FEEDBACK–INVARIANT OPTIMAL CONTROL THEORY AND DIFFERENTIAL GEOMETRY — I. REGULAR EXTREMALS

A. A. Agrachev, R. V. Gamkrelidze

Abstract. Feedback–invariant approach to smooth optimal control problems is considered. A Hamiltonian method of investigating regular extremals is developed, analogous to the differential–geometric method of investigation Riemannian geodesics in terms of the Levi–Civita connection and the curvature tensor.

§0. Introduction

1. Outline of the content. This is the first in a series of forthcoming papers, devoted to the unification of the Theory of Smooth Optimal Control Problems and that part of Differential Geometry which is dealing with geodesics of different kinds. The obtained results, we believe, not merely suggest a dictionary for translating the known results from one language into another, but they really extend the scope of applicability of both theories. The key notions brought into interplay are "Hamiltonian system" in optimal control and the "curvature tensor" in differential geometry.

Since the discovery of the Pontryagin maximum principle, cf. [11], finding extremals in problems of optimal control is reduced to solving Hamiltonian systems of differential equations. Even in the classical case of Riemannian geometry, the maximum principle approach to finding geodesics leads to the final result much simpler and shorter than the traditional method of using the Levi–Civita connection. If we consider more general geometric variational problems, dealing not only with regular extremals (geodesics), but with singular extremals as well, then we should admit that the maximum principle approach has no serious alternative.

Turning now from geodesics to the curvature tensor, which describes quite deeply not only local but also global behavior of geodesics without even solving any differential equations, we see that it is obtained in a standard way from the Levi–Civita connection, whereas a Hamiltonian approach to the curvature tensor or its analogue was never considered. The main content of this paper is devoted to such an approach.

We shall give now a brief overview of the content by sections. In §1 the notion of the $\mathcal{L}$–derivative is introduced. The intuitive meaning of this notion, which plays in the sequel an important role, could be described as follows.

1991 Mathematics Subject Classification. 49K15, 53B05(B15), 53C22.

Key words. Hamiltonian, extremal, geodesic, connection, curvature.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Let $f : U \longrightarrow M$ be a smooth mapping between two finite-dimensional manifolds, with the differential $f'_u : T_uU \longrightarrow T_{f(u)}M$ at $u \in U$. The point u is critical for f if and only if the image of f'_u is annihilated by some nonzero covector λ, the Lagrange multiplier,

$$\lambda f'_u = 0,\ \lambda \in T^*_{f(u)}M,\ \lambda \neq 0.$$

In this equation, the argument λ belongs to the symplectic manifold T^*M, the argument u — to the manifold U. Linearization of the equation at the point (u, λ) gives us a linear system of equations in variables $\delta u \in T_uU$, $\delta\lambda \in T_\lambda(T^*M)$. Let $\mathcal{L}_{(u,\lambda)}(f)$ be the set of all $\delta\lambda$ which satisfy the linear system (with some δu). The linear subspace $\mathcal{L}_{(u,\lambda)}(f) \subset T_\lambda(T^*M)$ is called the $\mathcal{L}$–derivative of f at (u, λ). It turns out that $\mathcal{L}_{(u,\lambda)}(f)$ is always a Lagrangian subspace of the symplectic space $T_\lambda(T^*M)$, in particular, $dim\,\mathcal{L}_{(u,\lambda)}(f) = dim\,M$. Thus the dimension of the $\mathcal{L}$–derivative is independent on the character of the critical point; for a constant mapping we have $\mathcal{L}_{(u,\lambda)}(f) = T_\lambda(T^*_{f(u)}M)$. The optimal control situation is more general, with an infinite–dimensional U. In §1 the $\mathcal{L}$–derivative for the infinite–dimensional case is considered, and the important chain rule for the $\mathcal{L}$–derivative is formulated.

In §2 we give a feedback–invariant definition of a smooth control system which includes as special cases many basic differential–geometric structures. The space of admissible trajectories is introduced and the boundary–value mapping is defined, which sends the trajectory into its boundary points. Critical points of the boundary–value mapping are the extremal trajectories, geodesics in the geometric terminology. At the end of the section a Hamiltonian characterization of extremal trajectories is given in form of a feedback–invariant analogue of the maximum–principle.

In §3 the $\mathcal{L}$–derivative of the boundary–value mapping and of its particular case, of the endpoint mapping, is computed. We also introduce regular extremals which are trajectories of a fixed Hamiltonian system, defined in a region of T^*M. For regular extremals, the $\mathcal{L}$–derivative of the boundary–value mapping is computed particularly simple. Let $\tau \mapsto \lambda_\tau \in T^*M,\ 0 \le \tau \le t$, be a regular extremal, $\lambda_\tau = P_\tau(\lambda_0)$, where $\tau \mapsto P_\tau$ is the Hamiltonian flow in T^*M such that all of its trajectories are regular extremals. Then the $\mathcal{L}$–derivative is the graph of the linear symplectic mapping $P_{t*} : T_{\lambda_0}(T^*M) \longrightarrow T_{\lambda_t}(T^*M)$.

In §4 Jacobi curves are introduced and investigated. Jacobi curves are curves in the Lagrangian Grassmannian corresponding to given extremals of a very general nature, certainly including all regular extremals, and are constructed, roughly, in the following way. An arbitrary segment of an extremal is again an extremal. Hence, varying the initial point of the extremal with the fixed endpoint (or vice versa), we obtain $\mathcal{L}$–derivatives of the endpoint mapping, which are Lagrangian subspaces in a fixed symplectic space depending on a time–variable, thus obtaining the Jacobi curve.

We also develop here the differential geometry of regular curves in a Lagrangian Grassmannian. Nonregular Jacobi curves, occurring in problems with nonholonomic constraints, will be considered in future publications. The most important infinitesimal invariants of a regular curve in the Lagrangian Grassmannian are the "derivative curve" and the "curvature tensor". The curves of constant scalar curvature are characterized. Formulas are derived which relate the number of conjugate points, Maslov index and curvature.

The obtained results are applied in §5 to develop the differential geometry of Hamiltonian systems on T^*M, and of differential equations of second order on M. To a trajectory of the Hamiltonian system, passing through a point $\lambda \in T^*M$, germ of a curve in the Lagrangian Grassmannian of the symplectic space $T_\lambda(T^*M)$ is assigned. Infinitesimal invariants of these germs define a (in general nonlinear) canonical connection on T^*M associated to the Hamiltonian. The main result of this section consists in deriving identities, connecting the curvature of the canonical connection with the curvatures of the germs of curves in the Lagrangian Grassmannian. A similar theory is developed for differential equations of the second order for which we have to substitute T^*M by TM, and the Lagrangian Grassmannian by ordinary Grassmannian. It turns out that the canonical connection of the equation of the geodesic flow of a (pseudo)–Riemannian structure coincides with the Levi–Civita connection of this structure.

In §6, for two–dimensional systems, the curvature of the extremals of smooth control systems is expressed through standard "state–invariants", the iterated Lie brackets of vector fields.

2. Preliminaries. Here we introduce some formulas of "Chronological calculus" and certain notions and relations related to linear symplectic spaces used in the article, cf. [3,8], [1,7,9].

Assume M is a smooth, i.e. of class C^∞, manifold, and $C^\infty(M)$ is the algebra of smooth functions on M. We identify an arbitrary diffeomorphism $P : M \longrightarrow M$ with the corresponding automorphism of the algebra $C^\infty(M)$,

$$a(\cdot) \mapsto a \circ P(\cdot) = a(P(\cdot)),\ a \in C^\infty(M).$$

Under this identification, the action of P on a, i.e. the substitution of P into a, is denoted by Pa, and the value of P at $x \in M$ is denoted by xP, $xPa \overset{\text{def}}{=} a(P(x))$. As usual, smooth vector fields X on M are identified with the derivations of the algebra $C^\infty(M)$, hence they are $\mathbb{R}$–linear mappings of $C^\infty(M)$ satisfying the Leibniz rule, $X(a_1a_2) = a_1Xa_2 + a_1Xa_2$. The Lie bracket, $[X_1, X_2] = X_1 \circ X_2 - X_2 \circ X_1$, turn the $\mathbb{R}$–linear space of vector fields into the Lie algebra, $Vect\, M$. For a given X, the inner derivation of the Lie algebra $Vect\, M$ is defined,

$$ad\, X : Vect\, M \longrightarrow Vect\, M,\ (adX)Y = [X, Y].$$

Every diffeomorphism P defines an inner automorphism $Ad\, P : Vect\, M \longrightarrow Vect\, M$, $(Ad\, P)X = P \circ X \circ P^{-1}$. It is easily seen that the differential of the inverse diffeomorphism P^{-1}, denoted by P_*^{-1}, acts on vector fields according to the formula $(Ad\, P)X = P_*^{-1}X$.

We also consider nonstationary vector fields, i.e. measurable essentially bounded mappings, $t \mapsto X_t,\ t \in \mathbb{R},\ X_t \in Vect\, M$, and nonstationary flows, i.e. Lipschitz mappings $t \mapsto P_t,\ t \in \mathbb{R},\ P_t \in Diff\, M$. Every nonstationary vector field defines the corresponding differential equation $\dot{x} = X_t(x)$ on M with an arbitrary initial condition, $x(t_0) = x_0$. If a solution of the differential equation exists for arbitrary $x_0 \in M,\ t \in \mathbb{R}$, i.e. if the field X_t is complete, then it uniquely defines for $t \in \mathbb{R}$ an absolutely continuous solution of the operator differential equation

$$\frac{dP}{dt} = P_t \circ X_t,\ P_{t_0} = id, \tag{0.1}$$

which we call the *flow on* M, defined by the nonstationary complete vector field X_t, and denote

$$P_t = \overrightarrow{exp}\int_{t_0}^{t} X_\tau d\tau.$$

We also call this flow the *right chronological exponential of* X_τ. For stationary vector fields $X_\tau \equiv X$, the corresponding flows are denoted by $P_t = e^{tX}$. In the sequel, all vector fields are assumed to be complete. This will not restrict the generality of our considerations.

The chronological exponential admits an asymptotical expansion as a Volterra series,

$$\overrightarrow{exp}\int_{t_0}^{t} X_\tau d\tau \approx id + \int_{t_0}^{t} X_\tau d\tau + \ldots + \int_{t_0}^{t} d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \ldots \int \lim its_{t_0}^{\tau_{i-1}} (X_{\tau_i} \circ \ldots \circ X_{\tau_1}) d\tau_i + \ldots \quad .$$

For a stationary field, we obtain

$$e^{(t-t_0)X} \approx id + (t-t_0)X + \ldots + \frac{(t-t_0)^i}{i!} X^i + \ldots \quad .$$

In the sequel, we shall need the following important *variation formula*, which represents the chronological exponential of the sum of two vector fields as a product of two chronological exponentials,

$$\begin{gathered}\overrightarrow{exp}\int_{t_0}^{t} (X_\tau + Y_\tau) d\tau = \overrightarrow{exp}\int_{t_0}^{t} X_\tau d\tau \circ \overrightarrow{exp}\int_{t_0}^{t} Ad\left(\overrightarrow{exp}\int_{t}^{\tau} X_\theta d\theta\right) Y_\tau d\tau = \\ \overrightarrow{exp}\int_{t_0}^{t} Ad\left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right) Y_\tau d\tau \circ \overrightarrow{exp}\int_{t_0}^{t} X_\tau d\tau.\end{gathered} \tag{0.2}$$

Differentiating the expression

$$Ad\left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right) Y = \overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta \circ Y \circ \left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right)^{-1}$$

with respect to τ, we obtain the equality

$$\frac{d}{d\tau} Ad\left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right) Y = Ad\left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right) \circ ad\, X_\tau Y \quad \forall Y \in Vect\, M.$$

In other words, the expression $Ad\left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right)$ satisfies an equation, similar to (0.1), where the field X_τ is substituted by the operator $ad\, X_\tau$. This remark should justify the notation

$$Ad\left(\overrightarrow{exp}\int_{t_0}^{\tau} X_\theta d\theta\right) = \overrightarrow{exp}\int_{t_0}^{\tau} ad\, X_\theta d\theta, \quad Ad\left(e^{\tau X}\right) = e^{\tau\, ad\, X},$$

which is also supported by the validity of the asymptotic expansions

$$\overrightarrow{exp}\int_{t_0}^{t} ad\,X_\tau d\tau \approx id + \int_{t_0}^{t} ad\,X_\tau d\tau + \ldots + \int_{t_0}^{t} d\tau_1 \int_{t_0}^{t_1} d\tau_2 \ldots \int_{t_0}^{\tau_{i-1}} (ad\,X_{\tau_i} \circ \ldots \circ ad\,X_{\tau_1}) d\tau_i + \ldots$$

$$e^{\tau\, ad\,X} \approx id + \tau\, ad\,X + \ldots + \frac{\tau^i}{i!}(ad\,X)^i + \ldots \quad .$$

In this notation, the variation formula (0.2) takes the form

$$\overrightarrow{exp}\int_{t_0}^{t} (X_\tau + Y_\tau) d\tau = \overrightarrow{exp}\int_{t_0}^{t} X_\tau d\tau \circ \overrightarrow{exp}\int_{t_0}^{t} \left(\overrightarrow{exp}\int_{t}^{\tau} ad\,X_\theta d\theta \right) Y_\tau d\tau =$$
$$\overrightarrow{exp}\int_{t_0}^{t} \left(\overrightarrow{exp}\int_{t_0}^{\tau} ad\,X_\theta d\theta \right) Y_\tau d\tau \circ \overrightarrow{exp}\int_{t_0}^{t} X_\tau d\tau. \tag{0.3}$$

We shall also need the Hamiltonian version of the variation formula. Let $TM,\ T^*M$ be tangent and cotangent bundles of M, with canonical projections denoted by one letter, π. Let θ be the canonical 1–form on T^*M, $< \theta_\lambda, v > = < \lambda, \pi_* v > \ \ \forall \lambda \in T^*M,\ v \in T_\lambda(T^*M)$. The 2–form $\sigma = d\theta$ is the canonical symplectic structure on T^*M. Every smooth function h, defined on an open region of T^*M, defines a Hamiltonian vector field $\vec{h}$ on the region by the formula

$$\vec{h} \rfloor \sigma = -dh.$$

The corresponding differential equation $\dot{\lambda} = \vec{h}(\lambda)$ is the Hamiltonian system associated to the Hamiltonian H. The Poisson bracket of two Hmiltonians is given by the relation

$$\{h_1, h_2\} \stackrel{\text{def}}{=} \sigma(\vec{h}_1, \vec{h}_2) = \vec{h}_1 h_2.$$

If the functions h_i are linear on fibres of T^*M, $h_i(\lambda) = < \lambda, X_i >,\ \ X_i \in Vect\,M,\ \ i = 1, 2,$ then

$$\{h_1, h_2\}(\lambda) = < \lambda, [X_1, X_2] > .$$

A measurable essentially bounded family of Hamiltonians $h_t,\ t \in \mathbb{R}$, will be called a *non-stationary Hamiltonian.* The corresponding *Hamiltonian flow* $\overrightarrow{exp}\int_{t_0}^{t} \vec{h}_\tau d\tau$ preserves the symplectic structure and satisfies the relation

$$\frac{\partial}{\partial t}\overrightarrow{exp}\int_{t_0}^{t} \vec{h}_\tau d\tau g = \overrightarrow{exp}\int_{t_0}^{t} \vec{h}_\tau d\tau \{h_t, g\} \ \ \forall g.$$

Finally, the variation formula for Hamiltonian flows could be reduced to the following relations,

$$\overrightarrow{exp}\int_{t_0}^{t}(\overrightarrow{h}_\tau+\overrightarrow{g})d\tau = \overrightarrow{exp}\int_{t_0}^{t}\overrightarrow{h}_\tau d\tau\circ\overrightarrow{exp}\int_{t_0}^{t}\overrightarrow{(\overrightarrow{exp}\int_{t}^{\tau}\overrightarrow{h}_\theta d\theta)g_\tau} d\tau =$$

$$\overrightarrow{exp}\int_{t_0}^{t}\overrightarrow{(\overrightarrow{exp}\int_{t_0}^{\tau}\overrightarrow{h}_\theta d\theta)g_\tau} d\tau\circ\overrightarrow{exp}\int_{t_0}^{t}\overrightarrow{h}_\tau d\tau$$

We introduce now some notions and formulate certain facts of linear symplectic geometry used in the sequel. For details cf. [1,7,9].

Let Σ be a $2n$–dimensional symplectic space with the skew-symmetric form $\sigma(\cdot,\cdot)$, for example, the cotangent space $T_\lambda(T^*M)$. For every subspace $S\subset\Sigma$ put $S^\angle=\{\zeta\in\Sigma\big|\ \sigma(S,\zeta)=0\}$, hence $dim\,S+dim\,S^\angle=2n$. The subspace S is called isotropic if $S^\angle\subset S$. An n–dimensional subspace $\Lambda\subset\Sigma$ is called Lagrangian if $\Lambda^\angle=\Lambda$. The set of all Lagrangian subspaces is organized into a smooth $\frac{n(n+1)}{2}$–dimensional manifold, the Lagrangian Grassmannian, $L(\Sigma)=\{\Lambda\subset\Sigma\big|\ \Lambda^\angle=\Lambda\}$.

The tangent space $T_\Lambda L(\Sigma)$, $\Lambda\in L(\Sigma)$, is identified in a natural way with the space of quadratic forms defined on the n–dimensional space Λ. Indeed, let $t\mapsto\Lambda_t$ be a germ of a smooth curve in $L(\Sigma)$. We correspond to the tangent vector $\frac{d}{dt}\Lambda$ the quadratic form $\dot\Lambda_0:\lambda_0\mapsto\sigma(\lambda_0,\frac{d}{dt}\lambda_0)$, where $t\mapsto\lambda_t$ is a germ of a smooth curve in Σ, $\lambda_t\in\Lambda_t$. It is easy to show that $\dot\Lambda_0(\lambda_0)$ is correctly defined, i.e. it does depend on $\frac{d}{dt}\Lambda_0$, λ_0, but is independent on the choice of the germs $\Lambda.$, $\lambda.$.

The symplectic group $Sp(\Sigma)$ is the group of linear transformations of Σ preserving σ, hence transforming Lagrangian subspaces into Lagrangian subspaces. $Sp(\Sigma)$ acts on $L(\Sigma)$ transitively, thus $L(\Sigma)$ is a homogeneous space for the group $Sp(\Sigma)$.

