Q). \square \quad (5.11)$$

To apply the result in our case we define the Hilbert space $\mathcal{H}$ as the set of pairs $(\delta\theta, u(\cdot))$ meeting the condition (5.6), while the subspace $\mathcal{N}$ is the set of pairs $(0, u(\cdot))$ meeting the condition (5.8). Let Q be the quadratic form (5.5). Evidently $\operatorname{codim}\mathcal{N} = 1$.

For index of the restriction $Q_{\mathcal{N}}$ we have formula (5.10). Let us compute the other addends in the right-hand side of (5.11).

To this end let us characterize the space $\mathcal{N}_Q^{\perp}$. Let us define firstly the constraints of (5.6) by the system of equations:

$$\forall \nu \in \Pi: \ \sigma(\nu, \underline{f})\delta\theta + \int_0^T \sigma(\nu, \Upsilon_\tau)u(\tau)d\tau = 0, \quad (5.12)$$

and (5.8) by:

$$\forall \nu \in \Pi: \ \int_0^T \sigma(\nu, \Upsilon_\tau)u(\tau)d\tau = 0. \quad (5.13)$$

Then let us introduce a symmetric bilinear form $\mathcal{B}$ on the space (5.6) which corresponds to the quadratic form (5.5). This is:

$$\mathcal{B}(\delta\theta_1, u_1(\cdot); \delta\theta_2, u_2(\cdot)) = \int_0^T \gamma_\tau(u_1(\tau), u_2(\tau))d\tau +$$
$$+ \int_0^T \sigma(\int_0^\tau \Upsilon_\xi u_1(\xi)d\xi, \Upsilon_\tau u_2(\tau))d\tau + \int_0^T \sigma(-\underline{f}\delta\theta_2, \int_0^T \Upsilon_\tau u_1(\tau)d\tau). \quad (5.14)$$

Assuming that $\delta\theta_2 = 0$ and $u_2(\cdot)$ satisfies (5.13) we obtain for any pair $(\delta\theta, u(\cdot)) \in \mathcal{N}_Q^{\perp}$ an equation: for some $\nu \in \Pi$

$$\gamma_\tau u(\tau) + \sigma(\int_0^\tau \Upsilon_\xi u_1(\xi)d\xi, \Upsilon_\tau) = -\sigma(\nu, \Upsilon_\tau),$$

or

$$u(\tau) = \bar{\gamma}_\tau^{-1}\sigma(\Upsilon_\tau, \nu + \int_0^\tau \Upsilon_\xi u_1(\xi)d\xi). \quad (5.15)$$

Putting

$$y(\tau) = \nu + \int_0^\tau \Upsilon_\xi u_1(\xi)d\xi, \quad (5.16)$$

and differentiating it w.r.t. τ, we obtain

$$\dot{y}(\tau) = \Upsilon_\tau \bar{\gamma}_\tau^{-1}\sigma(\Upsilon_\tau, y(\tau)),$$

i.e. $y(\tau)$ satisfies the Jacobi equation (5.9).

Substituting (5.15) in (5.5) we derive, that

$$Q|_{\mathcal{N}_Q^\perp} = \sigma(\int_0^T \Upsilon_\tau u(\tau)d\tau, \underline{f}\delta\theta + \nu) = \sigma(y(T) - \nu, \underline{f}\delta\theta + \nu).$$

Since in virtue of (5.6) $\underline{f}\delta\theta + \nu + \int_0^T \Upsilon_\tau u(\tau)d\tau = \underline{f}\delta\theta + y(T) \in \Pi$ and hence is skew orthogonal to $\nu \in \Pi$, then

$$Q|_{\mathcal{N}_Q^\perp} = \sigma(y(T), \underline{f}\delta\theta). \tag{5.17}$$

