Hamiltonian Systems and Optimal Control

Andrei Agrachev

Abstract Solutions of any optimal control problem are described by trajectories of a Hamiltonian system. The system is intrinsically associated to the problem by a procedure that is a geometric elaboration of the Lagrange multipliers rule. The intimate relation of the optimal control and Hamiltonian dynamics is fruitful for both domains; among other things, it leads to a clarification and far going generalization of important classical results about Riemannian geodesic flows.

Introduction

These are notes of the lectures given in June 25–28, 2007 for the school "Hamiltonian Dynamical Systems and Applications" in Montreal and written by Natalia Shcherbakova.

Hamiltonian systems play a very important role in the optimal control theory since the foundation of the this subject in the middle of the twenties Century. Indeed, the first fundamental result of the theory, the Pontryagin maximum principle, is formulated in the Hamiltonian form. As I learned from Ponryagin collaborators, it was the central role played by the Hamiltonian system that convinced Pontryagin in the importance and universality of the found by him optimality condition. See the pioneering book [5] for the original approach and the books [3], [4], [2] for some further developments.

The Pontryagin maximum principle is a natural Hamiltonian form for the first order optimality conditions. In these lectures, we explain the Hamiltonian nature of the second order information on the local structure of the optimal control problem which leads, among other things, to the curvature-type invariants of Hamiltonian systems on the cotangent bundles. These invariants control Hamiltonian dynamics in the way similar to the Riemannian sectional curvature does for the geodesic flows.

A. Agrachev
SISSA, Trieste and MIAN, Moscow, e-mail: agrachev@sissa.it

A more detailed and formal exposition of the presented here constructions and facts can be found in [1].

Lecture I

Consider two smooth[1] manifolds: M, $\dim M = n$, a *state space* and U, $\dim U = m \leq n$, a *space of control parameters*.

Control system is a family of ordinary differential equations on M:

$$\dot{q} = f(q,u), \quad ,q \in M,\ u \in U. \tag{1}$$

The vector field f is assumed to be smooth with respect to both variables. For $q \in M$ fixed, the set $f(q,U)$ is called the set of admissible velocities.

Any L_∞-curve $t \mapsto u(t),\ t \in [0,t_1],\ u(t) \in U$ is called the control function. Here t_1 is fixed. Substituting the control function $u(t)$ into (1) we get a non-autonomous differential equation

$$\dot{q} = f(q,u(t)).$$

A solution $t \mapsto q(t) \in M$ of this equation is called a trajectory of (1) associated to $u(t)$.

Consider

$$\mathscr{W} = \{(u(\cdot),q(\cdot)) : u(\cdot) \in L_\infty([0,t_1],U),\ \dot{q}(t) = f(q(t),u(t))\},$$

the space of admissible pairs.

Remark. $\mathscr{W}$ is a Banach manifold modelled on $\mathbb{R}^n \times L^m_\infty([0,t_1])$. This fact is a direct consequence of the standard theorem on the existence, uniqueness and smooth dependence of the data for solutions to the Cauchy problem of systems of ordinary differential equations.

Let

$$J^{t_1}(u(\cdot),q(\cdot)) = \int_0^{t_1} \varphi(q(t),u(t))\,dt,$$

where φ is a smooth scalar function. J^{t_1} is called the cost functional.

Optimal control problem: Given $q_0, q_1 \in M$ minimize J^{t_1} over admissible pairs $(u(\cdot),q(\cdot))$ such that $q(0) = q_0,\ q(t_1) = q_1$.

This problem generalizes standard problem of the Calculus of Variations:

$$\int_0^{t_1} \varphi(q(t),\dot{q}(t))\,dt \longrightarrow \min$$

[1] In these lecture notes smooth objects are C^∞ unless otherwise stated

over all curves such that $q(0) = q_0,\ q(t_1) = q_1$. This problem can be stated as an optimal control problem by setting $\dot{q} = u,\ u \in T_qM$.