Let h be a quadratic form (quadratic Hamiltonian) on Σ, then $e^{t\overrightarrow{h}}\in Sp(\Sigma)$. Put $\Lambda_t=e^{t\overrightarrow{h}}(\Lambda_0)$, then $\dot\Lambda_0=2h\Big|_{\lambda_0}$. A smooth curve in $L(\Sigma)$, $t\mapsto\Lambda_t$, is called *monotonically nondecreasing (nonincreasing)* if $\dot\Lambda_t\geq0$ ($\dot\Lambda_t\leq0$). The subset

$$\mathcal{M}_{\Lambda_0}=\Big\{\Lambda\in L(\Sigma)\Big|\ \Lambda\bigcap\Lambda_0\neq0\Big\}\subset\Sigma$$

is called the *train* of the Lagrangian subspace Λ_0. $\mathcal{M}_{\Lambda_0}$ is an algebraic hypersurface in $L(\Sigma)$, smooth beyond some set of codimension 3 in Σ, hence $\mathcal{M}_{\Lambda_0}$ is a pseudomanifold. Moreover, the hypersurface $\mathcal{M}_{\Lambda_0}$ carries a natural coorientation, defined in such a way that the monotonically increasing curves intersect $\mathcal{M}_{\Lambda_0}$ in the positive direction, and monotonically decreasing curves — in the negative direction.

Thus for every continuous curve $t\mapsto\Lambda(t)$, $t\in[t_0,t_1]$, such that $\Lambda(t_i)\notin\mathcal{M}_{\Lambda_0}$, $i=0,1$, the intersection number, $\Lambda(\cdot)\cdot\mathcal{M}_{\Lambda_0}$, is defined, and $\Lambda(\cdot)\cdot\mathcal{M}_{\Lambda_0}\geq0\,(\leq0)$, if the curve $\Lambda(\cdot)$ is nondecreasing (nonincreasing).

The curve is called *simple* if there exists $\Delta \in L(\Sigma)$ such that $\Lambda(t) \bigcap \Delta = 0 \ \forall t \in [t_0, t_1]$. If the curve $\Lambda(\cdot)$ is simple, then $\Lambda(\cdot) \cdot \mathcal{M}_{\Lambda_0} \leq n$. Finally, if $\Lambda(\cdot)$ is closed, $\Lambda(t_0) = \Lambda(t_1)$, then the intersection number $\Lambda(\cdot) \cdot \mathcal{M}_{\Lambda_0}$ does not depend on Λ_0 and is denoted $Ind\,\Lambda(\cdot)$. This is the Maslov index of the closed curve.

§1. $\mathcal{L}$–DERIVATIVES OF SMOOTH MAPPINGS

We start with some definitions and constructions relevant to critical points of smooth mappings. The exposition is carried out for infinite dimensional case, sufficiently general for handling variational and control problems discussed further.

The differential of a scalar–valued function on a Banach space (evaluated at an arbitrary point) is an element of the dual space. In the finite–dimensional situation we can make no difference between the initial space and its dual, but in the infinite–dimensional case the dual might be less comprehensible than the initial space. A standard example — the space $L_\infty[0,1]$ of admissible controls in optimal control problems, which is very natural and simplest possible to be considered in most situations, but its dual is pretty involved. Due to the restrictive nature of the functionals involved in smooth control problems, the final results, if appropriately formulated, do not use the dual space at all, though some cumbersome analytical efforts are needed for eliminating the dual space in final formulations. Meanwhile, a natural modification of some basic definitions makes it possible to avoid all artificial complications connected with this phenomenon. The simple trick consists in considering the initial space as a dual to some "acceptable" Banach space, in our example, considering from the beginning the space of controls $L_\infty[0,1]$ as the dual to $L_1[0,1]$, and appropriately defining (stiffening) the notion of the differential of a smooth mapping on such a space. Formally, we proceed in the following way.

Let B be a Banach space, B^* — its dual. We shall always suppose the natural (isometric) inclusion $B \subset B^{**}$.

A differentiable scalar–valued function a on B^* (a "nonlinear functional") is said *to be of class* $*$–C^1, or $*$*–differentiable of class* C^1, if it is of class C^1 in the usual sense and its differential $d_x a$ at an arbitrary point $x \in B$, which is an element of the second dual, $d_x a \in B^{**}$, belongs in fact to B,

$$d_x a \in B (\subset B^{**}) \ \forall x \in B^*.$$

Equivalently, we can say that a is of class $*$–C^1 if da is a continuous mapping from B^* to B,

$$da : B^* \longrightarrow B, \ x \mapsto d_x a \in B.$$

A scalar–valued function *a of class* $*$–C^1 *is said to be of class* $*$–C^k, if it is of class C^k (in the usual sense), and for $\forall \xi \in B^*$ the scalar valued function on B^*,

$$< \xi, da >: B^* \longrightarrow \mathbb{R}, \ x \mapsto < \xi, d_x a >,$$

where $< \cdot, \cdot >$ is the pairing between B^* and B, is of class $*$–C^{k-1} (definition by induction).

Finally, we say that a *differentiable mapping of class* C^k *(in the usual sense),* $\Phi : B_1^* \longrightarrow B_2^*$, *is of class* $*\text{–}C^k$, if for every scalar–valued function a of class $*\text{–}C^k$ on B_2^*, the composition

$$a \circ \Phi : B_1^* \xrightarrow{\Phi} B_2^* \xrightarrow{a} \mathbb{R}$$

is of class $*\text{–}C^k$ on B_1^*. We shall also say that Φ *is* $*$*–differentiable of class* C^k. Evidently, the composition of two $*$–differentiable mappings of class C^k is again a $*$–differentiable mapping of class C^k. The chain rule for differentiation implies that the mapping Φ is of class $*\text{–}C^k$, if every composition

$$\eta \circ \Phi : B_1^* \xrightarrow{\Phi} B_2^* \xrightarrow{\eta} \mathbb{R},\ \ x \mapsto < \eta, \Phi(x) >,\ \ \forall \eta \in B_2,$$

is a scalar–valued function of class $*\text{–}C^k$ on B_1^*.

A Banach manifold, modeled on a space B^*, is said to be $*$*–smooth of class* C^k, or *of class* $*\text{–}C^k$, if the mappings $B^* \longrightarrow B^*$, induced by the overlapping neighborhoods of corresponding atlases, are of class $*\text{–}C^k$. Between two such manifolds, mappings of class $*\text{–}C^k$ are well defined in an obvious way. For a fixed k, all $*\text{–}C^k$ manifolds and all $*\text{–}C^k$ mappings between them form a category. Cotangent spaces in this category consist, *by definition,* of differentials of $*$–smooth scalar valued functions at corresponding points, hence they are isomorphic to B, *not to* B^{**}. Tangent spaces are isomorphic to B^*.

Let U be a $*$–smooth manifold modeled on a Banach space B^*, $f : U \longrightarrow M$ be a $*$–smooth mapping into an n–dimensional manifold M of appropriate class. A pair

$$(u, \lambda),\quad u \in U,\ \lambda \in \left(T^*_{f(u)}M\right) \setminus \{0\},$$

will be called a *Lagrangian point of* f if $\lambda f'_u = 0$, where $\lambda f'_u$ is the composition of the differential $f'_u : T_u U \longrightarrow T_{f(u)}M$ with the linear functional $\lambda : T_{f(u)}U \longrightarrow \mathbb{R}$, i.e., is the image of λ under the adjoint $f_u^{'*}$,

$$\lambda f'_u \stackrel{\text{def}}{=} \lambda \circ f'_u = f_u^{'*}\lambda : T_u U \longrightarrow \mathbb{R}.$$

Thus for an arbitrary pair (u, λ), $u \in U$, $\lambda \in T^*_{f(u)}M$, we have

$$\lambda f'_u \in T^*_u U \approx B,$$

and, if (u, λ) is a Lagrangian point, $\lambda f'_u$ is the zero element of the fibre $T^*_u U$. The first component u of every Lagrangian point (u, λ) of f is a critical point of f, the second component λ is a *Lagrange multiplier* associated with the critical point u.

Consider the induced bundle $f^*(T^*M)$ over U defined by the mapping $f : U \longrightarrow M$ and the cotangent bundle T^*M,

$$f^*(T^*M) = \left\{(u, \lambda)\Big|\ u \in U,\ \lambda \in T^*_{f(u)}M\right\} = \bigcup_{u \in U} T^*_{f(u)}M.$$

Besides the canonical projection $(u, \lambda) \mapsto u$, it defines canonically two additional mappings

$$f^{'*} : (u, \lambda) \mapsto \lambda f'_u,\quad \varphi : (u, \lambda) \mapsto \lambda,$$

represented in the diagram

$$\begin{array}{ccc} & f^*(T^*M) & \\ & f^{'*} \quad \varphi & \\ & \swarrow \quad \searrow & \\ T^*U & & T^*M \end{array} .$$

Identifying U with the trivial section of T^*U, $U \subset T^*U$, we can assert that the set of the Lagrangian points of f is identical with the preimage of U under the mapping $f^{'*}$ less the trivial section of $f^*(T^*M)$. Furthermore, every tangent space to T^*U at an arbitrary point of the trivial section is canonically represented as a direct sum of its *horizontal and vertical subspaces,*

$$T_uT^*U = T_uU \oplus T_u^*U, \ \forall u \in U \subset T^*U.$$

Denote by

$$Ver_u : T_uT^*U \longrightarrow T_u^*U, \ u \in U \subset T^*U,$$

the projector onto the vertical subspace. Let (u, λ) be a Lagrangian point of f. The linear mapping

$$f''_{(u,\lambda)} \stackrel{\text{def}}{=} Ver_u \circ (f^{'*})'_{(u,\lambda)} : T_{(u,\lambda)}f^*(T^*M) \stackrel{(f^{'*})'_{(u,\lambda)}}{\longrightarrow} T_uT^*U \stackrel{Ver_u}{\longrightarrow} T_u^*U$$

contains a complete information about the second order approximation of f at u.

Note that

$$ker\ \varphi'_{(u,\lambda)} \approx ker\ f'_u \subset T_uU.$$

Thus $f''_{(u,\lambda)}\Big|_{ker\ \varphi'_{(u,\lambda)}}$ is a well–defined mapping from $ker\ f'_u$ to T_u^*U. We call this mapping the *second derivative of f at the Lagrangian point* (u, λ) and denote in the sequel by

$$D^2_{(u,\lambda)}f : ker\ f'_u \longrightarrow T_u^*U.$$

This definition needs some clarification. At the first site, the natural choice for the second derivative is the mapping $(f^{'*})'_{(u,\lambda)}$ defined at all points $(u, \lambda) \in f^*(T^*M)$. But such a definition would be completely useless, since, by virtue of the implicit function theorem, in some neighborhood of every regular point, local coordinates could be introduced in which f is linear, hence it is senseless to consider in such points second derivatives. Concerning the horizontal component of the mapping $(f^{'*})'_{(u,\lambda)}$ at a Lagrangian point, it is easy to see that it coincides with the differential of the canonical projection $(u, \lambda) \mapsto u$ and has no connections with the differential properties of the mapping f.

Define, finally,

$$\mathcal{L}_{(u,\lambda)}(f) = \varphi'_{(u,\lambda)}(ker\ f''_{(u,\lambda)}) \subset T_\lambda(T^*M).$$

The choice of arbitrary local coordinates in U leads to the representation

$$T_{(u,\lambda)}f^*(T^*M) = T_\lambda(T^*_{f(u)}M) \oplus T_uU = T_\lambda(T^*_{f(u)}M) \oplus B^*.$$

If local coordinates are introduced in M as well, then we obtain the representation

$$\lambda = (p, f(u)), \ p \in \mathbb{R}^{n*}, \ T^*_{f(u)}M = \mathbb{R}^{n*}, \ T_{(\lambda,u)}f^*(T^*M) = \mathbb{R}^{n*} \oplus B^*.$$

The mappings φ'_u and $f''_{(u,\lambda)}$ take the form

$$\varphi'_u : (\xi, v) \mapsto (\xi, f'_u v),\ \ \xi \in \mathbb{R}^{n*},\ v \in B^*;\ \ f''_{(u,\lambda)} : (\xi, v) \mapsto \xi f'_u + p f''_u v, \tag{1.1}$$

where pf''_u is the second derivative at u of the $*$–smooth function $pf : B^* \longrightarrow \mathbb{R}$. Thus the linear mapping $pf''_u : B^* \longrightarrow B$ is symmetric (selfadjoint). Denote

$$F = \left\{ \xi f'_u \middle| \xi \in \mathbb{R}^{n*} \right\},\ E = \left\{ pf''_u v \middle| v \in B^* \right\}.$$

Note that F is the image of the tangent space to the fibre of the bundle $f^*(T^*M)$ under the mapping $f''_{(u,\lambda)}$. At the same time the subspace E depends on the choice of the local coordinates in U. Denote

$$\delta_{(u,\lambda)} = dim(F \cap \overline{E}) - dim(F \cap E).$$

The following Proposition shows that the number $\delta_{(u,\lambda)}$ is independent on the coordinate choice.

Proposition 1.1. *$\mathcal{L}_{(u,\lambda)}(f)$ is an isotropic subspace of dimension $(n - \delta_{(u,\lambda)} f)$ in the ($2n$–dimensional) tangent space $T_\lambda(T^*M)$ to the cotangent bundle T^*M with the natural symplectic structure.*

Proof. First we proof the isotropy. The choice of local coordinates in M identifies $T_\lambda(T^*M)$ with $\mathbb{R}^{n*} \oplus \mathbb{R}^n$, and the canonical symplectic form σ takes the form

$$\sigma\left((\xi_1, \eta_1), (\xi_2, \eta_2)\right) = \xi_1 \eta_2 - \xi_2 \eta_1,\ \ \xi_i \in \mathbb{R}^{n*},\ \eta_i \in \mathbb{R}^n.$$

We must prove the implication

$$(\xi_i, \eta_i) \in \mathcal{L}_\lambda(f) \Longrightarrow \xi_1 \eta_2 = \xi_2 \eta_1.$$

We have $\eta_i = f'_u v_i$, where $v_i \in B^*,\ \xi_i f'_u + pf''_u v_i = 0$. Hence

$$\xi_1 \eta_2 = \xi_1 f'_u v_2 = - < v_2, pf''_u v_1 >,$$

and the identity to be proved follows from the symmetry of the operator pf''_u.

Now we turn to the dimension of $\mathcal{L}_{(u,\lambda)}(f)$. Formula (1.1) and the symmetry of pf''_u imply that

$$ker\ f''_{(u,\lambda)} \approx coker\ f'_u \oplus E \cap F \oplus E^\perp,\ \ E^\perp = ker\, pf''_u = ker\ f''_{(u,\lambda)} \cap B^*.$$

Furthermore, $ker\ f'_u = F^\perp$. Hence

$$\mathcal{L}_{(u,\lambda)}(f) \approx coker\ f'_u \oplus E \cap F \oplus E^\perp / E^\perp \cap F^\perp.$$

Thus

$$Dim\,\mathcal{L}_{(u,\lambda)}(f) = corank\, f'_u + dim\, F - \delta_{(u,\lambda)} f = corank\, f'_u + rank\, f'_u - \delta_{(u,\lambda)} f.$$

If $dim\,\mathcal{L}_{(u,\lambda)}(f) = n$, then $\mathcal{L}_{(u,\lambda)}(f)$ is a Lagrangian subspace in $T_\lambda(T^*M)$. In this case we call $\mathcal{L}_{(u,\lambda)}$ the *Lagrangian derivative* of f, or $\mathcal{L}$–*derivative*, at the Lagrangian point (u,λ).

The $\mathcal{L}$–derivative, i.e., the image $im\,\varphi'_{(u,\lambda)}\Big|_{ker\, f''_{(u,\lambda)}}$, could be considered as a dual object to the second derivative. Note that

$$ker\, D^2_{(u,\lambda)} f = ker\, \varphi'_{(u,\lambda)} \cap ker\, f''_{(u,\lambda)}.$$

In coordinates we obtain

$$D^2_{(u,\lambda)} f : v \mapsto p f''_u v,\;\; v \in ker\, f'_u.$$

The following assertion is a result of direct calculations.

Lemma. *The relations (1), (2) below are equivalent:*

(1) $$ker\, D^2_{(u,\lambda)} f = 0$$

(2) $$\overline{im f''_{(u,\lambda)}} = T^*_u U.$$

Suppose that $D^2_{(u,\lambda)} f$ is injective and U is finite–dimensional. Then the implicit function theorem implies that the germ at (u,λ) of the set of Lagrangian points is a germ of a smooth n–dimensional manifold. The restriction of φ on this germ is a Lagrange immersion into T^*M, and $\mathcal{L}_{(u,\lambda)}(f)$ is the tangent space to the obtained germ of a Lagrange submanifold.

Finally, we call *the Hessian of the mapping* f at (u,λ) the quadratic form

$$Hess_{(u,\lambda)} f : ker\, f'_u \longrightarrow \mathbb{R},\;\; Hess_{(u,\lambda)} f(v) = < v, D^2_{(u,\lambda)} f v > .$$

Negative and positive indices of the quadratic form $Hess_{(u,\lambda)} f(v)$*, (which are nonnegative integers or $+\infty$) are important characteristics of the Lagrangian point (u,λ). In particular, for the optimization problems they give essential information about the configuration of the image of a small neighborhood of the point under f. We formulate here only the simplest assertion of this kind. For deeper results in this respect and particularities cf. [2,5,6].

Proposition 1.2. *Let $\widetilde{\gamma}$ be a nonconstant germ of a smooth curve on M with the initial point $f(u)$. Assume that $im\,\gamma \cap f(O_u) = f(u)$ for a certain representative γ of the germ $\widetilde{\gamma}$*

*The negative (positive) index of a quadratic form Q on a linear space B, $ind_- Q$, $(ind_+ Q)$, is the maximal dimension of subspaces in B where Q is negative (positive).

and certain neighborhood O_u of u in U. Then there exists a Lagrangian point (u,λ) such that

$$<\lambda,\dot\gamma(0)>\le 0 \text{ and } ind_-\, Hess_{(u,\lambda)}f < corank\, f'_u.$$

There is an intimate tie between the indices of the Hessian and the Lagrangian derivative. Certainly, the Lagrangian derivative at a fixed point could not give any estimates for the indices of the Hessian in that point, but it is possible to express the increments of the indices along a one–parametric family of Lagrangian points through the Maslov index of the corresponding family of Lagrangian derivatives. Cf. [7] and §4 of this paper.