Let us consider the starting at Π Lacobi curve $\tau \to \Lambda_\tau$ and take the Lagrangian plane $\Lambda_T^f = \Lambda_T \cap \underline{f}^\flat + \text{span}\, \underline{f}$. Define for the triple $(\Lambda_T, \Pi, \Lambda_T^f)$ of Lagrangian planes the quadratic form β on $((\Lambda_T + \Lambda_T^f) \cap \Pi)/(\Pi \cap \Lambda_T \cap \Lambda_T^f)$ according to Definition 2.1. Evidently $\text{ind} Q|_{\mathcal{N}_Q^\perp} = \text{ind}\beta$ and according to the same definition the modified Maslov index of the triple of Lagrangian planes equals $\text{ind}\ \beta + \frac{1}{2}\dim\ker\ \beta$. Calculating

$$\dim\ker\ \beta = \dim(\Pi \cap \Lambda_T) + \dim(\Lambda_T^f \cap \Pi) - 2\dim(\Pi \cap \Lambda_T \cap \Lambda_T^f)$$

we come to the formula

$$\begin{aligned} \text{ind}\, Q|_{\mathcal{N}_Q^\perp} &= \text{ind}_\Pi(\Lambda_T, \Lambda_T^f) \\ &-\frac{1}{2}(\dim(\Lambda_T \cap \Pi) + \dim(\Lambda_T^f \cap \Pi)) + \dim(\Lambda_T \cap \Lambda_T^f \cap \Pi). \end{aligned} \tag{5.18}$$

We now proceed with the other addends of (5.11). Note, that the inclusion $(\delta\theta, u(\cdot)) \in (\mathcal{N} \cap \mathcal{N}_Q^\perp)$ implies the equation (5.15) and $\delta\theta = 0$. This means that the elements $(\delta\theta, u(\cdot)) \in (\mathcal{N} \cap \mathcal{N}_Q^\perp)$ correspond to the solutions (5.16) of the Jacobi equation (5.9), which start and finish in Π. These solutions form a space of dimension $\dim(\Pi \cap \Lambda_T)$. Note that the mapping $y(\tau) \longrightarrow u(\cdot)$ may have a nontrivial kernel, consisting of *constant* solutions of Jacobi equation. The dimension of this kernel is then equal to $\dim \cap_{\tau\in[0,T]}\Lambda_\tau$. Hence

$$\dim(\mathcal{N} \cap \mathcal{N}_Q^\perp) = \dim(\Pi \cap \Lambda_T) - \dim \cap_{\tau\in[0,T]}\Lambda_\tau. \tag{5.19}$$

In order to compute $\dim(\mathcal{N} \cap \ker Q)$ we should suppose (5.14) to vanish for all pairs $(\delta\theta_2, u_2(\cdot))$, meeting the condition (5.12). Then we obtain, that $u_1(\cdot)$ must satisfy the equation (5.15) and

$$\sigma(\nu + \int_0^T \Upsilon_\tau u_1(\tau)d\tau, \underline{f}) = \sigma(y(T), \underline{f}) = 0.$$

Together with the condition $y(T) = \nu + \int_0^T \Upsilon_\tau u_1(\tau)d\tau \in \Pi$ this implies $y(T) \in \Pi \cap \underline{f}^\flat$. Such solutions form the space of dimension $\dim(\Lambda_T \cap \Lambda_T^f \cap \Pi)$, and

$$\dim(\mathcal{N} \cap \ker Q) = \dim(\Lambda_T \cap \Lambda_T^f \cap \Pi) - \dim(\cap_{\tau \in [0,T]} \Lambda_\tau \cap \underline{f}^\flat).$$

Since $\cap_{\tau\in[0,T]}\Lambda_\tau \cap \underline{f}^\flat$ consists of $\nu \in \Pi$ that $\sigma(\nu, \Upsilon_\tau) = 0$ and $\sigma(\nu, \underline{f}) = 0$ for all $\tau \in [0, T]$, we have

$$\begin{aligned}\cap_{\tau\in[0,T]}\Lambda_\tau \cap \underline{f}^\flat \subseteq \Pi^\flat \cap \underline{f}^\flat \cap \cap_{\tau\in[0,T]}(\Upsilon_\tau)^\flat = \\ (\Pi + \mathrm{span}\{\underline{f}\} + \mathrm{span}\{\Upsilon_\tau \mid \tau \in [0,T]\})^\flat = \Sigma^\flat = \{0\},\end{aligned}$$

and we obtain: $\dim(\mathcal{N} \cap \ker Q) = \dim(\Lambda_T \cap \Lambda_T^f \cap \Pi)$.