Geometrically speaking, we state the problem as follows: consider a locally trivial bundle $V \to M$ over M whose typical fiber is U: $V = \bigcup\limits_{q \in M} V_q,\ V_q \cong U$. Then $f : V \to TM$ such that $f(V_q) \subset T_qM$. An admissible pair is a curve $t \to v(t)$ in V such that $v(t) \in V_{q(t)}$ and $\dot{q}(t) = f(v(t))$.

Given an admissible pair, we can trivialize V along the trajectory.

In order to describe extremals of the optimal control problem we will use the geometrically elaborated Lagrange multipliers method. This method provides a first order optimality conditions. In forthcoming lectures we will also discuss II-nd order conditions and related invariants.

First order conditions

Optimal control problem is a kind of a constrained optimization problem, where constrains are given by the boundary points condition. We set $F_t : \mathscr{W} \to M$ such that $F_t(u(\cdot 0, q(\cdot)) = q(t)$. The map F_t has the same smoothness as the field f. Our minimization problem is

$$\min J^{t_1}\Big|_{F_0 = q_0,\, F_{t_1} = q_1}. \tag{2}$$

We are looking for solutions of the problem among solutions of the equation

$$dJ^{t_1}\Big|_{\ker DF_0 \cap \ker DF_{t_1}} = 0$$

that is equivalent to the equation

$$dJ^{t_1} = \lambda_{t_1} DF_{t_1} - \lambda_0 DF_0, \tag{3}$$

where $\lambda_0 : T_{q_0}M \to \mathbb{R}$ is a linear form on $T_{q_0}M$ and λ_{t_1} is a linear form on $T_{q_1}M$, i. e. $\lambda_0 \in T^*_{q_0}M,\ \lambda_{t_1} \in T^*_{q_1}M$. Covectors λ_0, λ_{t_1} are nothing else but the Lagrange multipliers. The sign "$-$" in front of λ_0 is chosen for further convenience.

We have $DF_t : T_{(u,q)}\mathscr{W} \to T_{q(t)}M$ and

$$\lambda_t DF_t : T_{(u,q)}\mathscr{W} \xrightarrow{DF_t} T_{q(t)}M \xrightarrow{\lambda_t} \mathbb{R}.$$

Proposition 1 *Equation (3) implies that there exists a unique Lipschitz curve $\lambda_t \in T^*_{q(t)}M,\ 0 \le t \le t_1$ such that*

$$\lambda_t DF_t - \lambda_0 DF_0 = dJ^t,$$

where $J^t = \int\limits_0^t \varphi(q(\tau), u(\tau))\, d\tau$.

Proof. The uniqueness follows from the fact that F_t is a submersion. Let us prove the existence. Denote by $(\tilde{u}(\cdot), q(\cdot))$ the reference pair. By (3),

$$dJ^{t_1}_{(\tilde{u},q)} = \lambda_{t_1} DF_{t_1} - \lambda_0 DF_0.$$

Let us fix some $t \in [0, t_1]$. We restrict ourselves to the admissible pairs of the form

$$\mathscr{V}_t = \{(u(\cdot), q(\cdot)) : u(\tau) = \tilde{u}(\tau) \text{ for } \tau \geq t\}.$$

Then $F_t(u,q) = \Phi_t^{t_1} \circ F_t$, where $\Phi_t^{t_1}$ is a fixed diffeomorphism. The family $\tau \mapsto \Phi_t^\tau$ satisfies $\Phi_t^t(q) = q$,

$$\frac{\partial}{\partial \tau} \Phi_t^\tau(q) = f\left(\Phi_t^\tau(q), \tilde{u}(\tau)\right),$$

i. e. $\Phi_t^\tau : q(t) \mapsto q(\tau)$. Restricting to $\mathscr{V}_t$ we obtain at $v \in \mathscr{V}_t$

$$\lambda_{t_1} D_v(\Phi_t^{t_1} \circ F_t) - \lambda_0 D_v F_0 = d_v J^t + d_v(a_t \circ F_t),$$

where $a_t = \int\limits_t^{t_1} \varphi\left(\Phi_t^s(q(t)), \tilde{u}(s)\right) ds$. Hence

$$\lambda_t D_v F_t - \lambda_0 D_v F_0 = d_v J^t,$$

where $\lambda_t = \lambda_{t_1} D_v \Phi_t^{t_1} - d_{q(t)} a_t$.