We emphasize that the isotropic subspace $\mathcal{L}_{(u,\lambda)}(f)$ is called $\mathcal{L}$–derivative only in case when its dimension is n, i.e. when it is a Lagrangian subspace. This is always the case if U is finite–dimensional, but by far not always in infinite–dimensional case. It turns out that if one of the indices of $Hess_{(u,\lambda)}f$ is finite then $\mathcal{L}_{(u,\lambda)}(f)$ could always be extended in a natural way to an n–dimensional Lagrangian subspace, which should be called the $\mathcal{L}$–derivative. Below we give a precise formulation of this result. In this paper the result is not used, therefore the proof will be given in subsequent publications.

Let (u,λ) be a Lagrangian point of f and N be a germ of a submanifold in U at u, hence (u,λ) is a Lagrangian point of $f\Big|_N$. If N is finite–dimensional then $\mathcal{L}_{(u,\lambda)}(f\Big|_N)$ is a Lagrangian subspace. Denote by $\mathcal{N}$ the set of all such germs partially ordered by inclusion. Then $\left\{\mathcal{L}_{(u,\lambda)}(f\Big|_N)\right\}_{N\in\mathcal{N}}$ is a generalized sequence of points of the Lagrangian Grassmannian $L(T_\lambda(T^*M))$.

Theorem. *The limit $\mathcal{N}$-$lim\,\mathcal{L}_{(u,\lambda)}(f\Big|_N)$ exists if one of the indices of $Hess_{(u,\lambda)}f$ is finite.*

This limit is the precise definition of the $\mathcal{L}$–derivative at the Lagrangian point (u,λ). It contains the isotropic subspace $\mathcal{L}_{(u,\lambda)}(f)$. But it is not enough to prove the existence of the limit, we must compute it. Introducing local coordinates, we can assume that U is a Banach space and u its origin. Let $U_0\subset U$ be an arbitrary linear subspace which is dense in U, and $\mathcal{N}_0$ be the set of all finite–dimensional subspaces $\mathcal{N}_0\subset U_0$, partially ordered by inclusion. Thus, $\mathcal{N}_0\subset\mathcal{N}$. The following assertion is an essential addition to the Theorem, making possible to explicitly compute the limit indicated in the Theorem.

Proposition 1.3. *Under the hypothesis of the Theorem the limit*

$$\mathcal{N}_0\text{-}lim\,\mathcal{L}_{(u,\lambda)}(f\Big|_{N_0}) = \mathcal{N}\text{-}lim\,\mathcal{L}_{(u,\lambda)}(f\Big|_N).$$

<u>Remark</u>. According to the definition of the $\mathcal{L}$–derivative we have

$$T_\lambda(im f'_u)^\perp\subset\mathcal{L}_{(u,\lambda)}(f).$$

In particular,

$$T_\lambda(\mathbb{R}\lambda)\subset\mathcal{L}_{(u,\lambda)}(f).$$

Thus the subspace $\mathcal{L}_{(u,\lambda)}(f)$ contains a one–dimensional subspace which depends only on λ and not on f. This makes possible to make all our constructions in the $2(n-1)$–dimensional symplectic space $T_\lambda(\mathbb{R}\lambda)^\angle/T_\lambda(\mathbb{R}\lambda)$ and not in the $2n$–dimensional space $T_\lambda(T^*M)$, considering $\mathcal{L}_{(u,\lambda)}(f)$ as a ($(n-1)$–dimensional) Lagrangian subspace in $T_\lambda(\mathbb{R}\lambda)^\angle/T_\lambda(\mathbb{R}\lambda)$. Certainly, the same reduction could be described in the language of contact geometry, considering not T^*M but its projectivization $\mathbb{P}T^*M$, which possesses the natural structure of a $(2n-1)$–dimensional contact manifold. In a certain sense the contact formulation is more natural, though we shall not use it here for the following reasons. Even considering $\mathbb{P}T^*M$ instead of T^*M, we would be led to consider homogeneous coordinates on projective spaces $\mathbb{P}T_q^*M$, thus constantly returning to the same space T^*M.

The following assertion formulates the "chain rule" for $\mathcal{L}$–derivatives. It easily follows from the definitions and has many useful consequences.

Proposition 1.4. *Suppose $f_i : U \longrightarrow M_i$ are $*$–smooth mappings, $u \in U$, $\lambda_i \in T^*_{f_i(u)}M_i$, $i = 1,2,3$. Suppose further that $(u, (-\lambda_0, \lambda_1))$ is a Lagrangian point for the mapping $(f_0, f_1) : U \longrightarrow M_0 \times M_1$, $(u, (-\lambda_1, \lambda_2))$ is a Lagrangian point for $(f_1, f_2) : U \longrightarrow M_1 \times M_2$, and*

$$(-\eta_1, \eta_2) \in \mathcal{L}_{(u,(-\lambda_0,\lambda_1))}(f_0, f_1) \subset T_{-\lambda_0}(T^*M_0) \times T_{\lambda_1}(T^*M_1),$$
$$(-\eta_0, \eta_1) \in \mathcal{L}_{(u,(-\lambda_1,\lambda_2))}(f_1, f_2) \subset T_{-\lambda_1}(T^*M_1) \times T_{\lambda_2}(T^*M_2),$$

then $(u, (-\lambda_2, \lambda_0))$ is a Lagrangian point of the mapping $(f_2, f_0) : U \longrightarrow M_2 \times M_0$, and $(-\eta_2, \eta_0) \in \mathcal{L}_{(u,(-\lambda_2,\lambda_0))}(f_2, f_0)$.

Suppose now that $M_i = M$, $i = 1,2,3$, and that the projections

$$\pi_i^{(0,1)} : (-\eta_0, \eta_1) \mapsto \eta_i, \ (-\eta_0, \eta_1) \in \mathcal{L}_{(u,(-\lambda_0,\lambda_1))}(f_0, f_1), \ i = 0,1$$
$$\pi_j^{(1,2)} : (-\eta_1, \eta_2) \mapsto \eta_j, \ (-\eta_1, \eta_2) \in \mathcal{L}_{(u,(-\lambda_1,\lambda_2))}(f_1, f_2), \ j = 1,2,$$

are invertible mappings of $\mathcal{L}_{(u,(-\lambda_0,\lambda_1))}$ on $T_{\lambda_i}(T^*M)$, $i = 0,1$, and of $\mathcal{L}_{(u,(-\lambda_1,\lambda_2))}$ on $T_{\lambda_j}(T^*M)$, $j = 1,2$, respectively. Then, Proposition 1.3 implies that the mappings

$$\pi_k^{(2,0)} : (-\eta_2, \eta_0) \mapsto \eta_k, \ (-\eta_2, \eta_0) \in \mathcal{L}_{(u,(-\lambda_2,\lambda_0))}(f_2, f_0), \ k = 2,0,$$

are also invertible.

Set

$$\Phi_{ij} = \pi_i^{(i,j)} \circ (\pi_j^{(i,j)})^{-1}, \ \Phi_{ji} = \Phi_{ij}^{-1}.$$

The property for the subspaces $\mathcal{L}_{(u,(-\lambda_i,\lambda_j))}(f_i, f_j)$ to be Lagrangian is equivalent to the fact that the mappings $\Phi_{ij} : T_{\lambda_j}(T^*M) \longrightarrow T_{\lambda_i}(T^*M)$ are symplectic. Proposition 1.3 also implies the important composition rule

$$\Phi_{20} = \Phi_{21} \circ \Phi_{10}.$$

We shall meet below concrete applications of this rule to control systems considering the boundary–value mappings on admissible curves, cf. §3.

The projections $\pi_i^{(0,1)}$, $i = 0,1$, are invertible only if f_0 and f_1 are submersions at u. The symplectic mapping Φ_{10} which represents the $2n$–dimensional $\mathcal{L}$–derivative of the mapping (f_0, f_1) represents also the n–dimensional $\mathcal{L}$–derivative of the mapping $f_1\Big|_{f_0=const}$, the restriction of f_1 to the level of f_0 through u. Direct calculations imply $\mathcal{L}_{(u,\lambda_1)}(f_1\Big|_{f_0=const}) = \Phi_{10}(T_{\lambda_0}(T^*_{f_0(u)}M))$.

§2. Smooth control systems and basic structures of Differential Geometry

1. Definition of smooth control systems. Suppose a smooth (locally trivial) fibre bundle over a smooth n–dimensional manifold M is given,

$$pr: \ W \longrightarrow M, \tag{2.1}$$

with the typical fibre U, a smooth r–dimensional manifold. Furthermore, suppose an arbitrary smooth fibrewise mapping of W into the tangent bundle TM is defined over the identity mapping of M,

$$f: W \longrightarrow TM, \ f(W_x) \subset T_xM \ \forall x \in M; \ W_x = pr^{-1}\{x\}. \tag{2.2}$$

We call the data (1.1)–(1.2) a *smooth control system*, (2.1) (or W) is called the *control space* of the system, the typical fibre U is called the *space of control parameters*; M is the *state space*, the cotangent bundle T^*M is the *phase space* of the system.

Morphisms between two control systems $f_i : W_i \longrightarrow TM_i$, $i = 1,2$, are, by definition, arbitrary commutative diagrams

$$\begin{array}{ccc} W_1 & \xrightarrow{f_1} & TM_1 \\ \Phi\downarrow & & F_*\downarrow \\ W_2 & \xrightarrow{f_2} & TM_2, \end{array} \tag{2.3}$$

where Φ is a smooth fibrewise mapping of control spaces, $F: M_1 \longrightarrow M_2$ is a diffeomorphism. We denote the morphism (2.3) by (Φ, F_*). If (Φ', F'_*) is a second morphism,

$$\begin{array}{ccc} W_2 & \xrightarrow{f_2} & TM_2 \\ \Phi'\downarrow & & F'_*\downarrow \\ W_3 & \xrightarrow{f_3} & TM_3, \end{array}$$

then their composition is defined, $(\Phi' \circ \Phi, F'_* \circ F_*)$, which again is a morphism between $f_1 : W_1 \longrightarrow TM_1$, and $f_3 : W_3 \longrightarrow TM_3$. The identity morphisms are defined in an obvious way. Thus a *category of smooth control systems* is introduced.

If Φ in (2.3) is a diffeomorphism and $F_* = id$, $(M_1 = M_2)$, then the morphism (Φ, id) is called a *feedback transformation*, and the corresponding control systems are said to be *feedback equivalent*. Feedback transformations are smooth fibre transformations of the control space over the identity map. If $W_1 = W_2$, then the feedback transformations are also called *guage transformations*. Two feedback equivalent control systems are equivalent in our category, hence the given definition of a control system is "feedback–invariant".

According to the usual "state–invariant" definition, a smooth control system,

$$\dot{x} = f(x,u) \in T_xM, \ (x,u) \in M \times U, \tag{2.4}$$

is a family of smooth vector fields on the state manifold M, indexed by a control parameter $u \in U$. The control space is the direct product $W = M \times U$ with the canonical trivialization,

$$pr: \ M \times U \longrightarrow M, \ (x,u) \mapsto x, \tag{2.5}$$

and the mapping $W \longrightarrow TM$ is given by f. Evidently, this definition, though invariant under coordinate transformations in M, is not feedback–invariant.

A measurable essentially bounded curve in the control space,*

$$\xi : [0,t] \longrightarrow W, \tag{2.6}$$

is called an *admissible control space trajectory* of the system (2.1)–(2.2) if its projection on M,

$$x = pr\, \xi : \ [0,t] \longrightarrow M, \tag{2.7}$$

is a Lipschitz curve in M satisfying for almost all τ the differential equation

$$\frac{dx}{d\tau}(\tau) = f \circ \xi(\tau), \ 0 \le \tau \le t. \tag{2.8}$$

The curve $x(\tau)$ is called an *admissible state space trajectory.* Admissible control space trajectories will be also called *admissible controls.*

The set of all measurable essentially bounded mappings w will be considered as a space with the following natural topology. Consider a metric ρ, compatible with the topology of W, and define the ε–neighborhood of a given $\widetilde{w}$ by the relation

$$\mathcal{O}(\varepsilon, \widetilde{w}) = \left\{ w \,\middle|\, \underset{0 \le \tau \le t}{ess\ sup}\, \rho(w(\tau), \widetilde{w}(\tau)) = \parallel \rho(w, \widetilde{w}) \parallel_\infty < \varepsilon \right\}.$$

Since the closure of the image of $\tilde{w}$ is compact, the introduced topology is independent on the choice of the metric.

Denote by Ω^t the space of all admissible controls (2.6), (considered as a subspace of the space of all measurable essentially bounded curves $[0,t] \longrightarrow W$).

Proposition 2.1 *The space Ω^t of all admissible controls can be given, in a natural way, the structure of a $*$–smooth Banach manifold modeled on the direct product*

$$\mathbb{R}^n \times (L_1^r[0,t])^* = \mathbb{R}^n \times L_\infty^r[0,t], \ n = dim\, M, \ r = dim\, U.$$

---

*A measurable mapping $w : [0,t] \longrightarrow W$ is said to be essentially bounded if the closure of the image under w of a subset of $[0,t]$ of full measure is compact. For simplicity, we shall assume in the sequel that the closure of the image of w is compact.

Proof. First we proof that the space of essentially bounded mappings $u : [0, t] \longrightarrow U$ is, in a natural way, a $*$–smooth Banach manifold modeled on $L^r_\infty[0, t]$.

For this we define coordinate mappings of ε–neighborhoods of an arbitrary element $\widetilde{u} : [0, t] \longrightarrow U$ into $L^r_\infty[0, t]$, assuming that a Riemannian metric ρ is fixed on U.

Denote by Q the (compact) closure of the image of $\widetilde{u}$ and consider a finite cover of Q by ε–balls $B_\varepsilon(p_i),\ i = 1, \dots , s$, centered at $p_i \in Q \subset \bigcup\limits_{\alpha=1}^{s} B_\varepsilon(p_\alpha) = B$. Set ε small enough to secure the following two conditions.

(1) The exponential mapping,

$$exp : TU \longrightarrow U \times U,$$

is invertible on $\bigcup\limits_{p\in Q} (p, B_\varepsilon(p)) \subset Q \times B$,

$$exp^{-1} : \bigcup_{p\in Q} (p, B_\varepsilon(p)) \longrightarrow \bigcup_{p\in Q} T_pU.$$

(2) The tangent subbundles $TB_\varepsilon(p_i) \subset TU,\ i = 1, \dots , s$, are trivial, with trivializations

$$\zeta_i : TB_\varepsilon(p_i) \longrightarrow \mathbb{R}^r .$$

Introduce the (measurable) mapping

$$\zeta : TB \longrightarrow \mathbb{R}^r,\ \zeta(z) = \zeta_i(z),\ i = \min_\alpha \left\{\alpha \,\middle|\, z \in TB,\, pr(z) \in B_\varepsilon(p_\alpha)\right\}.$$

For every measurable curve $u(\tau), 0 \le \tau \le t$, in the ε–neighborhood of $\widetilde{u}$, $u(\tau) \in B_\varepsilon(\widetilde{u}(\tau))$ $0 \le \tau \le t$, we can define a measurable curve $v(\tau)$, $0 \le \tau \le t$, in $\mathbb{R}^r$ according to the correspondence

$$u(\tau) \mapsto \zeta \circ exp^{-1}(\widetilde{u}(\tau), u(\tau)) = v(\tau) \in \mathbb{R}^r,$$

which is an injection and satisfies the relation

$$\operatorname*{ess\,sup}_{0\le\tau\le t} |v(\tau)| = \| v \|_\infty \le \delta(\varepsilon) \to 0\, (\varepsilon \to 0).$$

Conversely, every measurable curve $v(\tau)$, $0 \le \tau \le t$, in $\mathbb{R}^r$, satisfying the relation $\| v \|_\infty \le \delta$, with δ sufficiently small, could be obtained in this way. Indeed, denote the restriction of ζ to T_pB, $p \in Q$, by ζ_p. Then

$$u(\tau) = exp \circ \zeta^{-1}_{\widetilde{u}(\tau)} v(\tau) = (\zeta \circ exp^{-1})^{-1} v(\tau) \in B_\varepsilon(\widetilde{u}(\tau)).$$

From here the assertion is easily deduced. We now turn to the proof of the Proposition.

We shall show that there is a natural one–to–one correspondence between admissible controls, sufficiently close to a given admissible control $\widetilde{\xi}$, and arbitrary pairs $(x(0), u)$, where $x(0)$ is sufficiently close to the initial condition $\widetilde{x}(0) = pr\,\widetilde{\xi}(0)$ and $u : [0, t] \longrightarrow U$

is sufficiently close (in the *ess sup* topology) to the mapping $\widetilde{u}:[0,t]\longrightarrow U$ corresponding to $\widetilde{\xi}$.

To define the correspondence consider the projection

$$\widetilde{x}(\tau)=pr\,\widetilde{\xi}(\tau),\ \tau\in[0,t],$$

and take a "tubular" ε–neighborhood of $\widetilde{x}$,

$$T:[0,t]\times B_\varepsilon\longrightarrow M,\ T(\tau,0)=\widetilde{x}(\tau),\ \tau\in[0,t], \tag{2.9}$$

where $B_\varepsilon\subset\mathbb{R}^n$ is an ε–ball centered at the origin, T is a diffeomorphism for every fixed τ, Lipschitz in τ,

$$T(\tau,\cdot)=T_\tau:B_\varepsilon\longrightarrow M;\ x=T_\tau(q),\ q=T_\tau^{-1}(x),\ q\in B_\varepsilon, x\in T_\tau(B_\varepsilon).$$

Since $[0,t]\times B_\varepsilon$ is contractible, the induced bundle $T^*(W)$ is trivial,

$$T^*(W)=\bigcup_{(\tau,q)}\left((\tau,q),W_{T(\tau,q)}\right)\approx([0,t]\times B_\varepsilon)\times U,\quad(\tau,q)\in[0,t]\times B_\varepsilon.$$

Every trivialization

$$\vartheta:([0,t]\times B_\varepsilon)\times U\longrightarrow T^*(W)$$

generates a continuous family of diffeomorphisms

$$\vartheta(\tau,q):U\longrightarrow W_{T(\tau,q)},\ \tau\in[0,t],q\in B_\varepsilon,u\in U,$$

smooth in q and Lipschitz in τ. Introduce the mapping

$$(\tau,x,u)\mapsto f_\tau(x,u)\overset{\text{def}}{=}f\circ\vartheta(\tau,T_\tau^{-1}(x))(u),\quad\tau\in[0,t],\,x\in T_\tau B,\,u\in U,$$

which is smooth in x,u and Lipschitz in τ.