Summarizing all the computations we conclude that the difference between the indices of the second variation (4.18)-(4.19) and of the reduced second variation (5.1)-(5.2) equals

$$\mathrm{ind}_\Pi(\Lambda_T, \Lambda_T^f) + \frac{1}{2}(\dim(\Lambda_T \cap \Pi) - \dim(\Lambda_T^f \cap \Pi)) - \dim \cap_{\tau\in[0,T]}\Lambda_\tau.$$

This last formula together with Proposition 5.1 gives for the index of an abnormal extremal the following expression

$$\begin{aligned}\sum_{i=0}^{m} \mathrm{ind}_\Pi(\Lambda_{\zeta_i}, \Lambda_{\zeta_{i+1}}) + \mathrm{ind}_\Pi(\Lambda_T, \Lambda_T^f) + \\ + \frac{1}{2}(\dim(\Lambda_T \cap \Pi) - \dim(\Lambda_T^f \cap \Pi)) - (n-1). \qquad (5.20)\end{aligned}$$

Note that

$$\frac{1}{2}\dim(\Lambda_T \cap \Pi) = \frac{n-1}{2} - \mathrm{ind}_\Pi(\Lambda_T, \Pi),$$

$$\frac{1}{2}\dim(\Lambda_T^f \cap \Pi) = \frac{n-1}{2} - \mathrm{ind}_\Pi(\Lambda_T^f, \Pi).$$

Since in Proposition 5.1 $\Lambda_{\zeta_m} = \Lambda_T, \Lambda_{\zeta_{m+1}} = \Pi$ then the last addend of the sum $\sum_{i=0}^m$ in (5.20) is $\mathrm{ind}_\Pi(\Lambda_T, \Pi)$ and the formula (5.20) can be transformed into

$$\sum_{i=0}^{m-1} \mathrm{ind}_\Pi(\Lambda_{\zeta_i}, \Lambda_{\zeta_{i+1}}) + \mathrm{ind}_\Pi(\Lambda_T, \Lambda_T^f) + \mathrm{ind}_\Pi(\Lambda_T^f, \Pi) - (n-1). \qquad (5.21)$$

According to Definition 2.3 this expression corresponds to the Maslov index of some curve which we will call *Jacobi curve corresponding to an abnormal extremal* $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ of the Lagrange problem (4.1)-(4.3).

Definition 5.2 *The Jacobi curve corresponding to the abnormal extremal* $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ *is the curve* $\tau \rightarrow \Lambda^a_\tau$ $(\tau \in [0,T])$ *in the Lagrangian Grassmanian such that* $\Lambda^a_0 = \Pi$, Λ^a_τ *coincides for* $\tau \in [0,T)$ *with the starting at* Π *solution* Λ_τ *of the Jacobi equation (5.9) and jumps at* $\tau = T - 0$ *to* $\Lambda^f_T = (\Lambda_T)^{R\underline{f}} = \Lambda_T \cap \underline{f}^\flat + \mathrm{span}\{\underline{f}\}$. □