Lecture II

Let us consider the Lipschitz curve $\lambda_t \in T^*_{q(t)}M$ whose existence was proved in Lecture I. This curve satisfies the following relation:

$$\lambda_t D_{(\tilde{u},q)} F_t - \lambda_0 D_{(\tilde{u},q)} F_0 = d_{(\tilde{u},q)} J^t. \tag{1}$$

Now we are going to derive a differential equation for λ_t. First of all, we can introduce local coordinates in the neighborhood of given $q(t) \in M$; then $\lambda = (p, q),\ p \in \mathbb{R}^{n*},\ q \in \mathbb{R}^n$ for any q from the coordinate neighborhood and any λ from the cotangent bundle to this neighborhood. In particular, $\lambda_t = (p_t, q(t))$. Then (1) becomes

$$d_{(\tilde{u},q)}(p_t F_t) - \lambda_0 D_{(\tilde{u},q)} F_0 = d_{(\tilde{u},q)} J^t, \tag{2}$$

where $d_{(\tilde{u},q)}$ is the differential in the Banach space of admissible pairs. After taking derivatives with respect to t we get

$$0 = d_{(\tilde{u},q)} \left(\frac{\partial}{\partial t}(p_t F_t) - \varphi(q(t), \tilde{u}(t)) \right) =$$

$$d_{(\tilde{u},q)}\left(\frac{dp}{dt}q(t)+p_t f(q(t),\tilde{u}(t))\right)-\varphi(q(t),\tilde{u}(t)).$$

Let us denote

$$h(p,q,u)=pf(q,u)-\varphi(q,u).$$

The function h is called the Hamiltonian of the optimal control problem. We have

$$d_{(\tilde{u},q)}\left(\dot{p}_t q(t)+h(p_t,q(t),\tilde{u}(t))\right)=0.$$

Hence we obtain

$$\dot{p}_t+\frac{\partial h}{\partial q}=0 \quad \frac{\partial h}{\partial u}=0.$$

Recall now that our original dynamics is given by $\dot{q}=f(q,u)=\frac{\partial h}{\partial p}$. So, we obtain a Hamiltonian system

$$\begin{cases} \dot{p}=-\dfrac{\partial h}{\partial q}(p,q,u) \\ \dot{q}=\dfrac{\partial h}{\partial p}(p,q,u) \end{cases} \tag{$(HS)_u$}$$

plus one extra equation

$$\frac{\partial h}{\partial u}(p,q,u)=0.$$

In a regular situation, one can find $u=u(p,q)$ from the last equation. Then we denote:

$$H(p,q)=h(p,q,u(p,q)),$$

and for remaining variables we obtain a standard Hamiltonian system

$$\begin{cases} \dot{p}=-\dfrac{\partial H}{\partial q}(p,q) \\ \dot{q}=\dfrac{\partial H}{\partial p}(p,q). \end{cases} \tag{(HS)}$$

From now on, we will always assume that u can be eliminated from the equation $(HS)_u$.

Second variation

Let $t\in[0,t_1]$, we are going to study the second derivative of J^t under the constraints

$$\dot{q}=f(q,u), \quad q(0)=q_0,\ q(t)=q_t.$$

The correctness of the Cauchy problem for ordinary differential equations allows us to immediately resolve one of the boundary constraints. From now on, we restrict ourselves to $F_t^{-1}(q_t)$ and slightly change notations. In our new notations:

$$J^t := J^t\big|_{F_t^{-1}(q(t))}, \quad F_0 := F_0\big|_{F_t^{-1}(q_t)}.$$

Since the trajectory $q(\cdot)$ is uniquely defined by the corresponding control $u(\cdot)$, the space of admissible pairs is now reduced to the space of control functions $\{u(\cdot)\}$.