Every admissible state space trajectory $x(\tau)=pr\,\xi(\tau),\ 0\le\tau\le t$, in the tubular neighborhood (2.9) of $\widetilde{x}$, $x(\tau)\in T(\tau,B_\varepsilon)$, is a solution of the equation

$$\frac{dx}{d\tau}=f_\tau(x,u(\tau)),\ 0\le\tau\le t, \tag{2.10}$$

where $u(\tau)$ is uniquely defined on $[0,t]$, (up to a set of measure zero), by the relation

$$u(\tau)=\vartheta^{-1}(\tau,T_\tau^{-1}(x(\tau)))(\xi(\tau)),\ 0\le\tau\le t. \tag{2.11}$$

For any preassigned $\delta>0$, all sufficiently small $\varepsilon>0$, and all admissible controls ξ, sufficiently close (in the *ess sup* topology) to $\widetilde{\xi}$, the inequality $\|\rho(u,\widetilde{u})\|_\infty<\delta$ holds, where $\widetilde{u}$ corresponds to $\widetilde{\xi}$ according to (2.11). Conversely, for every $u(\tau)$, satisfying the last inequality for a sufficiently small δ, there exists a solution $x(\tau)$, $0\le\tau\le t$, of the equation (2.10) with the initial condition $|x(0)-\widetilde{x}(0)|<\delta$. This solution is an admissible state space trajectory corresponding to the admissible control space trajectory ξ, obtained by inverting the relation (2.11),

$$\xi(\tau) = \vartheta(\tau, T_\tau^{-1}(x(\tau)))(u(\tau)), \ 0 \le \tau \le t.$$

ξ is arbitrarily close to $\widetilde{\xi}$ for sufficiently small $\delta > 0$. Thus, every pair

$$(x(0), u), \ u : [0, t] \longrightarrow U, \ \| \rho(u, \widetilde{u}) \|_\infty < \delta, \ |x(0) - \widetilde{x}(0)| < \delta,$$

uniquely defines, for sufficiently small δ, an admissible control in a certain neighborhood of $\widetilde{\xi}$, and all such controls could be obtained in this way. This proves the Proposition.

Definition of smooth control systems introduced here is general enough to include as special cases basic differential–geometric structures. Below we give several examples.

An extensive and important class of control systems is defined by (locally trivial) smooth subbundles of the tangent bundle TM, considered as control spaces W, and the corresponding inclusion maps

$$f : W \subset TM.$$

Many standard geometric structures are reduced to such systems, the structure type depending on the choice of the typical fibre U of the control space which, in this case, is a submanifold of $\mathbb{R}^n$,

$$U \subset \mathbb{R}^n, \ n = dim\, M.$$

(1) U is an ellipsoid with center at the origin. We obtain the Riemannian geometry. Admissible trajectories in the state space are arbitrary Lipschitz curves $x(t)$ of length t, parametrized by the arc length.
(2) U is a strongly convex body in $\mathbb{R}^n$, symmetric with respect to the origin. This is the case of the Finsler geometry.
(3) U is a hyperboloid, symmetric with respect to the origin — the case of pseudo–Riemannian geometry.
(4) U is a linear subspace in $\mathbb{R}^n$ of an arbitrary dimension. We come to the theory of distributions (in the differential–geometric sense). Admissible curves are the integral curves of the distribution.
(5) U is the intersection of an ellipsoid centered at the origin with a linear subspace. We obtain the sub–Riemannian geometry.
(6) The "affine" versions (i.e. translates with respect to the origin) of the structures (1)–(5). Though not very popular in geometry, they are of utmost importance in applications to Mechanics and Mathematical Physics.

Examples (1)–(6) could be generalized in the following way. Suppose the control space (2.1) is a (locally trivial) subbundle of an arbitrary vector bundle over M, $E \longrightarrow M$, with the typical fibre coinciding with one of the above mentioned types (1)–(6), and f in (2.2) is the restriction on W of a certain fibrewise mapping $E \longrightarrow TM$, linear on fibres; in the examples considered f was an embedding. These broader classes of systems include "singular" versions of geometric structures (1)–(6) with degenerations at certain points.

2. The boundary–value mapping. The mapping

$$F_\tau : \Omega^t \longrightarrow M, \ F_\tau(\xi) = pr \circ \xi(\tau),$$

which evaluates the admissible state–space trajectory $x = pr \circ \xi$ at the moment τ, is a $*$–smooth submersion for $\forall \tau \in [0,t]$. At the same time, the *boundary–value mapping*

$$(F_0, F_t) : \Omega^t \longrightarrow M \times M, \ (F_0, F_t)(\xi) = (F_0(\xi), F_t(\xi))$$

is, in general, not a submersion. Critical points of boundary–value mappings are called *extremal controls* of the control system.

Denote by $\Omega^t_{x_0}$ the set of admissible controls $\xi \in \Omega^t$ subject to the condition $x(0) = pr \circ \xi(0) = x_0$. In other words, $\Omega^t_{x_0}$ is the level set over the point $x_0 \in M$ of the submersion F_0. Evidently, $\Omega^t_{x_0}$ is a $*$–smooth Banach manifold modeled on $L^r_\infty[0,t]$.

We introduce the *endpoint mapping*

$$F_{0t} = F_t\Big|_{\Omega^t_{x_0}} : \Omega^t_{x_0} \longrightarrow M.$$

Critical points of F_{0t} are exactly the extremal controls in $\Omega^t_{x_0}$.

Let σ be the natural symplectic form on T^*M and put

$$H(\lambda, z) = \lambda f(z) \stackrel{\text{def}}{=} \lambda \circ f(z), \ z \in W, \ \lambda \in T^*_{pr(z)}M.$$

H is a smooth function on the direct product of the fibred manifolds W, T^*M over M. We call it the *Hamiltonian of the control system (2.1)–(2.2).*

Proposition 2.2. *The triple*

$$(\xi, (-\lambda_0, \lambda_t)), \ \xi \in \Omega^t, \ \lambda_0 \in T^*_{pr\,\xi(0)}M, \ \lambda_t \in T^*_{pr\,\xi(t)}M$$

is a Lagrangian point of the boundary value mapping (F_0, F_t) *iff there exists a Lipschitz curve*

$$\tau \mapsto \lambda_\tau \in T^*_{pr\,\xi(\tau)}M, \ \tau \in [0,t],$$

satisfying the condition

$$j_2^*\left(\frac{d\lambda_\tau}{d\tau} \rfloor \sigma\right) = -dH,$$

where j_1, j_2 *are the natural projections,*

$$\begin{array}{ccc} W & \underset{M}{\times} & T^*M \\ {\scriptstyle j_1}\swarrow & & \searrow{\scriptstyle j_2} \\ W & & T^*M. \end{array}$$

Proof. If the curve $\tau \mapsto \lambda_\tau$ exists then it is unique (for a given Lagrangian point). Indeed, according to Proposition 2.1, (see (2.10)), for an appropriate trivialization of W along the trajectory $pr\,\xi$ the admissible control ξ is represented as $\tau \mapsto (\widetilde{u}(\tau), \widetilde{x}(\tau)) \in U \times M$, and the controls close to ξ are exactly the solutions of the equation

$$\frac{dx}{d\tau} = f_\tau(x, u(\tau)),\ \ 0 \le \tau \le t,\ \ u \in U,\ \ x \in M,$$

close to $(\widetilde{u}(\tau), \widetilde{x}(\tau))$.

The Lipschitz curve $\tau \mapsto \lambda_\tau \in T^*_{\widetilde{x}(\tau)}M$ satisfies the assumptions of the Proposition 2.2 if and only if it satisfies the equations

$$\frac{d\lambda_\tau}{d\tau} = \vec{H}_\tau(\lambda_\tau, \widetilde{u}(\tau)), \tag{2.12}$$

$$\frac{\partial H_\tau}{\partial u}(\lambda_\tau, \widetilde{u}(\tau)) = 0, \tag{2.13}$$

where $\tau \mapsto \vec{H}_\tau(\lambda, u)$ is the Hamiltonian vector field on T^*M corresponding to the Hamiltonian

$$\lambda \mapsto \lambda f_\tau(x, u) = H_\tau(\lambda, u),\ \ \lambda \in T^*_x M,\ \ x \in M.$$

To prove the Proposition, we shall compute explicitly the differential of the mapping F_t, using the variation formula, cf. Introduction. We have

$$\begin{aligned} F_t(x_0, u(\cdot)) =& x_0\, \overrightarrow{exp} \int_0^t f_\tau(\cdot, u(\tau))d\tau = \\ & x_0\, \overrightarrow{exp} \int_0^t \widetilde{f}_\tau d\tau \circ \overrightarrow{exp} \int_0^t \overrightarrow{exp} \int_t^\tau ad\, \widetilde{f}_\theta d\theta (f_\tau(\cdot, u(\tau)) - \widetilde{f}_\tau) d\tau, \end{aligned}$$

where $\widetilde{f}_\tau(x) = f_\tau(x, \widetilde{u}(\tau))$. Thus,

$$\begin{aligned} \lambda F'_t\Big|_{\widetilde{u}}(\delta u, \delta x_0) = \int_0^t \lambda \overrightarrow{exp} \int_t^\tau ad\, \widetilde{f}_\theta d\theta \frac{\partial f_\tau}{\partial u}\Big|_{\widetilde{u}(\tau)} \delta u(\tau) d\tau + \lambda(\overrightarrow{exp} \int_0^t \widetilde{f}_\tau d\tau)_* \delta x_0 = \\ \int_0^t \overrightarrow{exp} \int_t^\tau \vec{\widetilde{H}}_\theta d\theta \frac{\partial H_\tau}{\partial u}(\lambda, \widetilde{u}(\tau)) \delta u(\tau) d\tau + \lambda(\overrightarrow{exp} \int_0^t \widetilde{f}_\tau d\tau)_* \delta x_0, \end{aligned}$$

where $\widetilde{H}_\theta(\lambda) = H_\theta(\lambda, \widetilde{u}(\theta))$. Hence,

$$\lambda_t F'_t\Big|_{\widetilde{u}}(\delta u, \delta x_0) = \int_0^t \frac{\partial H_\tau}{\partial u}(\lambda_\tau, \widetilde{u}(\tau)) \delta u(\tau) d\tau + \lambda_0 \delta x_0,$$

where $\tau \mapsto \lambda_\tau$ satisfies (2.12). The equality $\lambda_t F'_t\Big|_{\widetilde{u}} - \lambda_0 = 0$ implies (2.13).

The curves $\tau \mapsto (\xi(\tau), \lambda_\tau)$ in Proposition 2.2 will be called *extremals*.

§3. $\mathcal{L}$–Derivatives of Endpoint Mappings. Regular Extremals

We preserve the notations of the previous section. Put

$$g_\tau(\lambda, u) = \lambda(\overrightarrow{exp}\int_t^\tau ad\,\widetilde{f}_\theta d\theta(f_\tau(\cdot, u) - \widetilde{f}_\tau)).$$

All Lagrangian points (ξ', λ'_t) of the mapping F_{0t}, sufficiently close to (ξ, λ_t), are characterized by the condition: there exists a Lipschitz curve $\tau \mapsto \lambda'_\tau \in T^*_{x'(\tau)}M$ such that

$$\dot{\lambda}'_\tau = \vec{g}_\tau(\lambda'_\tau, u'(\tau)),\ \frac{\partial}{\partial u} g_\tau((\lambda'_\tau, u'(\tau)) = 0, \tag{3.1}$$

where $\xi'(\tau) = (x'(\tau), u'(\tau)),\ x'(0) = \widetilde{x}(t)$. Note also, that $g_\tau\Big|_{u=\widetilde{u}(\tau)} \equiv 0$.
Linearising (3.1), we obtain

Proposition 3.1. *The relations*

$$\eta_t \in \mathcal{L}_{(\xi,\lambda_t)}(F_{0t}) \subset T_{\lambda_t}(T^*M)$$

are equivalent to the relations:

there exist curves $\tau \mapsto \eta_\tau \in T_{\lambda_t}(T^*M),\ \tau \mapsto v(\tau) \in T_{\widetilde{u}(\tau)}U$ *such that*

$$\dot{\eta}_\tau = \frac{\partial}{\partial u}\vec{g}_\tau v(\tau),\ \frac{\partial^2 g_\tau}{\partial u^2}(v(\tau), \cdot) = - < d_{\lambda_\tau}(\frac{\partial g_\tau}{\partial u}\Big|_{\widetilde{u}(\tau)} \cdot), \eta_\tau >,$$

$$\eta_0 \in T_{\lambda_t}(T^*_{\widetilde{x}(t)}M),\quad 0 \le \tau \le t.$$

The last two equations can be written in a more symmetric form as

$$\dot{\eta}_\tau = \frac{\partial \vec{g}_\tau}{\partial u} v(\tau),\ \frac{\partial^2 g_\tau}{\partial u^2}(v(\tau), \cdot) = \sigma(\frac{\partial \vec{g}_\tau}{\partial u}\cdot, \eta_\tau) \tag{3.2}$$

In §1, from general considerations, we derived that $\mathcal{L}_{(\xi,\lambda_\tau)}(F_{0t})$ is an isotropic space. This follows easily also from (3.2). Indeed,

$$\sigma(\eta^1_t, \eta^2_t) = \int_0^t \frac{d}{d\tau}\sigma(\eta^1_\tau, \eta^2_\tau)d\tau;\ \ \frac{d}{d\tau}\sigma(\eta^1_\tau, \eta^2_\tau) = \sigma(\frac{\partial \vec{g}_\tau}{\partial u} v^1(\tau), \eta^2_\tau) + \sigma(\eta^1_\tau, \frac{\partial \vec{g}_\tau}{\partial u} v^2(\tau)) =$$
$$\frac{\partial^2 g_\tau}{\partial u^2}(v^2(\tau), v^1(\tau)) - \frac{\partial^2 g_\tau}{\partial u^2}(v^1(\tau), v^2(\tau)) = 0.$$

We use further the following notation: $\frac{\partial^2 g_\tau}{\partial u^2}(\cdot, \cdot)$ denotes a symmetric bilinear form on $T_{\widetilde{u}(\tau)}U$; the corresponding quadratic form is denoted by $\frac{\partial^2 g_\tau}{\partial u^2}(\cdot)$. Furthermore,

$\dfrac{\partial^2 g_\tau}{\partial u^2} : T_{\widetilde{u}(\tau)}U \longrightarrow T^*_{\widetilde{u}(\tau)}U$ is a selfadjoint linear mapping. If it is nondegenerate, then the mapping $\left(\dfrac{\partial^2 g_\tau}{\partial u^2}\right)^{-1} : T^*_{\widetilde{u}(\tau)}U \longrightarrow T_{\widetilde{u}(\tau)}U$ is defined and is also selfadjoint. The corresponding quadratic form on $T^*_{\widetilde{u}(\tau)}U$ is denoted by $\left(\dfrac{\partial^2 g_\tau}{\partial u^2}\right)^{-1}(\cdot)$. Observe also that the expression

$$\frac{\partial^k g_\tau}{\partial u^k}\Big|_{\lambda_\tau} = \frac{\partial^k}{\partial u^k}\lambda_\tau f_\tau, \ \ k = 1, 2, \dots \tag{3.3}$$

is the k–th derivative of the function λf, restricted on the fibre $W_{\widetilde{x}(\tau)}$ of the bundle W. The first derivative has an invariant meaning, independent on the local trivialization of W and on the choice of coordinates in the typical fibre. The first derivative vanishes along the extremals, hence at these points the second derivative has an invariant meaning, which is the Hessian of the function $u \mapsto \lambda_\tau f_\tau(\widetilde{x}(\tau), u)$ at the point $\widetilde{u}(\tau)$.

An extremal $\tau \mapsto (\xi(\tau), \lambda_\tau)$ is called *regular* if the quadratic form $\dfrac{\partial^2}{\partial u^2}\lambda_\tau f_\tau\Big|_{\widetilde{u}(\tau)}$ is nondegenerate at every $\tau \in [0, t]$. The following proposition is an evident consequence of the relations (3.2).

Proposition 3.2. *Assume $\tau \mapsto (\xi(\tau), \lambda_\tau)$ is a regular extremal. Then the relation $\eta_t \in \mathcal{L}_{(\xi,\lambda_t)}(F_t)$ is equivalent to the requirement that the solution η_τ, $0 \le \tau \le t$, of the linear Hamiltonian system on $T_{\lambda_t}(T^*M)$ with the nonstationary quadratic Hamiltonian*

$$q_\tau(\eta) = -\frac{1}{2}\left(\frac{\partial^2 g_\tau}{\partial u^2}\right)^{-1}\left(\sigma(\frac{\partial \vec{g}_\tau}{\partial u}\cdot, \eta)\right), \ \ \eta \in T_{\lambda_t}(T^*M),$$

and the "end–condition", η_t satisfies the "initial condition" $\eta_0 \in T_{\lambda_t}(T^*_{\widetilde{x}(t)}M)$.

Corollary. *If the Lagrangian point (ξ, λ_t) defines a regular extremal, then $\mathcal{L}_{(\xi,\lambda_t)}(F_{0t})$ is a Lagrangian subspace, hence an $\mathcal{L}$–derivative of the mapping F_{0t} at (ξ, λ_t).*

In the sequel only regular extremals will be considered.

Put

$$\mathcal{D} = \left\{(w, \lambda) \in W \times_M (T^*M) \Big| \ \frac{\partial}{\partial u}(\lambda f)\Big|_w = 0, \ \frac{\partial^2}{\partial u^2}(\lambda f)\Big|_w \text{ is invertible.}\right\}$$

Proposition 3.3

(1) *All regular extremals are contained in $\mathcal{D}$.*
(2) *$\mathcal{D}$ is a smooth submanifold of dimension $2n$.*
(3) *Through every point of $\mathcal{D}$ passes a unique <u>continuous</u> regular extremal.*

Proof.