Now we are able to state

Theorem 5.3 (Index Theorem for Abnormal Extremals)
Let $\tau \rightarrow \Lambda^a_\tau$ *be the Jacobi curve in the Lagrangian Grassmanian* $\mathcal{L}(\Sigma)$, *which corresponds to the corank* 1 *abnormal extremal* $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$. *Then for any subdivision* $\eta_{s+1} = 0 = \eta_0 < \eta_1 < \cdots < \eta_s = T$ *of* $[0,T]$ *such that all arcs* $\Lambda|_{[\eta_i, \eta_{i+1}]}$, $(i = 0, \ldots s-1)$ *are simple, the Morse index of the abnormal extremal equals*

$$\sum_{i=0}^{s} \mathrm{ind}_\Pi(\Lambda^a_{\eta_i}, \Lambda^a_{\eta_{i+1}}) - (n-1). \ \square \tag{5.22}$$

In order to formulate the Nullity Theorem for an abnormal extremal we have to investigate the kernel of the bilinear form (5.14) defined on the subspace (5.6). If a pair $(\delta\theta_1, u_1(\cdot))$ belongs to $\ker \mathcal{B}$, then $u_1(\cdot)$ must satisfy the equation (5.15) and the condition $\nu + \int_0^T \Upsilon_\tau u(\tau) d\tau = y(T)\flat\underline{f}$. Taking into account the condition $y(T) + \underline{f}\delta\theta \in \Pi$ we conclude that the dimension of the space of these solutions equals $\dim(\Lambda^f_T \cap \Pi)$ and the dimension of the space of those (constant) solutions of (5.9), which correspond to zero elements of $\ker Q$, is $\dim(\cap_{\tau \in [0,T]} \Lambda_\tau \cap \underline{f}^\flat) = 0$.

Hence

$$\dim \ker Q = \dim(\Lambda^f_T \cap \Pi) = \dim(\Lambda^a_T \cap \Pi).$$

Thus we obtain

Theorem 5.4 (Nullity Theorem for Abnormal Extremals) *Let the Jacobi curve* $\tau \longrightarrow \Lambda^a_\tau$ *in the Lagrangian Grassmanian* $\mathcal{L}(\Sigma)$, *correspond to the abnormal extremal* $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$. *Then the Nullity of the abnormal extremal, i.e. the dimension of the kernel of the second variation (5.5)-(5.6), equals* $\dim(\Lambda^a_T \cap \Pi)$. □

What follows is a corollary of theorems 5.3 and 5.4.

Corollary 5.5 (Local Rigidity of Abnormal Extremals)

Let the Strong Legendre Condition hold along an abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ *whose restrictions* $(\hat{u}(\cdot)|_{[0,t]}, \hat{q}(\cdot)|_{[0,t]}, 0, \hat{\psi}(\cdot)|_{[0,t]}, t)$ *on* $[0,t]$ *for* $t \in (0,T]$ *all have corank* 1. *Then for any small enough* $t > 0$ *the* *restriction* $(\hat{u}(\cdot)|_{[0,\bar{t}]},\bar{t})$ *is a weak minimizer for the Lagrange problem (4.1)-(4.2) with endpoint condition* $q(\bar{t})=\hat{q}(\bar{t})$. *For some (depending on* $\bar{t}$*)* $\epsilon>0$ *no one other admissible control from the* $\epsilon-$*neighborhood of* $\hat{u}(\cdot)$ *in* L_∞ *is able to steer the system (4.2) from* q^0 *to* $\hat{q}(\bar{t})$ *in any time* $t'\in[\bar{t}-\epsilon,\bar{t}+\epsilon]$. □

Proof: It is enough to prove that both the index and nullity along the restriction vanish. In fact then the Strong Legendre Condition implies positive definiteness of the second variation along the restriction and we can apply Theorem 4.3.