Equation (1) now reads

$$\lambda_0 D_{\tilde{u}} F_0 + d_{\tilde{u}} J^t = 0. \tag{1$'$}$$

The Hessian of $J^t\big|_{F_0^{-1}(q_0)}$ is

$$Hess_{\tilde{u}}\left(J^t\big|_{F_0^{-1}(q_0)}\right) = \left(D^2_{\tilde{u}} J^t + \lambda_0 D^2 F_0\right)\big|_{\ker D_{\tilde{u}} F_0}.$$

The Hessian is a quadratic form; we are interested in the Morse index of this form that is the supremum of the dimensions of subspaces where the form is negative definite. In particular, Morse index equals 0 if and only if the form is nonnegative. We use notation ind$Q$ for the Morse index of a quadratic form Q; the value of this index is a nonnegative integer or $+\infty$.

Let us express everything in terms of the Lagrange multipliers. After the elimination of one of the boundary constraints, equation (2) reads:

$$\begin{cases} p\dfrac{\partial F_0}{\partial u} + \dfrac{\partial J^t}{\partial u} = 0 \\ q = F_0(u). \end{cases}$$

Linearizing this equation we obtain

$$\xi \frac{\partial F_0}{\partial u} + \left(p\frac{\partial^2 F_0}{\partial u^2} + \frac{\partial^2 J^t}{\partial u^2}\right) v = 0 \quad \eta = \frac{\partial F_0}{\partial u} v. \tag{LS}$$

Here $(\xi, \eta, v) = (\delta p, \delta q, \delta u)$. Denote

$$\Lambda_t = \{(\xi, \eta) : \exists v \text{ satisfying } (LS)\},$$

the projection of the space of solutions of (LS) into the (p,q)-space.

We made computations in coordinates but the construction of Λ_t (linearization and projection) is intrinsic, and Λ_t is actually a well-defined subspace of $T_{\lambda_0}(T^*M)$. Recall that T^*M is a symplectic manifold endowed with the standard symplectic structure $dp \wedge dq$. Hence $T_{\lambda_0}(T^*M)$ is a symplectic space. It is not hard to check that Λ_t is a Lagrangian subspace of $T_{\lambda_0}(T^*M)$.

The family $t \mapsto \Lambda_t,\ t \in [0, t_1]$ is called the *Jacobi curve*.

We will use Λ_t to calculate the Morse index of the second variation of J^t under constraints. Indeed, system (LS) is of the type

$$\xi A + Bv = 0 \quad \eta = Av, \tag{3}$$

where
$$A : V \to \mathbb{R}^n, \quad B = p\frac{\partial^2 F_0}{\partial u^2} + \frac{\partial^2 J^t}{\partial u^2},$$
so that $B : V \to V^*$ and $B^* = B$. Then
$$Hess\left(J^t\big|_{F_0^{-1}(q_0)}\right) : v \mapsto \langle Bv, v\rangle, \quad v \in \ker A. \tag{4}$$

Exercise. Assume that the linear mapping A is surjective and B is non-degenerate. Then quadratic form (4) is non-degenerate if and only if

$$\Lambda \bigcap \{(\xi,0) : \xi \in \mathbb{R}^{n*}\} = 0,$$

where $\Lambda = \{(\xi,\eta) : \exists v \text{ satisfying}(3)\}$.

Let us follow the evolution of the Morse index of the quadratic form $Hess\left(J^t\big|_{F_0^{-1}(q_0)}\right)$, when t runs from 0 to t_1. The Morse index changes for those t for which Λ_t has a nonzero intersection with $\Lambda_0 = \{(\xi,0) : \xi \in \mathbb{R}^{n*}\}$. Moreover, the Morse index of $Hess\left(J^t\big|_{F_0^{-1}(q_0)}\right)$ grows monotonically with t simply due to the fact that the past does not depend on the future for our control system.