(1) is evident.
(2) follows from implicit function theorem.
(3) follows from implicit function theorem and the relations (2.12), (2.13).

The mapping $\varphi : (w, \lambda) \mapsto \lambda,\ (w, \lambda) \in \mathcal{D}$, is locally one–to–one, and its image, $\mathcal{D}_r$, (possibly empty), is open in T^*M. For many important problems, which include all examples of §2, this mapping is globally one–to–one, hence is a diffeomorphism. In this case the smooth function is defined,

$$h(\lambda) = \lambda f(\varphi^{-1}(\lambda)),\ \lambda \in \mathcal{D}_r,$$

which will be called the *master–Hamiltonian* of the corresponding control system. Evidently, the restrictions of h to fibres $\mathcal{D} \bigcap T_x^*M$ are positively homogeneous functions of degree one. If the master–Hamiltonian exists, then the regular extremals are exactly the curves $\tau \mapsto (\varphi^{-1}(\lambda_\tau), \lambda_\tau)$, where λ_τ is an arbitrary trajectory of the Hamiltonian system

$$\dot{\lambda} = \vec{h}(\lambda),\ \lambda \in \mathcal{D}_r.$$

All regular extremals are smooth.

We now describe the domains $\mathcal{D}_r$ and the master–Hamiltonians for control systems (1) — (6) enumerated in §2. The corresponding computations are straightforward. For Riemannian geometry we have $\mathcal{D}_r = T^*M \setminus M$, and the restriction of the master–Hamiltonian, $h\Big|_{T_x^*M}$ is the square root of a positive quadratic form on T_x^*M. For the Finsler structure, $\mathcal{D}_r$ again coincides with $T^*M \setminus M$, and $h\Big|_{T_x^*M}$ is the support function to the unit Finsler ball in T_xM, hence, is convex. For pseudo–Riemannian geometry of a given signature, $h\Big|_{T_x^*M}$ is the square root of a quadratic form of the same signature, $\mathcal{D}_r \bigcap T_x^*M$ is the positive cone of the quadratic form. For a distribution, $\mathcal{D}_r = \emptyset$. Finally, in case of sub–Riemannian geometry we have $\mathcal{D}_r \bigcap T_x^*M = T_x^*M \setminus (span\, W_x)^\perp$, the master–Hamiltonian $h\Big|_{T_x^*M}$ is the square root of a nonnegative quadratic form, with kernel equal to $(span\, W_x)^\perp$. For the "affine versions" of the considered structures the domains $\mathcal{D}_r$ remain unchanged, and to the Hamiltonians scalar functions are added, which are linear on fibres.

Returning to the $\mathcal{L}$–derivative, we note that, for regular extremals, the projection of the Lagrangian subspace

$$\mathcal{L}_{(\xi,(-\lambda_0,\lambda_t))}(F_0, F_t) \subset T_{\lambda_0}(T^*M) \times T_{\lambda_t}(T^*M)$$

onto the factors in the right–hand side is one–to–one. Thus we are in a situation discussed at the end of §1. Hence the symplectic mappings are defined, $\Phi_{t_1t_0} : T_{\lambda_0}(T^*M) \longrightarrow T_{\lambda_t}(T^*M)$, satisfying the conditions

$$\Phi_{t_2t_0} = \Phi_{t_2t_1} \circ \Phi_{t_1t_0},\ \mathcal{L}_{(\xi,\lambda_t)}(F_{0t}) = \Phi_{t0}(T_{\lambda_0}(T_{q_0}^*M)).$$

It is easily seen that $\Phi_{t_1t_0} = \left(e^{(t_1-t_0)\vec{h}}\right)_*$, where h is the master–Hamiltonian.

§4. Jacobi Curves

For an arbitrary $\lambda \in T^*M$ consider the hyperplane $(\mathbb{R}\lambda)^\angle$ in the symplectic space $T_\lambda(T^*M)$† and consider the factor space $\Sigma_\lambda = T_\lambda(T^*M)/(\mathbb{R}\lambda)^\angle$, which is a symplectic space of dimension $2(n-1)$. Let $L(\Sigma_\lambda)$ be the corresponding Lagrangian Grassmannian, the manifold of the Lagrangian subspaces in Σ_λ. At the same time $L(\Sigma_\lambda)$ is the manifold of Lagrangian subspaces in $T_\lambda(T^*M)$ containing λ.

Let $\tau \mapsto (\widetilde{\xi}(\tau), \lambda_\tau)$ be a regular extremal, $\widetilde{x}(\tau) = pr\,\widetilde{\xi}(\tau), 0 \le \tau \le t$. Denote by $F_{\tau,t} : \Omega^{t-\tau}_{\widetilde{x}(\tau)} \longrightarrow M$ the end–point mapping, defined on admissible state space trajectories, starting at $\widetilde{x}(\tau)$. For every $\tau \in [0,t]$ the $\mathcal{L}$–derivative $\mathcal{L}_{(\xi,\lambda_t)}(F_{\tau,t}) \in L(\Sigma_{\lambda_t})$ is defined. Consider the curve

$$J_{(\xi,\lambda)} : \tau \mapsto \mathcal{L}_{(\xi,\lambda_t)}(F_{\tau,t})$$

in the Lagrangian Grassmannian $L(\Sigma_{\lambda_t})$, which will be called the *Jacobi curve* associated with the regular extremal $\tau \mapsto (\xi(\tau), \lambda_\tau)$.

Note that the line $\mathbb{R}\lambda_t$ belongs to the kernel of the quadratic forms q_τ from Proposition 3.2. Hence the Hamiltonians q_τ are correctly defined on Σ_{λ_t}. Every linear Hamiltonian field on Σ_{λ_t} defines a vector field on $L(\Sigma_{\lambda_t})$ which we also call Hamiltonian. From the Proposition 3.2 and the variation formula for Hamiltonian systems, cf. Introduction, it follows that $J_{(\xi,\lambda)}$ is a trajectory of the Hamiltonian system on $L(\Sigma_{\lambda_t})$, defined by the Hamiltonian $-(\overrightarrow{exp}\int\limits_\tau^t \vec{q}_0 d\theta)q_\tau$. Furthermore, the relations at the end of §3 imply

$$J_{(\xi,\lambda)}(\tau) = \left(e^{(t-t_1)\vec{h}}\right)_* J_{(\xi,\lambda.)|_{[t_0,t_1]}}(\tau),\ \ 0 \le t_0 \le \tau \le t_1 \le t. \tag{4.1}$$

The Jacobi curve belongs to the Lagrangian Grassmannian, which is a homogeneous space for the symplectic group. We shall consider two curves in a Lagrangian Grassmannian to be *equivalent* if one is transformed into the other by a symplectic transformation. From (4.1) the following basic assertion follows: the germ of the Jacobi curve $J_{(\xi,\lambda)}$ at τ is defined, (up to the equivalence), by the germ at τ of the extremal $(\xi, \lambda.)$. To fully appreciate this fact, we should emphasize that the Jacobi curve is a curve in a special remarkable homogeneous space of the symplectic group, whereas the extremal belongs to a smooth manifold with a completely incomprehensible group of transformations.

We remind that the tangent vectors to the Lagrangian Grassmannian at the point $\Lambda \in L(\Sigma)$ are quadratic forms on $\Lambda \subset \Sigma$. Descending from germs of curves to 1–jets, we obtain

Proposition 4.1. *$\frac{d}{d\tau}J_{(\xi,\lambda)}(\tau)$ is a quadratic form of rank $dim\,U$ and signature $sgn\frac{\partial^2}{\partial u^2}\lambda_\tau f_\tau\Big|_{\widetilde{\xi}(\tau)}$.*

Before investigating the germs of Jacobi curves, we remind an important result, related to the curve as a whole, namely, to the indices of the Hessian of the end–point

†We identify λ with the corresponding tangent vector to the linear space $T^*_{\pi(\lambda)}M$.

mapping. First of all, the negative (positive) index of $Hess_{(\xi,\lambda_t)}F_t$ is finite if and only if $\frac{\partial^2}{\partial u^2}\lambda_\tau f_\tau\Big|_{\tilde{\xi}(\tau)} > 0\,(<0)$, i.e. when the Jacobi curve $J_{(\xi,\lambda)}$ is monotonically nondecreasing (nonincreasing). Suppose the finiteness condition is satisfied, and let $\overline{J}_{(\xi,\lambda)}$ be the closed curve in $L(\Sigma_\lambda)$, obtained by adding to $J_{(\xi,\lambda)}$ of an arbitrary nondecreasing (nonincreasing) simple curve connecting $J_{(\xi,\lambda)}(t)$ with $J_{(\xi,\lambda)}(0)$.† Then,

$$\pm ind_{\mp}\, Hess_{(\xi,\lambda_t)}F \pm rank\, F'_{0t}\big|_\xi = Ind\,\overline{J}_{(\xi,\lambda)}, \tag{4.2}$$

where $dim\,U \le rank\, F'_{0t}\Big|_\xi \le n-1$. Here, Ind is the Maslov index of a closed curve on a Lagrangian Grassmannian, cf. Introduction. Details and proofs could be found in [1,4].

Now we turn to the geometry of germs of a curve on the Lagrangian Grassmannian $L(\Sigma)$ of a given $2(n-1)$–dimensional symplectic space Σ.

Lemma. *Take an arbitrary $\Lambda_0 \in L(\Sigma)$. The set*

$$\Lambda_0^{\pitchfork} = \left\{\Lambda \in L(\Sigma)\Big|\ \Lambda_0 \bigcap \Lambda = 0\right\}$$

can be given invariantly the structure of an affine space over the vector space of linear selfadjoint mappings of Λ_0^* into Λ_0.

Proof. First, we remark that the set of all $(n-1)$–dimensional subspaces in Σ, transversal to Λ_0, has the structure of an affine space over the space of all linear mappings from Σ/Λ_0 into Λ_0, and this affine structure does not depend on the symplectic structure in Σ. Indeed, if $\Sigma = \Lambda_0 \oplus \Delta$, then the subspace Δ intersects every coset $(z+\Lambda_0) \in \Sigma/\Lambda_0$ exactly at one point. Define the mapping $(\Delta_1 - \Delta_0) : \Sigma/\Lambda_0 \longrightarrow \Lambda_0$ by the formula $(\Delta_1-\Delta_0)(z+\Lambda_0) = \Delta_1 \bigcap (z+\Lambda_0) - \Delta_0 \bigcap (z+\Lambda_0)$. It is easy to see that the introduced operation of difference of two subspaces defines the desired affine structure. Furthermore, the symplectic structure on Σ defines a nondegenerate pairing between Λ_0 and Σ/Λ_0, hence we can identify Σ/Λ_0 with Λ_0^*. Since the subspaces $\Delta_0, \Delta_1 \in \Lambda_0^{\pitchfork}$ are Lagrangian, their difference, $(\Delta_1 - \Delta_0) : \Lambda_0^* \longrightarrow \Lambda_0$ is selfadjoint. By counting dimensions it is easily seen that every selfadjoint mapping from Λ_0^* into Λ_0 is realized as such a difference.

Let $\tau \mapsto \Lambda(\tau)$, $\tau \in [0,t]$, be a smooth curve in $L(\Sigma)$. We call the curve $\Lambda(\cdot)$ *regular* if $\dot\Lambda(\tau)$ is a nondegenerate quadratic form on $\Lambda(\tau)$ for every τ. In this article we restrict to considering only regular curves, postponing more general cases, (highly important and informative), to further publications. The Jacobi curve of a regular extremal is a regular curve if and only if $dim\,U = n-1$, cf. Proposition 4.1.

Let $\Lambda(\cdot)$ be a regular curve, and consider its germ at an arbitrary point t. We have $\Lambda(\tau) \in \Lambda(t)^{\pitchfork}$ for all $\tau \ne t$, sufficiently close to t. More precisely, $\tau \mapsto \Lambda(\tau)$ defines the germ of the curve in the affine space $\Lambda(t)^{\pitchfork}$ with a simple pole at t. We shall give a coordinate representation of this fact.

Let $\Sigma = \left\{(p,q)\Big|\ p,q \in \mathbb{R}^{(n-1)}\right\}$, and the symplectic form has the canonical expression $\sigma((p_1,q_1),(p_2,q_2)) = <p_1,q_2> - <p_2,q_1>$. Without restricting generality, we can

†A curve in a Lagrangian Grassmannian is a family of Lagrangian subspaces of a symplectic space. The curve is said to be simple if there exists a Lagrangian subspace, transversal to all subspaces of the family.

assume that $\Lambda(t) = \left\{(p,0) \middle| \; p \in \mathbb{R}^{(n-1)}\right\}$. Then for every τ close to t, $\Lambda(\tau)$ is represented as $\Lambda(\tau) = \left\{(p, S_\tau p) \middle| \; p \in \mathbb{R}^{(n-1)}\right\}$, where S_τ is a smooth family of symmetric $(n-1)\times(n-1)$–matrices, $S_t = 0$. The regularity of the curve $\Lambda(\cdot)$ means that $det\, \dot{S}_\tau \neq 0$. Every $\Delta \in \Lambda_0^\pitchfork$ has the form $\Delta = \left\{(A_\Delta q, q) \middle| q \in \mathbb{R}^{(n-1)}\right\}$, where A_Δ is a $(n-1)\times(n-1)$ symmetric matrix, and the mapping $\Delta \mapsto A_\Delta$ from $\Lambda_t^\pitchfork$ onto the space of symmetric matrices defines coordinates on $\Lambda_t^\pitchfork$, compatible with the invariant affine structure. In these coordinates the curve $\tau \mapsto \Lambda(\tau) \in \Lambda_t^\pitchfork$ has the expression

$$\tau \mapsto S_\tau^{-1} = \frac{1}{\tau - t}\dot{S}_t^{-1} - \frac{1}{2}\dot{S}_t^{-1}\ddot{S}_t\dot{S}_t^{-1} - \frac{(\tau-t)}{3}\left(((2\dot{S}_t)^{-1}\ddot{S}_t)^{\cdot} - ((2\dot{S}_t)^{-1}\ddot{S}_t)^2\right)\dot{S}_t^{-1} + O((\tau-t)^2). \tag{4.3}$$

To write down the Laurent series of the curve in the affine space we have to use some coordinates, but the coefficients of the series have a clear invariant meaning. Indeed, translation of the affine space by a vector of the corresponding linear space leaves all coefficients of the series unchanged, with the exception of the free term, to which the translating vector is added. Thus, all coefficients of the Laurent series, except the free term, are elements of the linear space, and the free term is an element of the affine space. For a regular curve $\Lambda(\cdot)$ in $L(\Sigma)$ we obtain

$$\Lambda(\tau) \equiv \frac{1}{\tau - t}\Lambda_{-1}(t) + \Lambda_0(t) + \sum_{i=1}^{\infty}(\tau-t)^i\Lambda_i(t),$$

where $\Lambda_0(t) \in \Lambda(t)^\pitchfork$, and $\Lambda_{-1}(t), \Lambda_1(t), \Lambda_2(t), \ldots$ are selfadjoint linear mappings from $\Lambda^*(t)$ into $\Lambda(t)$. Note that $\Lambda_{-1}(t) = \left(\dot{\Lambda}(t)\right)^{-1}$. Put

$$R(t) = -3\Lambda_1(t) \circ (\Lambda_{-1}(t))^{-1} = -3\Lambda_1(t) \circ \dot{\Lambda}(t).$$

Then, $R(t) : \Lambda(t) \longrightarrow \Lambda(t)$ is a linear operator, symmetric with respect to the (pseudo)–Euclidean structure on $\Lambda(t)$, defined by the quadratic form $\dot{\Lambda}(t)$. According to (4.3), the operator $R(t)$ is expressed in coordinates as the Schwarz derivative

$$R(t) = \left((2\dot{S}_t)^{-1}\ddot{S}_t\right)^{\cdot} - \left((2\dot{S}_t)^{-1}\ddot{S}_t\right)^2 \tag{4.4}$$

The curve in $L(\Sigma)$, $t \mapsto \Lambda_0(t)$, is called the *derivative curve* of $\Lambda(\cdot)$. The operator $R(t) : \Lambda(t) \longrightarrow \Lambda(t)$ is called the *curvature operator* of the curve $\Lambda(\cdot)$ at the point t. Straightforward calculations in coordinates show that the derivative curve of a regular curve is smooth, though not necessarily regular. We have $\Sigma = \Lambda(t) \oplus \Lambda_0(t)$. Hence

$$\Lambda_0(t) \cong \Sigma/\Lambda(t) = \Lambda(t)^*, \quad \Lambda(t) \cong \Sigma/\Lambda_0(t) = \Lambda_0(t)^*.$$

Hence, $\dot{\Lambda}(t)$ also is a linear mapping from $\Lambda(t)$ into $\Lambda_0(t)$, and $\dot{\Lambda}_0(t)$ is a linear mapping from $\Lambda_0(t)$ into $\Lambda(t)$. Their calculation in coordinates leads us to the following important identity

$$R(t) = -\dot{\Lambda}_0(t) \circ \dot{\Lambda}(t), \tag{4.5}$$

which could be considered as another equivalent definition of the curvature operator. Since $\dot\Lambda(t)$ is nondegenerate, we have $\dot\Lambda_0(t) = 0 \iff R(t) = 0$. We shall call a regular curve *flat* if it satisfies one of the equivalent conditions, $\dot\Lambda_0(t) \equiv 0 \iff R(t) \equiv 0$. If $R(t) \equiv \varkappa\, id$, $\varkappa \in \mathbb{R}$, then the curve $\Lambda(\cdot)$ is said to have a constant curvature $\varkappa$.

Proposition 4.2. *Germs of two regular curves of constant curvature $\varkappa$ are equivalent iff the signatures of their velocities are equal.*

A regular curve is flat iff its Laurent series in the powers of $(\tau - t)$ has no positive power terms for every t.