To compute the index of the restriction $(\hat{u}(\cdot)|_{[0,t]},\hat{q}(\cdot)|_{[0,t]},0,\hat{\psi}(\cdot)|_{[0,t]},t)$ let us consider the corresponding Jacobi curve of Definition 5.2. Since $\underline{f}\notin\Pi$ then $\underline{f}\notin\Lambda^a_\tau$ for any small enough $\tau>0$ and therefore starting at Π^f curve $\tau\to\Lambda^f_\tau$ where $\Lambda^f_\tau=\Lambda^a_\tau\cap\underline{f}^\flat+\underline{f}$ is continuous on a sufficiently small interval $[0,t]$. Then according to Lemma 2.2 there exist $t>0$ and a Lagrangian plane Δ such that for any $\tau\in[0,t]$ Λ^a_τ can be connected with Λ^f_τ by a simple nondecreasing curve $\Lambda_\tau(s),\ 0\leq s\leq 1$ such that $\Lambda_\tau(s)\cap\Delta=0,\ \forall s\in[0,1]$. Then the concatenation of the curve $\Lambda^a|_{[0,t]}$ with the corresponding curve $\Lambda_t(s)$ is also simple and evidently nondecreasing.

According to Proposition 5.1 and Theorem 5.3 the index of the (having corank 1) restriction $(\hat{u}(\cdot)|_{[0,t]},\hat{q}(\cdot)|_{[0,t]},0,\hat{\psi}(\cdot)|_{[0,t]},t)$ equals

$$\mathrm{ind}_\Pi(\Pi,\Lambda_t)+\mathrm{ind}_\Pi(\Lambda_t,\Lambda^f_t)+\mathrm{ind}_\Pi(\Lambda^f_t,\Pi)-(n-1).$$

According to Lemma 2.1 $\mathrm{ind}_\Pi(\Pi,\Lambda_t)+\mathrm{ind}_\Pi(\Lambda_t,\Lambda^f_t)=\mathrm{ind}_\Pi(\Pi,\Lambda^f_t)$ for all small enough $t>0$ and we obtain for the Morse index the expression:

$$\begin{aligned}&\mathrm{ind}_\Pi(\Pi,\Lambda^f_t)+\mathrm{ind}_\Pi(\Lambda^f_t,\Pi)-(n-1)=\\&\quad=2\frac{1}{2}(n-1-\dim(\Lambda^f_t\cap\Pi))-(n-1)=-\dim(\Lambda^a_t\cap\Pi)\leq 0.\end{aligned}$$

Being nonnegative this Morse index must vanish, i.e. $\dim(\Lambda^f_t\cap\Pi)=0$. According to Theorem 5.4 the last dimension coincides with the nullity of the restriction $(\hat{u}(\cdot)|_{[0,t]},\hat{q}(\cdot)|_{[0,t]},0,\hat{\psi}(\cdot)|_{[0,t]},t)$ provided the restriction has corank 1. □

Let us consider a corank 1 abnormal extremal $(\hat{u}(\cdot),\hat{q}(\cdot),0,\hat{\psi}(\cdot),T)$ meeting the Strong Legendre Condition. We would like to characterize the time evolution of the attainable sets $\mathcal{A}^\epsilon_t=F(t,\mathcal{U}^\epsilon)$ (here $\mathcal{U}^\epsilon$ is a small $\epsilon-$ neighborhood of $\hat{u}(\cdot)$ in L_∞).

Considering the restrictions $(\hat{u}(\cdot)|_{[0,t]},\hat{q}(\cdot)|_{[0,t]},0,\hat{\psi}(\cdot)|_{[0,t]},t)$ of the abnormal extremal to subintervals $[0,t]\subset[0,T]$ let us put

$$i(t)=\mathrm{ind}F''|_{(t,\hat{u}|_{[0,t]})}[\psi_T],\ i_r(t)=\mathrm{ind}F''_r|_{(t,\hat{u}|_{[0,t]})}[\psi_T]$$

for the indices of the second variation and the reduced second variation along the restricted extremals. It is known, that $i(t)$ and $i_r(t)$ are nondecreasing functions of t, $i_r(t) \leq i(t) \leq i_r(t) + 1$, and according to the Corollary 5.5 both $i(t)$ and $i_r(t)$ vanish for small $t > 0$.

Let us put

$$\Theta_r^* = \inf\{t | i_r(t) > 0\}, \ \Theta^* = \inf\{t | i(t) > 0\}.$$

Obviously $\Theta^* \leq \Theta_r^*$.