Definition 1. We say that t is conjugate to 0 if $\Lambda_t \cap \Lambda_0 \neq 0$.

Theorem 1. *If conjugate times are isolated, then*

$$\mathrm{ind} Hess\left(J^t\big|_{F_0^{-1}(q_0)}\right) - Hess\left(J^{t_0}\big|_{F_0^{-1}(q_0)}\right) = \sum_{t_0 \leq t < t_1} \dim(\Lambda_t \cap \Lambda_0)$$

for any $t_0 \in (0, t_1)$.

Lecture III

Let us consider the trajectories of the Hamiltonian system corresponding to the Hamiltonian
$$H : T^*M \to \mathbb{R}.$$
We denote by $\mathbf{H}$ the associated Hamiltonian vector field and by $e^{t\mathbf{H}}$ its flow on T^*M. Recall that the trajectories of $\mathbf{H}$ describe the extremals of our optimal control problem together with the associated Lagrange multipliers.

In Lecture II, we saw that second variation of the cost functional J^t under constraints is related to a family of the Lagrangian subspaces Λ_t, $t \in [0,t_1]$, where Λ_t is the tangent space at λ_0 of the submanifold of T^*M formed by the values at time 0 of the solutions to the Hamiltonian system $\dot{\lambda} = \mathbf{H}(\lambda)$ whose values at time t belong to $T^*_{q(t)}M$.

In other words, we consider the Jacobi curve

$$\Lambda_t = T_{\lambda_0}\left(e^{-t\mathbf{H}}(T^*_{q(t)}M)\right).$$

Our aim is to obtain some information on the conjugate points and on the long-time behavior of the flow without solving differential equations. To do that, we introduce a kind of curvature-type invariants associated to the problem.

Curves in the Lagrange Grassmannians

Consider a $2n$-dimensional symplectic space Σ. Denote by σ the symplectic form. To any pair of transversal Lagrangian subspaces Λ_0, Δ, we can associate coordinates on Σ such that

$$\Sigma = \Lambda_0 \oplus \Delta = \{(p,q) : p,q \in \mathbb{R}^n\}, \quad \sigma = \sum_{i=1}^n dp_i \wedge dq_i$$

and

$$\Lambda_0 = \{(p,0) : p \in \mathbb{R}^n\}, \quad \Delta = \{(0,q) : q \in \mathbb{R}^n\}.$$

Any transversal to Δ n-dimensional subspace Λ has a form

$$\Lambda = \{(p, S_\Lambda p) : p \in \mathbb{R}^n\},$$

for some $n \times n$-matrix S_Λ. The subspace Λ is Lagrangian (i. e. $\sigma|_\Lambda = 0$) if and only if $S_\Lambda = S^*_\Lambda$, in other words, if S_Λ belongs to the space $\mathrm{Sym}(n)$ of symmetric $n \times n$-matrices. In particular, it follows that the map $\Lambda \mapsto S_\Lambda$ gives local coordinates on the *Lagrange Grassmannian* $L(\Sigma)$ of all Lagrangian subspaces of Σ.

Now consider a curve $t \mapsto \Lambda_t$ in $L(\Sigma)$ and the corresponding curve $t \mapsto S_t = S_{\Lambda_t}$ in $\mathrm{Sym}(n)$.

Lemma 1. *The quadratic form $p \mapsto \langle \dot{S}_t p, p\rangle$ is an intrinsically defined quadratic form on the subspace Λ_t.*

Proof. Pick $\lambda_t \in \Lambda_t$ and insert it in some curve $\tau \mapsto \lambda_\tau \in \Lambda_\tau$. Then $\lambda_\tau = (p_\tau, S_\tau p_\tau)$ and we have:

$$\sigma(\lambda_t, \dot{\lambda}_t) = \sigma\left((p_t, S_t p_t), (\dot{p}_t, S_t \dot{p}_t) + (0, \dot{S}_t p_t)\right) = \langle \dot{S} p_t, p_t\rangle.$$