Proof. Introduce in Σ coordinates in which the symplectic structure has canonical form and

$$\Sigma = \left\{(p,q)\Big|\ p,q \in \mathbb{R}^{n-1}\right\},\quad \Lambda(t) = \left\{(p,0)\Big|\ p \in \mathbb{R}^{n-1}\right\},\quad \Lambda_0(t) = \left\{(0,q)\Big|\ q \in \mathbb{R}^{n-1}\right\}.$$

Then, $\Lambda(\tau) = \left\{(p, S_\tau p)\Big|\ p \in \mathbb{R}^{n-1}\right\}$, where $S(t) = \ddot S(t) = 0$. If $\Lambda(\cdot)$ is a curve of constant curvature $\varkappa$, then

$$\left((2\dot S_\tau)^{-1}\ddot S_\tau\right)^{\cdot} = \left((2\dot S_\tau)^{-1}\ddot S_\tau\right)^2 + \varkappa\, id.$$

Solving the matrix differential equation with the initial condition at $\tau = t$, we obtain

$$S_\tau = \begin{cases} |2\varkappa|^{-\frac12} tg\left(|2\varkappa|^{\frac12}(\tau - t)\right)\dot S_t, & \varkappa > 0,\\ (\tau - t)\dot S_t, & \varkappa = 0,\\ |2\varkappa|^{-\frac12} th\left(|2\varkappa|^{\frac12}(\tau - t)\right)\dot S_t, & \varkappa < 0. \end{cases} \tag{4.6}$$

Furthermore, under symplectic transformations, which leave fixed $\Lambda(t)$ and $\Lambda_0(t)$, the matrix $\dot S_t$ is transformed as the matrix of a quadratic form, and since it is, by assumption, nondegenerate, the signature is its only invariant.

We give now another equivalent definitions of the derivative curve and the curvature operator, which are more geometric and justify the choice of the term "curvature". We shall use a natural approach to constructing differential geometry of curves on arbitrary homogeneous manifolds. The structure of a homogeneous space, i.e. a transitive action of a given Lie group, singles out a class of "distinguished" curves — the orbits of one–parameter subgroups of the group. Consider an arbitrary germ of a smooth curve, and find a "distinguished" curve which has the same jet of the maximal possible order, as the corresponding jet of the given jet. On the space of "distinguished" curves the group is, in general, nontransitive. The invariants of the approximating "distinguished" curves are the most important differential invariants of the initial germ. This is how the curvature and torsion appear in $\mathbb{R}^3$. Certainly, every homogeneous space brings its own specific features into the general methodology.

One–parameter subgroups of $Sp(\Sigma)$ are the flows of linear *stationary* Hamiltonian systems. They define the family of distinguished curves on $L(\Sigma)$. Elementary calculations imply to the following

Proposition 4.3. *Let $\Lambda(\cdot)$ be a regular curve in $L(\Sigma)$, and h be a quadratic form on Σ, such that the 2–jet of the curve $\tau \mapsto e^{(\tau-t)\vec{h}}(\Lambda(t))$ coincides with the 2–jet of the curve $\Lambda(\cdot)$ at t, and the subspace $\Lambda(t)^{\perp}_{h} = \left\{ y \in \Sigma \middle| \ h(y, \Lambda(t)) = 0 \right\}$ is Lagrangian. Then $\Lambda(t)^{\perp}_{h} = \Lambda_0(t)$. If, in addition, the 3–jet of the curve $\tau \mapsto e^{(\tau-t)\vec{h}}\Lambda(t)$ coincides with the 3–jet of the curve $\Lambda(\cdot)$ at t, then the form h is uniquely defined, where $\dot{\Lambda}(t) = 2h\Big|_{\Lambda(t)}$, $\dot{\Lambda}_0(t) = 2h\Big|_{\Lambda_0(t)}$.*

The curvature operator is defined by the germ of the curve, but it also enables to make conclusions about global properties of the curve.

The points t_0, t_1 are said to be *conjugate* for the curve $\Lambda(\cdot)$ if $\Lambda(t_0) \bigcap \Lambda(t_1) \neq 0$; the number $dim\,(\Lambda(t_0) \bigcap \ \Lambda(t_1))$ is the *multiplicity* of the conjugate pair.

The following Proposition is a direct consequence of elementary facts of symplectic geometry, and we formulate it as a separate assertion for the convenience of references.

Proposition 4.4. *Let $\Lambda : [0, T] \longrightarrow L(\Sigma)$ be a smooth curve, $\dot{\Lambda}(t) > 0 \ \forall t$, $\Lambda(0) \bigcap \Lambda(T) = 0$, and $\overline{\Lambda}(\cdot)$ is a closed curve obtained from $\Lambda(\cdot)$ by adding to it of a regular simple non-decreasing curve. Then every $t \in [0, T]$ is conjugate only to a finite number of points, and*

$$Ind\,\overline{\Lambda}(\cdot) = n - 1 + \sum_{0<t<T} dim(\Lambda(0) \bigcap \Lambda(t)) = n - 1 + \sum_{0<t<T} dim(\Lambda(t) \bigcap \Lambda(T)).$$

The last assertion of this section is the "comparison theorem", which estimates the index through the curvature.

Theorem. *Assume $\Lambda(\cdot)$ is a smooth curve in $L(\Sigma)$ and $\dot{\Lambda}(t) > 0$ for $\forall t$. If $R(t) \leq C\,id$ for some $C \geq 0$ and $\forall t$, then $|t_1 - t_0| \geq \frac{\pi}{\sqrt{C}}$ for every pair of conjugate points t_0, t_1. In particular, if $R(t) \leq 0$, then there are no conjugate points.*

Assume $tr\,R(t) \geq (n-1)c$ for some $c > 0$ and for $\forall t$, then for arbitrary $t_0 \leq t$ the interval $[t, t + \frac{\pi}{\sqrt{c}}]$ contains a point conjugate to t_0. Both estimates are exact.

Proof. We start with the case $R(t) \leq 0$. The absence of conjugate points under this assumption easily follows from the standard facts about Lagrangian Grassmannian contained, for example, in [7]. In this case, local coordinates exist from the standard affine atlas in which $\Lambda(t_0)$ is represented by the zero matrix, and $\Lambda_0(t_0)$ — by a symmetric positive matrix. Let S_t be the matrix corresponding to $\Lambda(t)$, S^0_t be the matrix corresponding to $\Lambda_0(t)$. Then $\dot{S}_t > 0$ and $det(S^0_t - S_t) \neq 0$, since $\Lambda(\cdot)$ is monotonically increasing and $\Lambda_0(t) \bigcap \Lambda(t) = 0$. The operator $R(t)$ is represented by the matrix $(S^0_t - S_t)^{-1}\dot{S}^0_t(S^0_t - S_t)^{-1}\dot{S}_t$. Hence, $\dot{S}^0_t \leq 0$. The given relations hold until $\Lambda(t)$ and $\Lambda_0(t)$ remain in the coordinate neighborhood. However, the relations imply that $0 \leq S_t < S^0_t \leq S^0_{t_0}$, hence, $\Lambda(t)$ and $\Lambda_0(t)$ do not leave the coordinate neighborhood at all.

Now assume that $R(t) \leq C\,id$. We shall make use of the following formula, a direct consequence of (4.4). Assume $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ is a smooth function, $\dot{\varphi}(\tau) > 0$, and $R_{\varphi}(\tau)$ is the curvature operator of the curve $\tau \mapsto \Lambda(\varphi(\tau))$. Then

$$R_\varphi(\tau) = \dot{\varphi}^2(\tau) R(\varphi(\tau)) + \left((\frac{\ddot{\varphi}}{2\dot{\varphi}})^{\cdot}(t) - (\frac{\ddot{\varphi}(t)}{2\dot{\varphi}(t)})^2 \right) id \tag{4.7}$$

Put

$$\varphi_{\bar{t}}(\tau) = \frac{1}{\sqrt{C}} \left(arctg(\sqrt{C}\tau) + \frac{\pi}{2} \right) + \bar{t}, \quad \varphi_{\bar{t}}(\mathbb{R}) = (\bar{t}, \bar{t} + \frac{\pi}{\sqrt{C}}).$$

We obtain,

$$R_{\varphi_{\bar{t}}}(\tau) = \frac{1}{(C\tau^2+1)^2} \left(R(\varphi_{\bar{t}}(\tau)) - C\, id \right) \leq 0.$$

Hence the curve $\tau \mapsto \Lambda(\varphi_{\bar{t}}(\tau))$ has no conjugate points on the interval $(\bar{t}, \bar{t} + \frac{\pi}{\sqrt{C}})$.

Assume now that $tr\, R(t) \geq (n-1)c$. We shall prove that, if $\Delta \bigcap \Lambda(\tau) = 0$ for some $\Delta \in L(\Sigma)$ and $\forall \tau \in [t, \bar{t}]$, then $\bar{t} - t < \frac{\pi}{\sqrt{c}}$.

Indeed, if such a Δ exists, then $\Lambda\Big|_{[t,\bar{t}]}$ is completely contained in a coordinate neighborhood, therefore the curvature operator $R(\tau)$ is defined by the formula (4.4). Put $Z(\tau) = (2\dot{S}_\tau)^{-1}\ddot{S}_\tau$, $z(\tau) = tr\, Z(\tau)$, $\tau \in [\tau, t]$. Then,

$$\dot{Z}(\tau) = Z^2(\tau) + R(\tau), \quad \dot{z}(\tau) = tr\, Z^2(\tau) + tr\, R(\tau).$$

Since for an arbitrary symmetric $(n-1)\times(n-1)$–matrix A we have $tr\, A^2 \geq \frac{1}{n-1}(tr\, A)^2$, the inequality $\dot{z} \geq \dfrac{z^2}{n-1} + (n-1)c$ holds. Hence, $z(\tau) \geq x(\tau)$, $t \leq \tau \leq \bar{t}$, where $x(\cdot)$ is a solution of the equation

$$\dot{x} = \frac{x^2}{n-1} + (n-1)c,$$

i.e. $x(\tau) = (n-1)\sqrt{c}\, tg(\sqrt{c}(\tau - \tau_0))$.

The function $z(\cdot)$, together with $x(\cdot)$, are bounded on the interval $[t, \bar{t}]$. Hence, $\bar{t} - t < \frac{\pi}{\sqrt{c}}$.

To verify that the estimates are exact, it is enough to consider curves of constant curvature.

Applying the theorem to the Jacobi curve $J_{(\xi,\lambda)}$, one can obtain explicit estimates of the index of $Hess_{(\xi,\lambda_t)} f$ through the curvature of the Jacobi curve in case of a finite index. Indeed, formula (4.2) and Proposition 4.4 imply the following form of the classical Morse formula,

$$ind\, Hess_{(\xi,\lambda_t)} f = \sum_{0<\tau<t} dim(J_{(\xi,\lambda)}(\tau) \bigcap J_{(\xi,\lambda)}(t)).$$

In other words, the more conjugate points, the bigger the index. If there are no conjugate points at all, then $Hess_{(\xi,\lambda_t)} f$ is sign–definite.

§5. Canonical Connections of Hamiltonian Systems and of Differential Equations of Second Order.

1. Nonlinear connections on fibre bundles. Assume a smooth (locally trivial) fibre bundle $E = \bigcup_{x\in M} E_x$ over M is given, with the canonical projection $\pi : E \longrightarrow M$. In the tangent bundle TE the *"vertical"* subbundle is defined,

$$T^{ver}E = \bigcup_{e\in E} T_e E_{\pi(e)} \subset TE = \bigcup_{e\in E} T_e E, \quad ker\, \pi_* = T^{ver}E, \ im\, \pi_* = TM.$$

Any direct complement to $T^{ver}E$ in TE will be called a *(nonlinear) connection* on E, vector fields on TE with values in the direct complement will be called *horizontal*, vector fields with values in $T^{ver}E$ will be called *vertical.*

Assume a connection is fixed on E. Then, for every $e \in E$, the restriction $\pi_*\Big|_{T_e E}$ defines a one–to–one mapping of the space of horizontal tangent vectors at e onto $T_{\pi(e)}M$. Hence there exists a uniquely defined mapping, $X \mapsto \nabla_X$, $X \in Vect\, M$, of the space of vector fields on M into the space of horizontal fields on E, satisfying the relation $\pi_*\nabla_X = X \ \forall X \in Vect\, M$. Evidently, the correspondence $X \mapsto \nabla_X$ is a $C^\infty(M)$–linear (or tensorial) mapping:

$$\nabla_{X+Y} = \nabla_X + \nabla_Y, \ \nabla_{aX} = (a\circ\pi)\nabla_X \quad \forall a \in C^\infty(M).$$

For every vertical field V, the commutator $[\nabla_X, V]$ is a vertical field, and the mapping $X \mapsto [\nabla_X, V]$ is $C^\infty(M)$–linear (tensorial). In particular, the restriction $[\nabla_X, V]\Big|_{E_x}$ is uniquely defined by $X(x)$ and V. To emphasize this remark explicitly, as well as for some technical reasons which will be clear below, we omit the brackets in the commutator $[\nabla_X, V]$, and call the expression $\nabla_X V \stackrel{\text{def}}{=} [\nabla_X, V]$ the *covariant derivative* of the (vertical) field V along X. For every $v \in T_x M$ the covariant derivative, $\nabla_v V \in Vect\, E_x$, is correctly defined.

Every horizontal vector field is represented as $e \mapsto \nabla_{\xi(e)}$, $e \in E$, where $\xi(e) \in T_{\pi(e)}M$. The restriction of the mapping $e \mapsto \xi(e)$ to E_x is a vector–function with values in the vector space $T_x M$. Since we can act on every smooth vector–function by an arbitrary vector field, defined on the domain of definition of the function, by differentiating the vector–function along the corresponding directions, the action of vertical fields on the mapping $e \mapsto \xi(e)$ is correctly defined. The following evident, though very useful, formula gives the decomposition of the commutator of a horizontal and a vertical field into the horizontal and vertical components,

$$[\nabla_\xi, V](e) = (\nabla_{\xi(e)}V - \nabla_{V\xi})(e). \tag{5.1}$$

The commutator of vertical fields is vertical, at the same time the commutator of horizontal fields might not be horizontal. The description of the vertical component of the commutator of two horizontal vector fields leads us to the important notion of the curvature of a connection. Put

$$R^\nabla(X, Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}, \ \forall X, Y \in Vect\, M.$$

Evidently, the field $R^\nabla(X,Y)$ is vertical, and the mapping $X\wedge Y\mapsto R^\nabla(X,Y)$ is tensorial, (is a homomorphism of a $C^\infty(M)$–module $Vect\,M\wedge Vect\,M$ into the module of vertical vector fields). In particular, $R^\nabla(X,Y)\Big|_{E_x}$ is depending only on $X(x),Y(x)$, hence for $\forall v_1,v_2\in T_xM$ the field $R^\nabla(v_1,v_2)\in Vect\,E_x$ is correctly defined. Now let $e\mapsto\nabla_{\xi_i(e)},\ i=1,2$, be arbitrary horizontal fields. It is easy to show that the field $e\mapsto R^\nabla(\xi_1(e),\xi_2(e))$ is the vertical component of the field $[\nabla_{\xi_1},\nabla_{\xi_2}]$, i.e. $[\nabla_{\xi_1},\nabla_{\xi_2}]-R^\nabla(\xi_1,\xi_2)$ is a horizontal field.

For a nonstationary field X_τ on M the flow $t\mapsto\overrightarrow{exp}\int\limits_{t_0}^{t}\nabla_{X_\tau}d\tau$ consists of fibrewise diffeomorphisms of the bundle E. Let $x(t)=x_0\overrightarrow{exp}\int\limits_{t_0}^{t}\nabla_{X_\tau}d\tau$, hence $t\mapsto x(t)$ is a trajectory of the flow on M, defined by the field X_τ. Since the mapping $X\mapsto\nabla_X$ is tensorial, the diffeomorphism

$$\overrightarrow{exp}\int\limits_{t_0}^{t_1}\nabla_{X_\tau}d\tau\Big|_{E_x(t_0)}:E_x(t_0)\longrightarrow E_x(t_1)$$

depends only on the connection and on the curve $x(\tau),0\le\tau\le t$, and is independent on the values of the field X_τ off the curve $x(\tau)$. This diffeomorphism is called the *parallel translation along the curve* $x(\tau),0\le\tau\le t$.

Assume that E is a linear bundle. The connection is called *linear* if the fields ∇_X preserve the space of functions, linear on fibres. If the connection is linear, then the parallel translation $\overrightarrow{exp}\int\limits_{t_0}^{t}\nabla_{X_\tau}d\tau\Big|_{E_{x(t_0)}}$ is a linear mapping. We shall consider below nonlinear connections on linear bundles.

So far we were concerned only with main definitions related to connections, and all assertions were almost trivial. They could be checked by introducing local coordinates, or algebraically, identifying vector fields on M or E with corresponding derivations of algebras $C^\infty(M)$ or $C^\infty(E)$, vertical fields on E — with the annihilator of the subalgebra in $C^\infty(E)$ of functions, constant on fibres.

2. Connections associated with Hamiltonians. Assume now that the bundle E is not arbitrary, but rather a region in T^*M, $E_x=E\cap T^*_xM,\ x\in M$. Denote by $h:E\longrightarrow\mathbb{R}$ a smooth Hamiltonian on E, by Dh — the *vertical differential* of h,

$$D_\lambda h\stackrel{\text{def}}{=}d_\lambda(h\Big|_{E_x}),\ \lambda\in E_x,x\in M.$$

Thus $D_\lambda h\in(T^*_xM)^*=T_xM$, hence $Dh:E\longrightarrow TM$ is a smooth fibrewise mapping. Note that $D_\lambda h=\pi_*\vec{h}(\lambda)$. We assume that $D_\lambda h\ne0,\ \lambda\in E$. Since E_x is a region in a linear space, the *second vertical derivative* is also well defined, $D^2_\lambda h=d^2_\lambda(h\Big|_{E_x})$, and is a quadratic form on T^*M.

To every λ there corresponds a curve J_λ in the Lagrangian Grassmannian $L(T_\lambda(T^*M))$, according to the formula

$$J_\lambda(t)=(e^{-t\vec{h}})_*T_{\lambda_t}(T^*_{\pi(\lambda_t)}M),\ \lambda_t=\lambda e^{t\vec{h}}.$$

If h is a composition of the master–Hamiltonian of a control system with a function of a real variable, for example, some power of the master–Hamiltonian, then the curve

$$t \mapsto J_\lambda(t) \bigcap ker\, d_\lambda h \in L\left(\vec{h}(\lambda)^{\angle} / \vec{h}(\lambda)\right)$$

is the Jacobi curve of the corresponding geodesic, cf. §4. Let $\pi(\lambda) = x$, then $J_\lambda(0) = T_\lambda(E_x) = T_x^*M$. It is easily seen that $\dot{J}_\lambda(0) = D_\lambda^2 h$.