The points Θ^* and Θ_r^*, where the indices $i(t)$ and $i_r(t)$ jump, will be called correspondingly first conjugate point and first reduced conjugate point for the abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot), T)$ (compare with [9], where the difference between these two points was pointed out).

If $\Theta^* < \Theta_r^*$, then the time evolution of the attainable sets $\mathcal{A}_t^\epsilon$ looks like Fig. 1 (compare with [9]).

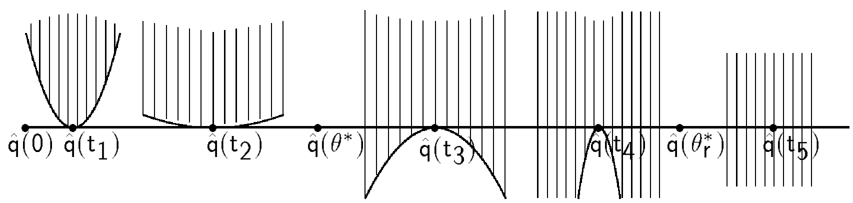

Fig. 1

Let us assume now $\Theta^* = \Theta_r^*$. It happens, for example, when $\hat{f}_T(q^1) \notin \text{span}\{Y_\tau^1(q^1) | \tau \in [0, T]\}$ and the reduced second variation (5.1)-(5.2) coincides with the second variation (4.18)-(4.19). As an illustrative example one may consider the control system:

$$\dot{x} = 1 - z, \ \dot{y} = u, \ \dot{z} = u^2 - y^2, \ x(0) = 0, \ y(0) = 0, z(0) = 0$$

driven by the control $\hat{u}(t) \equiv 0$. Constructing 'abnormal' Hamiltonian: $H = \psi_x(1 - z) + \psi_y u + \psi_z(u^2 - y^2)$ and introducing the adjoint system $\dot{\psi}_x = 0, \ \dot{\psi}_y = 2y\psi_z, \ \dot{\psi}_z = \psi_x$ we establish easily that

$$\hat{u}(t) = 0, \ \hat{x}(t) = t, \ \hat{y}(t) = 0, \hat{z}(t) = 0, \ \hat{\psi}_x(t) = 0, \ \hat{\psi}_y(t) = 0, \ \hat{\psi}_z(t) = 1$$

is an abnormal extremal for this system on any interval $[0, T]$. To find the conjugate point $\Theta^* = \Theta_r^*$ we write the Jacobi equation (5.9) for the reduced second variation. It is one-degree time-dependent Hamiltonian system, which in canonical coordinates (q, p) has form: $\dot{q} = \frac{1}{2}(p - 2tq), \ \dot{p} = t(p - 2tq)$. The corresponding initial condition is $q(0) = 0$. One derives from the Hamiltonian system the equation $\ddot{q} = -q$. Starting at 0 solutions of this equation are $q(t) = A \sin t$. The conjugate point $\Theta^* = \Theta_r^* = \pi$ is the first point where these solutions vanish.

We draw the projections of the portraits of the attainable sets $\mathcal{A}_t^\epsilon$ onto xz plane. The evolution of $\mathcal{A}_t^\epsilon$ with the growth of t differs from the one shown in Fig. 1 and looks like Fig. 2.

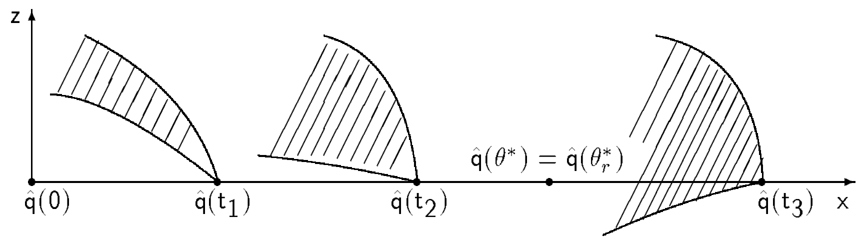

Fig. 2

References

[1] A.A. Agrachev. Quadratic mappings in geometric control theory, in *Itogi Nauki i Tekhniki; Problemy Geometrii, VINITI*, Acad. Nauk SSSR **20** (1988), 11-205. English transl. in *J. Soviet Math.* **51** (1990), 2667-2734.