We see that $\sigma(\lambda_t, \dot{\lambda}_t)$ depends only on λ_t and $\dot{\Lambda}_t$ and not on the choice of the curve $\tau \mapsto \lambda_\tau$. Hence $\lambda_t \mapsto \sigma(\lambda_t, \dot{\lambda}_t)$ is a well-defined quadratic form on Λ_t presented by the matrix $\dot{S}_t$. $\square$

Corollary 1. *The tangent space $T_\Lambda L(\Sigma)$ to the Lagrange Grassmannian is intrinsically identified with the space of quadratic forms on Λ.*

Definition 2. We say that the curve $t \mapsto \Lambda_t$ is monotone if the quadratic forms $p \mapsto \dot{S}_t$ are sign-definite.

Curvature-type invariants

Let us come back to the Jacobi curves in $T_z(T^*M),\ z \in T^*M$. Recall that $T_z(T^*M) = \Sigma_z$ is a symplectic space, so we are in the situation described above.

We want to consider the curves $t \mapsto \Lambda_z(t)$ such that $\Lambda_z(0) = T_z(T_q^*M)$, where $q = \pi(z)$ and $\pi : T^*M \to M$ is the standard projection. As before, choosing appropriate Darboux coordinates in $T_z(T^*M)$ we have:

$$\Lambda_t = \{(p, S_t p) : p \in \mathbb{R}^n\}, \quad S_0 = 0.$$

By a direct computation one verifies that

$$\dot{S}_0 = -\frac{\partial^2 H}{\partial p^2}\Big|_z.$$

From now on, we will deal with the Hamiltonians that are convex on fibers, i. e. we will assume that

$$\frac{\partial^2 H(p,q)}{\partial p^2} \geq 0, \quad (p,q) \in T^*M.$$

In what follows we identify $\dot{\Lambda} \cong \dot{S}_0$.

Let us show that the convexity assumption on H implies the monotonicity of the Jacobi curves. In other words,

$$\dot{\Lambda}_z(0) \leq 0 \ \forall z \quad \Rightarrow \quad \dot{\Lambda}_z(t) \leq 0 \ \forall t.$$

Indeed, $\Lambda_z(t) \subset T_z(T^*M)$ and

$$\Lambda_z(t) = e_*^{-t\mathbf{h}}\left(T_{e^{t\mathbf{h}}(z)}(T^*_{q(t)}M)\right);$$

then

$$\frac{d}{dt}\Lambda_z(t) = e_*^{-t\mathbf{H}}\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} e_*^{-\varepsilon\mathbf{h}}\left(T_{e^{(t+\varepsilon)\mathbf{h}}(z)}(T^*_{q(t+\varepsilon)}M)\right) =$$
$$e_*^{-t\mathbf{H}}\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\Lambda_{e^{t\mathbf{H}}(z)}(\varepsilon) \leq 0.$$

Here we use the fact that the quadratic form $\dot{\Lambda}_t$ is defined intrinsically and its sign does not change under a symplectic transformation of the curve Λ_t.

To any pair of transversal Lagrangian subspaces $\Lambda, \Lambda_0 :\ \Lambda \cap \Lambda_0 = 0,\ \Lambda, \Lambda_0 \subset \Sigma$, we can associate the projector

$$\pi_{\Lambda\Lambda_0} : \Sigma \to \Lambda_0$$

such that

$$\pi_{\Lambda\Lambda_0}\big|_{\Lambda} = 0, \quad \pi_{\Lambda\Lambda_0}\big|_{\Lambda_0} = Id.$$

Lemma 2. *The space $\{\pi_{\Lambda\Lambda_0} : \Lambda \cap \Lambda_0\}$ for a fixed Λ_0 is an affine subspace of the space of linear operators on Σ.*