As in §4, we restrict ourself to regular curves, J_λ, postponing more general cases, (quite important and interesting for optimal problems), until further publications. Thus, we suppose that $D_\lambda^2 h$ is a nondegenerate quadratic form on T_x^*M.

Remark. At first sight, the nondegeneracy requirement might seem too excessive. For example, the master–Hamiltonian, being positive–definite of first degree, does not satisfy it. But, if the master–Hamiltonian is not zero at λ and generates a regular Jacobi curve, then the square of the master–Hamiltonian satisfies the condition.

To every regular curve $t \mapsto J_\lambda(t)$ corresponds the derivative curve $t \mapsto J_{\lambda 0}(t)$. According to its definition, cf. §4,

$$J_{\lambda 0} \in L(T_\lambda(T^*M)), \quad T_\lambda(T^*M) = T_\lambda(E_x) \oplus J_{\lambda 0}(0).$$

Evidently, $J_{\lambda 0}(0)$ is smooth in λ. We call the Lagrangian bundle $J_{\lambda 0}(0), \lambda \in E$, the *canonical connection on E, associated with the Hamiltonian h.*

Let ∇_X be a horizontal field for the canonical connection, $\pi_* \nabla_X = X,\ X \in Vect\, M$.

Lemma 5.1. *Assume that the restriction $h\Big|_{E_x}$ of the Hamiltonian h to an arbitrary fibre E_x is a positively homogeneous function of degree $r+1$, $r \neq 0, -1$. Then $\vec{h} = \nabla_{D_\lambda h}$, $\lambda \in E$, hence the field $\vec{h}$ is horizontal.*

Proof. We identify $\lambda \in E_x \subset T_x^*M$ with the corresponding tangent vector from $T_\lambda E_x = T_x^*M$. The homogeneity of the Hamiltonian implies the identity

$$(s\lambda)e^{t\vec{h}} = s\left(\lambda e^{s^r t \vec{h}}\right) \quad \forall s > 0,$$

from which the relations follow, $\left(\lambda - rt\vec{h}(\lambda)\right) \in J_\lambda(t) \subset T_\lambda(T^*M)$. Moreover, since the Hamiltonian flow preserves h, and $\vec{h}(\lambda)^{\angle} = ker\, d_\lambda h$, we obtain,

$$J_\lambda(t) = \mathbb{R}\left(\lambda - rt\vec{h}(\lambda)\right) \oplus J_\lambda(t) \bigcap \vec{h}(\lambda)^{\angle}. \tag{5.2}$$

From here we conclude that $\vec{h}(\lambda) \in J_{\lambda 0}(t)$, hence $\vec{h}$ is a horizontal field. To complete the prove, we remark that $\pi_* \vec{h} = D\, h$.

Denote by R_{J_λ} the curvature operator of the curve $\lambda \mapsto J_\lambda(t)$ for $t = 0$. Since $J_\lambda(0) = T_\lambda E_x = T_x^*M$, where $\lambda \in E_x$, $R_{J_\lambda} : T_x^*M \longrightarrow T_x^*M$ is a linear operator. By R^∇ we denoted the curvature of the canonical connection. Despite of completely different ways of definition of these two curvatures, they are intimately connected and the use of the same term "curvature" in both cases is completely justified.

Theorem 5.1. *Under the conditions of Lemma 5.1 the following identity holds,*

$$R_{J_\lambda} l = R^\nabla(D_\lambda h, l \rfloor D_\lambda^2 h), \ \forall \lambda \in E_x, \ x \in M, \ l \in T_x^*M.$$

Proof. Let $x_\lambda(t)$ be the projection on M of the point $\lambda e^{t\vec{h}}$, and $\bar{l}$ — a vertical field, which has a restriction to $E_{x_\lambda(t)}$ coinciding with the parallel translation of the constant field l on E_x along the curve $x_(\cdot)$ for $\forall t$. Then,

$$\nabla_{Dh}\bar{l}\Big|_{\lambda e^{t\vec{h}}} = 0. \tag{5.3}$$

Since the action of $\left(e^{-t\vec{h}}\right)_*$ on the vector fields coincides with the the action of $e^{t\,ad\vec{h}}$, cf. Introduction, we obtain $J_\lambda(t) = \left\{ \left(e^{t\,ad\vec{h}}\bar{l}\right)(\lambda) \Big| \ l \in T_x^*M \right\}$. Let $\nabla_{A_t l}$ be the horizontal component of the vector $\left(e^{t\,ad\vec{h}}\bar{l}\right)(\lambda)$, and $B_t l$ be its vertical component, so that $A_t : T_x^*M \longrightarrow T_xM$, $B_t : T_x^*M \longrightarrow T_x^*M$, are linear mappings, where

$$J_\lambda(t) = \left\{ B_t l + \nabla_{A_t l} \Big| \ l \in T_x^*M \right\}, \ A_0 = 0, \ B_0 = id.$$

Thus the germ at zero of the curve J_λ is represented by the matrix curve $t \mapsto S_t = A_t B_t^{-1}$, $J_\lambda(t) = \left\{ l + \nabla_{S_t l} \Big| \ l \in T_x^*M \right\}$, and the curvature operator R_{J_λ} has, according to (4.4), the form

$$R_{J_\lambda} = \left((2\dot{S}_0)^{-1}\ddot{S}_0\right)^{\cdot} - \left((2\dot{S}_0)^{-1}\ddot{S}_0\right)^2. \tag{5.4}$$

Formulas (5.1) and (5.3) imply

$$\frac{d}{dt} e^{t\,ad\vec{h}}\bar{l}(\lambda) = e^{t\,ad\vec{h}}[\vec{h}, \bar{l}](\lambda) = e^{t\,ad\vec{h}}[\nabla_{Dh}, \bar{l}](\lambda) = -\left(e^{t\,ad\vec{h}}\nabla_{\bar{l}\rfloor D^2 h}\right)(\lambda).$$

Hence $\dot{A}_0 l = -\nabla_{l\rfloor D_\lambda^2 h}$, $\dot{B}_0 = 0$. Furthermore, the derivative curve of J_λ at $t = 0$ has the form, cf. (4.3), $J_{\lambda 0} = \left\{ -\frac{1}{2}\dot{S}_0^{-1}\ddot{S}_0\dot{S}_0^{-1}v + \nabla_v \Big| \ v \in T_xM \right\}$. At the same time, according to the definition of the canonical connection, $J_{\lambda 0}$ consists of horizontal vectors. Therefore $\ddot{S}_0 = 0$, and the formula (5.4) takes the form $R_{J_\lambda} = \frac{1}{2}\dot{S}_0^{-1}\dddot{S}_0$. Furthermore, since $A_0 = 0$ and $\ddot{S}_0 = 0$, we have $\ddot{A}_0 = 0$. Hence the vector field

$$\frac{d^2}{dt^2} e^{t\,ad\vec{h}}\bar{l}\Big|_{t=0} = -\frac{d}{dt} e^{t\,ad\vec{h}}\nabla_{\bar{l}\rfloor D^2 h} = [\nabla_{\bar{l}\rfloor D^2 h}, \vec{h}] = [\nabla_{\bar{l}\rfloor D^2 h}, \nabla_{Dh}]$$

is vertical at λ. Since the vertical component of the commutator of two horizontal fields is the curvature, we have

$$[\nabla_{\bar{l}\rfloor D^2h}, \nabla_{Dh}](\lambda) = R^{\nabla}(l\rfloor D^2_\lambda h, D_\lambda h).$$

Furthermore, the point λ is indistinguishable from any other point of the form $\lambda e^{t\vec{h}}$, hence the last identity is satisfied for all such points, we have only to substitute l by the value of the field $\bar{l}$ at $\lambda e^{t\vec{h}}$. We obtain,

$$\frac{d^2}{dt^2} e^{t\, ad\vec{h}}\,\bar{l}(\lambda) = e^{t\, ad\vec{h}} R^{\nabla}(\bar{l}\rfloor D^2h, Dh)(\lambda).$$

Thus $\ddot{B}_0 l = R^{\nabla}(l\rfloor D^2_\lambda h, D_\lambda h)$, and $\dddot{A}_0 l$ is the horizontal component of the vector

$$[\vec{h}, R^{\nabla}(\bar{l}\rfloor D^2h, Dh)](\lambda) = [\nabla_{Dh}, R^{\nabla}(\bar{l}\rfloor D^2h, Dh)](\lambda).$$

According to formula (5.1), we have $\dddot{A}_0 l = R^{\nabla}(D_\lambda h, l\rfloor D^2_\lambda h)\rfloor D^2_\lambda h$. Collecting the obtained formulas together, we can write,

$$\dddot{S}_0 l = \dddot{A}_0 l - 3\dot{A}_0 \ddot{B}_0 l = 2R^{\nabla}(l\rfloor D^2_\lambda h, D_\lambda h)\rfloor D^2_\lambda h, \quad R_{J_\lambda} l = \frac{1}{2}\dot{A}_0^{-1}\dddot{S}_0 l = R^{\nabla}(D_\lambda h, l\rfloor D^2_\lambda h).$$

3. Connections associated with second order differential equations. We have considered above canonical connections associated with Hamiltonian systems, a natural class of differential equations on the cotangent bundle. Now we describe connections with similar properties for differential equations of the second order, a natural class of differential equations on the tangent bundle.

Assume E is a region in TM, $E_x = E\bigcap T_xM$, $x \in M$. We shall say that a vector field Z on E is a *differential equation of the second order* or that it defines a differential equation of the second order, if $\pi_* Z(v) = v\ \forall v \in E$. For $s \in \mathbb{R}$ denote by $s_* : T(TM) \longrightarrow T(TM)$ the differential of the homothety $v \mapsto sv$, $v \in TM$. The field Z, defining the differential equation of the second order, is called a *spray* if $Z(sv) = s_* Z(v)\ \forall v \in E, s \in \mathbb{R}$ such that $sv \in E$. In local coordinates, the differential equation of the second order, defined by a spray, has the form $\ddot{x} = \varphi(x, \dot{x})$, where φ is homogeneous in $\dot{x}$ of degree 2. We should note that no other degrees of homogeneity are preserved under the coordinate change on M.

For every differential equation of the second order Z and every $v \in E$ we define a curve I_v in the Grassmannian $G_n(T_v(TM))$ of all n–dimensional subspaces in $T_v(TM))$ by the formula

$$I_v(t) = \left(e^{-tZ}\right)_* T_{v_t} E_{\pi(v_t)}, \ v_t = ve^{tZ}.$$

Before moving further we shall make few remarks about curves in the Grassmannian $G_n(\mathbb{R}^{2n})$. So far, we considered curves only in Lagrangian Grassmannians. Definitions and properties mentioned below are similar to those in the Lagrangian Grassmannian, and are proved even easier since no additional symplectic structure should be taken in consideration.

Put

$$K_0 \in G_n(\mathbb{R}^{2n}),\ K_0^{\pitchfork} = \left\{ K \in G_n(\mathbb{R}^{2n}) \middle| \ K_0 \bigcap K = 0 \right\}.$$

Then $K_0^{\pitchfork}$ has a natural structure, (independent on the choice of a basis in $\mathbb{R}^{2n}$), of an affine space over the vector space $Hom(\mathbb{R}^{2n}/K_0, K_0)$. We already described this affine structure in the proof of the Lemma in §4. Furthermore, the tangent space $T_{K_0}G_n(\mathbb{R}^{2n})$ is naturally identified with the space $Hom(K_0, \mathbb{R}^{2n}/K_0)$ in the following way. Assume $t \mapsto K_t$ is a smooth curve in $G_n(\mathbb{R}^{2n})$. We correspond to the tangent vector $\dot{K}_0 = \frac{d}{dt}K_t\Big|_{t=0}$ the mapping $k_0 \mapsto \frac{d}{dt}k_t\Big|_{t=0} + K_0$, where $k_t \in K_t$. It is easy to show that this mapping depends only on $\dot{K}_0$ and does not depend on the choice of the curves K_t and k_t.

The curve $\tau \mapsto K_\tau$ in $G_n(\mathbb{R}^{2n})$ is *regular* if its velocities $\frac{dK_\tau}{d\tau}$ are regular linear mappings from K_τ into $\mathbb{R}^{2n}/K_\tau\ \forall\tau$. It is easy to show that the curve I_v in $G_n(T_v(TM))$ is nondegenerate. In particular, for $\tau = 0$ we have $I_v(0) = I_v(T_xM) = ker\,\pi_*\Big|_{T_v(TM)}$, where $x = \pi(v)$. Identifying the spaces $T_v(T_xM)$ and $T_v(TM)/T_v(T_xM) \approx \pi_*T_v(TM)$ with T_xM, we obtain $\frac{d}{d\tau}I_v(0) = id$.

The germ at t of a regular curve $\tau \mapsto K(\tau)$ in $G_n(\mathbb{R}^{2n})$ defines a curve in the affine space $K(t)^{\pitchfork}$ with a simple pole at $\tau = t$. In other words,

$$K(\tau) \approx \frac{1}{\tau - t}K_{-1}(t) + K_0(t) + \sum_{i=1}^{\infty}(\tau - t)^i K_i(t), \quad K_0(t) \in K(t)^{\pitchfork},$$
$$K_i(t) \in Hom(\mathbb{R}^{2n}/K(t), K(t)),\ i \neq 0; \quad K_{-1}(t) = (\dot{K}(t))^{-1}.$$

Put $R(t) = -3K_1(t)\dot{K}(t)$. The curve $t \mapsto K_0(t)$ in $G_n(\mathbb{R}^{2n})$ is called the *derivative curve* of $K(\cdot)$. The operator $R(t) : K(t) \longrightarrow K(t)$ is called the *curvature operator* of the curve $K(\cdot)$ at t. In local coordinates the curvature operator is represented, as in the Lagrangian case, by the matrix Schwarz derivative (4.4), with matrices not necessarily symmetric.

Let $t \mapsto I_{v0}(t)$ be the derivative curve of I_v. We have $T_v(TM) = I_vE_x + I_{v0}(0)$, and $I_{v0}(0)$ smoothly depends on v. The subbundle in TE with the fibres $I_{v0}(0), v \in E$, is called the *canonical connection associated with the field Z*. Below, in this subsection, we assume that the symbol ∇_X denotes the horizontal field for the defined canonical connection, such that $\pi_*\nabla_X = X,\ X \in Vect\,M$.

Lemma 5.2. *If Z is a spray, then Z is a horizontal field for the canonical connection, $Z(v) = \nabla_v\ \forall v \in E$.*

Proof. We identify the vector $v \in E_x \subset T_xM$ with the corresponding vertical tangent vector in $T_vE_x = T_xM$. Since Z is a spray, we have $(sv)e^{tZ} = s(ve^{stZ}),\ s \in \mathbb{R}$. From here, we obtain

$$(v - tZ(v)) \in I_v(t) \subset T_v(TM). \tag{5.5}$$

The subspace $I_v(t)$ is represented as $I_v(t) = \left\{ l + \nabla_{S_t l} \middle| \ l \in T_vE_x = T_xM \right\}$, where $S_t : T_xM \longrightarrow T_xM$ is a linear operator smooth in t, $S_0 = 0$. Moreover, $\dot{S}_0 = id$.

From (4.3) it follows that the value at $t = 0$ of the derivative curve of I_v has the form

$$I_{v0} = \left\{ -\frac{1}{2}\ddot{S}_0 l + \nabla_l \middle| \; l \in T_x M \right\}.$$

Since I_{v0} consists of horizontal vectors, we have $\ddot{S}_0 = 0$. Let $Z(v) = l_0 + \nabla_v$, i.e., l_0 is the vertical component of the vector $Z(v)$. Then according to (5.5), $tv = S_t(tl_0 - v)$. Differentiating 2 times in t yields $\ddot{S}_0 v = l_0$. Hence $l_0 = 0,\ Z(v) = \nabla_v$.

Denote by R_{I_v} the curvature operator of the curve $t \mapsto I_v(t)$ at $t = 0$. Since $I_v(0) = T_x M$, where $x = \pi(v)$, the mapping $R_{I_v} : T_x M \longrightarrow T_x M$ is a linear operator. The symbol R^∇ denotes the curvature of the canonical connection associated with the field Z.

Theorem 5.2. *If Z is a spray, then*

$$R_{I_v} l = R^\nabla(v, l) \quad \forall v \in E_x,\ x \in M,\ l \in T_x M.$$

The proof is a repeating of the proof of Theorem 5.1 with corresponding simplifications.

4. Linear Connections. Here we shall consider in more detail linear connections on the bundles TM and T^*M. Assume a linear connection is given on TM, hence for every $X \in Vect\, M$ a horizontal vector field, ∇_X, on TM is given, which preserves the space of functions, linear on the fibres of TM. In this case, the diffeomorphism $e^{-t\nabla_X}$ is a *linear* mapping of the fibre $T_{xe^{tX}}M$ onto the fibre $T_x M$ for $\forall x \in M,\ t \in \mathbb{R}$. Considering the adjoint linear mappings

$$\left(e^{-t\nabla_X} \Big|_{T_{xe^{tX}}M} \right)^* : T_x^* M \longrightarrow T^*_{xe^{tX}} M,$$

we obtain the *"adjoint"* flow on the bundle T^*M. The generating vector field for this flow will be denoted by ∇_X^*. The mapping $X \mapsto \nabla_X^*,\ X \in Vect\, M$, defines a linear connection on T^*M, the *adjoint* to the connection $X \mapsto \nabla_X$. Evidently, we could start with an arbitrary linear connection on T^*M, and define the adjoint connection on TM, obtaining the involution $\nabla_X^{**} = \nabla_X$. For appropriately chosen notation the expressions for ∇_X and ∇_X^* are indistinguishable.

Indeed, every $Y \in Vect\, M$ could be considered as a cross–section of the vector bundle TM, and as such could be identified with the vertical vector field on TM, constant on fibres. Hence, the covariant derivative of the field Y along the field X is defined, denoted by $\nabla_X Y$, which is vertical and constant on fibres, $\nabla_X Y \in Vect\, M$. Furthermore, Y could be considered as a scalar–valued function on T^*M, linear on fibres. The image $\nabla_X^* Y$ of this function under the action of the vector field $\nabla_X^* \in Vect\, T^*M$ is again linear on fibres, in other words, $\nabla_X^* Y \in Vect\, M$. It is easily proved that

$$\nabla_X Y = \nabla_X^* Y. \tag{5.6}$$

For a linear connection on TM we define in a usual way the *torsion*

$$T^\nabla(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y],\ T^\nabla(X, Y) \in Vect\, M.$$

Lemma 5.3. *The following identity holds,*

$$T^{\nabla}(X,Y) = \sigma(\nabla^*_X, \nabla^*_Y),\ \ X, Y \in Vect\, M$$

where σ is the canonical symplectic structure on T^*M.