[2] A.A. Agrachev. Topology of quadratic mappings and hessians of smooth mappings, in *Itogi Nauki i Tekhniki; Algebra, Topologia, Geometria; VINITI*, Acad. Nauk SSSR **26** (1988), 85-124.

[3] A.A. Agrachev. Necessary 2-nd order optimality condition for a general nonlinear problem, *Matem. Sbornik* **102** (1977), 551-568. English transl. in *Math. USSR Sbornik* **31** (1977), 493-506.

[4] A.A. Agrachev and R.V. Gamkrelidze. Exponential representation of flows and chronological calculus, *Matem. Sbornik* **107** (1978), 467-532. English transl. in *Math. USSR Sbornik* **35** (1979), 727-785.

[5] A.A. Agrachev, R.V. Gamkrelidze and A.V. Sarychev. Local invariants of smooth control systems, *Acta Applicandae Mathematicae* **14** (1989), 191-237.

[6] V.I. Arnol'd. *Mathematical Methods of Classical Mechanics.* Berlin: Springer-Verlag, 1978.

[7] V.I. Arnol'd, A.N. Varchenko and S.M. Gusein-Zade. *Singularities of Differentiable Mappings* **1**. Boston: Birkhäuser, 1985.

[8] A.V. Arutjunov. Necessary extremality conditions for an abnormal problem with equality-type constraints, *Uspekhi Matem.Nauk* **45(5)** (1990), 181-182. English transl. in *Russian Mathematical Surveys* **45(5)** (1990).

[9] B. Bonnard and I. Kupka. Theorie de singularites de l'application entree/sortie et optimalite des trajectoires singulieres dans le probleme du temps minimal, preprint 1990.

[10] V. Guillemin and S. Sternberg. *Geometric Asymptotics.* Providence, RI: Amer. Math. Soc., 1977.

[11] M.R. Hestenes. *Optimization Theory - the Finite Dimensional Case.* New York: Wiley, 1975.

[12] M.R. Hestenes. On quadratic control problems in (D.L. Russel, ed.) *Calculus of Variations and Control Theory.* New York: Academic Press, 1976, 289-304.

[13] A.D. Ioffe and V.M. Tichomirov. *Theorie der Extremalaufgaben.* Berlin: VEB, 1979.

[14] A.J. Krener. The high-order maximum principle and its applications to singular extremals, *SIAM J. on Control and Optimiz.* **15** (1977), 256-293.

[15] G. Lion and M. Vergne. *The Weyl Representation, Maslov Index and Theta-series.* Boston: Birkhäuser, 1980.

[16] R. Montgomery. *Geodesics, which do not satisfy geodesic equations,* preprint 1991.

[17] M. Morse. *The Calculus of Variations in the Large.* New York: Amer.Math.Soc., 1934.

[18] A.V. Sarychev. The index of the second variation of the control system, *Matemat. Sbornik* **113** (1980), 464-486. English transl. in *Math.USSR Sbornik* **41** (1982), 383-401.

[19] L.C. Young. *Lectures on the Calculus of Variations and Optimal Control Theory.* New York: Chelsea, 1980.

Steklov Mathematical Inst., Russian Acad. of Sciences, ul. Vavilova 42, 117966, Moscow, Russia

A.A. AGRACHEV AND A.V. SARYCHEV

Departamento de Matemática, Universidade de Aveiro, 3800, Aveiro, Portugal. On leave from Inst. for Control Sciences, Russian Acad. of Sciences, Profsojuznaja ul. 65, 117342, Moscow, Russia

Communicated by Clyde F. Martin