In local coordinates such that

$$\Lambda_0 = \{(p,0) : p \in \mathbb{R}^n\}, \quad \Lambda = \{(Aq,q) : q \in \mathbb{R}^n\},$$

we have:

$$\pi_{\Lambda\Lambda_0} = \begin{pmatrix} Id & -A \\ 0 & 0 \end{pmatrix}.$$

Assume that the curve $t \mapsto \Lambda_t$ is such that $\Lambda_t \cap \Lambda_\tau = 0$ if $t \neq \tau$ and $|t-\tau|$ is sufficiently small. Pick τ and take coordinates such that the germ of the curve at τ has the form

$$\Lambda_t = \{(p, S_t p) : p \in \mathbb{R}^n\}, \quad S_\tau = 0.$$

Then we obtain:

$$\pi_{\Lambda_t\Lambda_\tau} = \begin{pmatrix} Id & -S_t^{-1} \\ 0 & 0 \end{pmatrix}.$$

If $\det S_t$ has a finite order root at τ, then $\pi_{\Lambda_t\Lambda_\tau}$ admits the Laurent expansion

$$\pi_{\Lambda_t\Lambda_\tau} = \sum_{i=-k}^{m} (t-\tau)^i \pi_\tau^i + o(t-\tau)^{m+1}.$$

The free term of the Laurent expansion π_τ^0 belongs to the described in the Lemma affine space of the projectors (where Λ_0 is substituted by Λ_τ). In other words, $\pi_\tau^0 = \pi_{\Lambda_\tau^\circ\Lambda_\tau}$ for some intrinsically defined Lagrangian subspace Λ_τ°, and $\Lambda_\tau^\circ \cap \Lambda_\tau = 0$ by the construction.

In particular, in the simplest case $k=1$ the coordinate expression of Λ_τ° has the form:

$$\Lambda_\tau^\circ = \left\{ (-\frac{1}{2}\dot{S}_\tau^{-1}\ddot{S}_\tau\dot{S}_\tau^{-1}q, q) : q \in \mathbb{R}^n \right\}.$$

Now recall that we identify $\dot{\Lambda}_t$ with a quadratic form on Λ_t, i. e.

$$\dot{\Lambda}_t : \Lambda_t \to \Lambda_t^*,$$

a self-adjoint linear mapping. Moreover, the symplectic form σ gives the non-degenerate pairing of the Lagrangian subspaces Λ_t and Λ_t°, so that $\Lambda_t^* \cong \Lambda_t^\circ$ and $(\Lambda_t^\circ)^* \cong \Lambda_t$. In particular, $\dot{\Lambda}_t$ and $\dot{\Lambda}_t^\circ$ can be treated as the mappings

$$\dot{\Lambda}_t : \Lambda_t \to \Lambda_t^\circ, \quad \dot{\Lambda}_t^\circ : \Lambda_t^\circ \to \Lambda_t.$$

Definition 3. The operator $R_\lambda(t) : \Lambda_t \to \Lambda_t$ defined by the formula

$$R_\Lambda(t) = \dot{\Lambda}_t^\circ \circ \dot{\Lambda}_t$$

is called the curvature operator of the curve $\Lambda_.$ at point t.

The curvature is a kind of "relative velocity" of the curves Λ_t and $\dot{\Lambda}_t$. In the regular case of a non-degenerate $\dot{\Lambda}_t$, the coordinate form of the curvature is as follows:

$$R_\Lambda(t) = \frac{1}{2}\dot{S}_t^{-1}\dddot{S}_t - \frac{3}{4}(\dot{S}_t^{-1}\ddot{S}_t)^2.$$

This is nothing else but the matrix version of the Schwartzian derivative of S_t.

Lecture IV

In this lecture, we consider the curves Λ_t in the Lagrange Grassmannian that is defined for all $t \in \mathbb{R}$.

Theorem 2. *If $\dot{\Lambda}(t) \geq 0$ (non-negative quadratic form) and $\dot{\Lambda}^\circ(t) \leq 0$ (non-positive quadratic form), then no conjugate points occur and*

$$\exists \lim_{t\to\pm\infty} \Lambda(t) = \Lambda(\pm\infty).$$

Remark. The same statement is true also in the case $\dot{\Lambda}(t) \leq 0$, $\dot{\Lambda}^\circ(t) \geq 0$. What is important is that $\dot{\Lambda}(t)$ and $\dot{\Lambda}^\circ(t)$ have opposite signs.