Proof. Let θ be the canonical 1–form on T^*M, $\sigma = d\theta$. Since $\pi_* \nabla^*_X = X$ we have $X =< \theta, \nabla^*_X >$, where X in the left–hand side of the last identity is considered as a function on T^*M, linear on fibres. Taking into account that θ vanishes on vertical fields, we obtain

$$T^{\nabla}(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] = \nabla^*_X < \theta, \nabla^*_Y > - \nabla^*_Y < \theta, \nabla^*_X > - < \theta, \nabla^*_{[X,Y]} >= \sigma(\nabla^*_X, \nabla^*_Y) + < \theta, R(X,Y) >= \sigma(\nabla^*_X, \nabla^*_Y).$$

Corollary. *Connection ∇ has a zero torsion iff ∇^* defines a Lagrangian subbundle in $T(T^*M)$.*

Every linear connection on TM defines a spray Z according to the formula $Z(v) = \nabla_v(v)\ \forall v \in TM$. The trajectories of this spray are called *geodesics of the connection* ∇. Different connections can define identical sprays, but among them there exists a unique connection with vanishing torsion. Not every spray can be obtained in this way, only the sprays which are quadratic on fibres.*

Proposition 5.1. *The canonical connection associated with a spray, quadratic on fibres, is linear and has a vanishing torsion.*

Proof. Fix an arbitrary point $x_0 \in M$ and local coordinates $x = (x^1, \dots, x^n)$ in the neighborhood of x_0, such that the corresponding differential equation of the second order in these coordinates has the form $\ddot{x} = \varphi(x, \dot{x}),\ \varphi(x_0, \dot{x}) = 0\ \forall \dot{x}$. We obtain,

$$I_{(x_0,\dot{x})0} = span\left\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right\}\ \forall \dot{x}.$$

In other words, for the canonical connection we have $\nabla_{\frac{\partial}{\partial x^i}} = \dfrac{\partial}{\partial x^i}$, hence ∇ is linear and has a vanishing torsion.

Thus the canonical connection associated with a spray, quadratic on fibres, is linear with vanishing torsion, and its geodesics are the trajectories of the spray. In geometry, the geodesics of a connection on TM are considered as the "straightest" lines, whereas the extremals of a variational problem as the "shortest" lines. We see that the trajectories of

*In local coordinates, the differential equation of the second order defined by a spray has the form $\ddot{x} = \varphi(x, \dot{x})$, and the property of φ to be quadratic in $\dot{x}$ is independent on the choice of the coordinates in M. This is a natural class of sprays, since for a spray, quadratic on a given fibre $T_{x_0}M$, the corresponding differential equation of the second order has, in appropriate coordinates, the form $\ddot{x} = \varphi(x, \dot{x})$, where $\varphi(x_0, \dot{x}) = 0\ \forall \dot{x}$.

every spray, not necessarily linear on fibres, could be considered as a system of " straightest" lines, corresponding to the canonical connection on TM, generally not linear. At the same time, the shortest lines, being the trajectories of the master–Hamiltonian, generate the canonical connection on T^*M.

According to the classical Riemannian geometry, the "shortest" lines for the Riemannian variational problem are also the "straightest" lines for the Levi–Civita connection. Therefore, it is natural to expect that the Levi–Civita connection on TM is conjugate to the canonical connection on T^*M, associated with the corresponding Hamiltonian. Indeed, we have the following

Proposition 5.2. *Assume $Q : TM \longrightarrow T^*M$ is a selfadjoint isomorphism, defined by a pseudo–Riemannian structure on M. Then the Levi–Civita connection of this structure is the adjoint connection to the canonical connection, associated with the Hamiltonian $h : \lambda \mapsto \frac{1}{2} < \lambda, Q^{-1}\lambda >,\ \lambda \in T^*M$.*

Proof. The equation $\dot{\lambda} = \vec{h}(\lambda)$ defines the pseudo–Riemannian geodesic flow in T^*M, where $\pi(\lambda)^\cdot = Q^{-1}\lambda$ along every trajectory of this flow. Hence the isomorphism $Q^{-1} : T^*M \longrightarrow TM$ transforms the geodesic flow in T^*M into the geodesic flow in TM, where the last flow is defined by a spray Z, quadratic on fibres. Let $X \mapsto \nabla_X$ be the canonical connection for h. From the definition, it follows that $X \mapsto Q_*^{-1}\nabla_X$ is the canonical connection for Z. Exploiting the fact that the parallel translation generated by the canonical connection preserves the Hamiltonian,(the canonical connection is tangent to the levels of the Hamiltonian), we obtain $Q_*^{-1}\nabla_X = \nabla_X^*$. From Proposition 5.1 follows now that ∇_X^* is the Levi–Civita connection, cf. also the remark after the Proof of Proposition 5.1.

Remark. For pseudo–Riemannian structure, the master–Hamiltonian has the form $\lambda \mapsto < \lambda, Q^{-1}\lambda >^{\frac{1}{2}}$, which is defined, in general, not for all $\lambda \in T^*M$.

§6. Two–Dimensional Control Systems

Consider a standard "state invariant" control system

$$\dot{x} = f(x, u), \quad x \in M,\ u \in U.$$

It is interesting to find explicit expressions of such a fundamental "state" and "feedback" invariant as the curvature tensor through classical "state"–invariants — linear relations between iterated Lie brackets of vector fields $f(\cdot, u),\ u \in U$. We already have all necessary means to derive such expressions, though some efforts are still needed, and the obtained expressions turn out to be pretty complicated even in the simplest two–dimensional case. We restrict ourself to the two–dimensional case only.

Assume that $dim\, M = 2,\ dim\, U = 1$, hence U is $\mathbb{R}$ or $S^1 = \mathbb{R}/2\pi\mathbb{Z}$. Put

$$\mathcal{O} = \left\{ (x, u) \in M \times U \Big| \ \frac{\partial f}{\partial u} \wedge \frac{\partial^2 f}{\partial u^2}\Big|_{(x,u)} \neq 0 \right\}.$$

All extremal controls, contained in the region $\mathcal{O}$, correspond to regular extremals, and through every point of the region passes exactly one extremal control, corresponding to a uniquely determined extremal, (up to a nonzero factor for λ). Assume $R(x,u)$, $(x,u) \in \mathcal{O}$ is the curvature tensor of the uniquely determined germ of extremal, with the corresponding control through (x,u). The tensor is a linear operator, and since it acts in the one–dimensional space, $(x,u) \mapsto R(x,u)$ is a real–valued function on $\mathcal{O}$.

Proposition 6.1. *Assume $\alpha_{i_1\dots i_k}$, $\beta_{i_1\dots i_k}$ are smooth functions on $\mathcal{O}$, defined by the formula*

$$\left[\frac{\partial^{i_1} f}{\partial u^{i_1}}, \left[\frac{\partial^{i_2} f}{\partial u^{i_2}}, \left[\dots, \frac{\partial^{i_k} f}{\partial u^{i_k}}\right] \dots\right]\right] = \alpha_{i_1\dots i_k}\frac{\partial f}{\partial u} + \beta_{i_1\dots i_k}\frac{\partial^2 f}{\partial u^2}.$$

Then

$$\begin{aligned} R =& \alpha_{001} + \beta_{101} - \frac{1}{2}\beta_{002} + \frac{1}{2}\beta_{001}\beta_3 - \frac{1}{2}\beta_{01}\beta_{12} + \frac{3}{2}\alpha_{01}\beta_{02} - \alpha_{02}\beta_{01} + \beta_{01}\beta_{03} - \\ & 2\alpha_{01}^2 + \frac{1}{4}\beta_{02}^2 - \alpha_{01}\beta_{01}\beta_3 - \frac{3}{2}\beta_{01}\beta_{02}\beta_3 + \alpha_3\beta_{01}^2 - \frac{1}{2}\beta_{01}^2\beta_4 + \frac{3}{4}\beta_{01}^2\beta_3^2 \end{aligned}$$

Proof. We shall use Proposition 3.2, which expresses the $\mathcal{L}$–derivative of the endpoint–mapping, hence the Jacobi curve, through solutions of the linear Hamiltonian system.

Assume $\tau \mapsto (x_\tau, u_\tau)$ is an extremal control in $\mathcal{O}$. Put

$$g_\tau(\lambda, u) = \lambda \overrightarrow{exp} \int_t^\tau ad\, f(\cdot, u_\theta) d\theta f(\cdot, u)\Big|_x, \quad g_\tau^{(i)}(\lambda) = \frac{\partial^i}{\partial u^i} g_\tau(\lambda, u_\tau), \ (x,u) \in \mathcal{O}, \ \lambda \in T_x^*M.$$

Let $\lambda_t \in T_{x_t}^*M$ be the Lagrange multiplier corresponding to the given extremal control. Then $g_\tau^{(1)}(\lambda_t) = 0$ identically in τ. Differentiating the last identity with respect to τ yields the following useful formula

$$\frac{d}{d\tau}u_\tau = \frac{\left\{g_\tau^{(1)}, g_\tau^{(0)}\right\}}{g_\tau^{(2)}}\Big|_{\lambda_t},$$

where $\{\cdot,\cdot\}$ are the Poisson brackets. Note that the inclusion $(x_\tau, u_\tau) \in \mathcal{O}$ implies that $g_\tau^{(1)}(\lambda_t), g_\tau^{(2)}(\lambda_t)$ can not vanish simultaneously. Assume, for definiteness, that $g_\tau^{(2)}(\lambda_t) < 0$. Certainly, substituting λ_t by $-\lambda_t$ and, accordingly, $g_\tau^{(2)}(\lambda_t)$ by $g_\tau^{(2)}(-\lambda_t) = -g_\tau^{(2)}(\lambda_t)$, we do not change the curvature. The quadratic Hamiltonian q_τ from Proposition 3.2 has in our case the form

$$q_\tau(\eta) = -\frac{\sigma\left(\vec{g}_\tau^{(1)}(\lambda_t), \eta\right)^2}{2g_\tau^{(2)}(\lambda_t)}.$$

The Hamiltonian q_τ is defined on a 4–dimensional symplectic space $T_{\lambda_t}(T^*M)$, though the vertical line $\mathbb{R}\lambda_\tau \subset T_{x_\tau}^* \approx T_{\lambda_t}(T_{x_t}^*M)$ consists entirely of fixed points of the Hamiltonian system. The Jacobi curve is constructed with the help of the solutions of the reduced system, defined on $\Sigma_{\lambda_t} = (\mathbb{R}\lambda_t)^\angle/\mathbb{R}\lambda_t$, cf. the beginning of §4. There exists a unique covector $e \in T_{\lambda_t}(T_{x_t}^*M)$ satisfying the conditions $\sigma(e, \vec{g}_t^{(1)}) = 1$, $\sigma(e, \vec{g}_t^{(2)}) = 0$. Consider the restriction of the Hamiltonian q_τ onto the symplectic subspace $span(e, \vec{g}_t^{(1)}) \subset (\mathbb{R}\lambda_t)^\angle$. Let $\eta = ye + z\vec{g}_t^{(1)}$, then

$$q_\tau(\eta) = \frac{1}{2}(a_\tau y + b_\tau z)^2,\ a_\tau = \frac{\sigma(\vec{g}_\tau^{(1)}, e)}{\sqrt{|g_\tau^{(2)}|}},\ b_\tau = \frac{\sigma(\vec{g}_\tau^{(1)}, \vec{g}_t^{(1)})}{\sqrt{|g_\tau^{(2)}|}},\ b_t = 0.$$

Nonstationary Hamiltonian q_τ defines a linear flow on $\mathbb{R}^2$ and, accordingly, a flow on $L(\mathbb{R}^2) = \mathbb{R}P^1$. The Jacobi curve $\tau \mapsto J(\tau)$ is a trajectory of the flow on $\mathbb{R}P^1$, inverse to the flow defined by the Hamiltonian q_t, $J(t) = \mathbb{R}e$. Let $\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix}$, $\Phi(t) = id$, be the fundamental matrix of the linear flow on $\mathbb{R}^2$, inverse to the flow defined by the Hamiltonian q_τ. Then

$$\dot{\Phi} = \Phi \begin{pmatrix} ab & b^2 \\ -a^2 & -ab \end{pmatrix},\ det\,\Phi = 1. \tag{6.2}$$

As a local coordinate on $\mathbb{R}P^1$ take $\frac{z}{y}$, then $J(\tau)$ is represented by the 1×1–matrix $S_\tau = \dfrac{\varphi_{21}(\tau)}{\varphi_{11}(\tau)}$. For the curvature we obtain the expression

$$R(x_t, u_t) = \left(\frac{\ddot{S}_t}{2\dot{S}_t}\right)^{\cdot} - \left(\frac{\ddot{S}_t}{2\dot{S}_t}\right)^2 = \frac{\ddot{a}_t}{a_t} - a_t\dot{b}_t - 2\left(\frac{\dot{a}_t}{a_t}\right)^2.$$

Note that

$$\frac{\partial}{\partial\tau} g_\tau^{(i)} = g_\tau^{(i+1)}\frac{du}{d\tau} + \left\{g_\tau^{(0)}, g_\tau^{(i)}\right\} = \frac{g_\tau^{(i+1)}}{g_\tau^{(2)}}\left\{g_\tau^{(1)}, g_\tau^{(0)}\right\} + \left\{g_\tau^{(0)}, g_\tau^{(i)}\right\}.$$

Furthermore, the quantities $g_t^{(i)}(\lambda) = \lambda\dfrac{\partial^i}{\partial u^i}f$, hence, the Poisson brackets of $g_t^{(i)}$, are expressed through the Lie brackets of the fields $\frac{\partial^i}{\partial u^i}f$. Therefore, the consecutive derivatives of the functions a_τ, b_τ, with respect to τ for $\tau = t$, are expressed explicitly, though quite cumbersome, through $\alpha_{i_1\dots i_k}(x_t, u_t), \beta_{i_1\dots i_k}(x_t, u_t)$. Direct calculations give the expression (6.1). Though pretty awkward, this formula is strongly simplified in some important special cases. Consider two–dimensional Riemannian and Lorentzian geometries. In the Riemannian case we have

$$f(x, u) = (cos\,u)v_1(x) + (sin\,u)v_2(x),$$

where v_1, v_2 is an arbitrary orthonormal frame of the considered Riemannian structure. It is easy to eliminate in (6.1) all indices ≥ 2. Indeed, from every such index we can subtract 2, at the same time changing the sign of the corresponding coefficient. Taking into account also the symmetries of the coefficients as the "structure constants", we obtain

$$R = \alpha_{001} + \beta_{101} - 2(\alpha_{01}^2 + \beta_{01}^2).$$

Note that R is the Gaussian curvature of the Riemannian surface, according to Proposition 5.2 and Theorem 5.1.

Let $[v_1, v_2] = c_1 v_1 + c_2 v_2$, for some smooth functions c_1, c_2. Then

$$R = v_1 c_2 - v_2 c_1 - c_1^2 - c_2^2.$$

In the Lorentzian case

$$f(x,u) = (ch\, u) v_1(x) + (sh\, u) v_2(x).$$

Again, 2 could be subtracted from the indices in (6.1), this time without changing the coefficients. For the Gaussian curvature we obtain,

$$R = \alpha_{001} - 2(\alpha_{01}^2 - \beta_{01}^2) = v_1 c_2 + v_2 c_1 + c_1^2 - c_2^2.$$

References

1. A. A. Agrachev, *Quadratic Mappings in Geometric Control Theory*, Itogi Nauki i Tekhniki; Problemy Geometrii, VINITI, Acad. Nauk SSSR **20** (1988), VINITI, Moscow, 111-205. (Russian)
   English transl. in J. Soviet Math., Plenum Publ. Corp., Vol.51, 1990, pp. 2667-2734.
2. A. A. Agrachev, *Topology of Quadratic Mappings and Hessians of Smooth Mappings*, Itogi Nauki i Tekhniki; Algebra, Topologia, Geometria; VINITI, Acad. Nauk SSSR **26** (1988), VINITI, Moscow, 85-124. (Russian)
   English transl. in J. Soviet Math., Plenum Publ. Corp., 1990, pp. 990-1013
3. A. A. Agrachev, R. V. Gamkrelidze, *Exponential Representation of Flows and Chronological Calculus*, Matem. Sbornik **107** (1978), Moscow, 467-532. (Russian)
   English transl. in Math. USSR Sbornik, Vol.35, pp. 727-785, 1979.
4. A. A. Agrachev, R. V. Gamkrelidze, *The Morse index and the Maslov index for extremals of controlled systems*, Dokl. Acad. Nauk SSSR **287** (1986), VINITI Moscow. (Russian)
   English transl. in Soviet Math. Dokl., Vol.33, pp. 392-395, 1986
5. A. A. Agrachev, R. V. Gamkrelidze, *The extremality index and quasi-extremal controls*, Dokl. Acad. Nauk SSSR **284** (1985). (Russian)
   English transl. in Soviet Math. Dokl., Vol.32, pp. 478-481, 1985.
6. A. A. Agrachev, R. V. Gamkrelidze, *The quasi-extremality for controlled systems*, Dokl. Itogi Nauki. VINITI. Sovremennye problemy matematiki. Novejshie dostigeniya **35** (1989). (Russian)
   English transl. in J. Soviet Math., Plenum Publ. Corp., pp. 1849-1864, 1991.
7. A. A. Agrachev, R. V. Gamkrelidze, *Symplectic Methods for optimization and control*, to appear in: "Geometry of Feedback and Optimal Control", Marcel Dekker.
8. A. A. Agrachev, R. V. Gamkrelidze and A. V. Sarychev, *Local Invariants of Smooth Control Systems*, Acta Applicandae Mathematicae **14** (1989), 191-237.
9. V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 3-ed, Nauka, Moscow, 1989, p. 472.
10. B. Bonnard, *Feedback equivalence for nonlinear systems and the time optimal control problems*, SIAM J. Control and Optim. **29** (1991), 1300-1321.
11. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko, *Mathematical theory of optimal processes*, Fiz.-mat. giz, Moscow, 1961, p. 462.