Let us explain this result in the simplest case $n = 1$. Then Σ is the plane equipped with the area form σ, and the Lagrange Grassmanninan is the oriented real projective line that is actually the oriented circle. A monotone curve $\Lambda(t)$ is simply a monotone curve on the circle. A conjugate point occurs when $\Lambda(t)$ makes a complete revolution. On the other hand, $\Lambda(t)$ never coincides with $\Lambda^\circ(t)$, hence there are no conjugate points and $\exists \lim\limits_{t\to\pm\infty} \Lambda(t)$.

If $n > 1$, then the proof remains essentially the same. Indeed, all pairs of transversal Lagrangian subspaces are equivalent by the action of the symplectic group. We may take coordinates in such a way that

$$\Lambda(t) = \{(p, S_t p) : p \in \mathbb{R}^n\}, \quad \Lambda^\circ(t) = \{(p, S_t^\circ p) : p \in \mathbb{R}^n\}$$

for t close to 0 and $S_0 < S_0^\circ$. The relation $\Lambda(t) \cap \Lambda^\circ(t) = 0$ is equivalent to the inequality $\det(S_t - S_t^\circ) \neq 0$. By the monotonicity assumption,

$$\langle p, \dot{S}_t p\rangle \geq 0, \quad \langle p, \dot{S}_t^\circ p\rangle \leq 0, \quad \forall p \in \mathbb{R}^n.$$

Hence the inequality

$$\langle p, \dot{S}_t p\rangle < \langle p, \dot{S}_t^\circ p\rangle, \quad \forall p \in \mathbb{R}^n$$

remains true for all positive t and each of two sides of the last inequality monotonically tends to a limit as $t \to +\infty$. The case of $t \to -\infty$ is handled by the reversion of time.

Let us come back to the Hamiltonian setting. We have $H : T^*M \to \mathbb{R}$, $z \in T^*M$, and $\Lambda_z(t) \subset T_z(T^*M)$ such that

$$\Lambda_z(t) = e_*^{-t\mathbf{H}}\left(T_{e^{t\mathbf{H}}(z)}(T^*_{q(t)}M)\right),$$

where $q(t) = \pi\left(e^{t\mathbf{H}}(z)\right)$. The limits $\Lambda_z(\pm\infty)$ are invariant for the flow $e^{t\mathbf{H}}$. Indeed,

$$e_*^{s\mathbf{H}}\Lambda_z(t) = e_*^{s\mathbf{H}}\Lambda_z(t) \lim_{t\to\pm\infty} e_*^{-t\mathbf{H}}\left(T_{e^{t\mathbf{H}}(z)}(T^*_{q(t)}M\right) =$$

$$\lim_{t\to\pm\infty} e_*^{(s-t)\mathbf{H}}\left(T_{e^{(t-s)\mathbf{H}}(z)}(T^*_{q(t-s)}M\right) = \Lambda_z(\pm\infty).$$

So we have a pair of invariant Lagrangian distributions $\Lambda_z(\pm\infty)$, $z \in T^*M$. Recall that the curvature $R_\Lambda(t) = \dot{\Lambda}_t^\circ \circ \dot{\Lambda}_t$ is in certain sense the relative velocity of $\Lambda(t)$ and $\Lambda^\circ(t)$. If we assume the strong monotonicity: $\dot{\Lambda}(t) > 0$ or $\dot{\Lambda}(t) < 0$, then $|\dot{\Lambda}(t)|$ defines a Euclidean structure on $\Lambda(t)$ and $R_\Lambda(t)$ is a self-adjoint operator for this Euclidean structure. In particular, the operator $R_\Lambda(t)$ is diagonalizable and all its eigenvalues are real. We say that the curvature is positive or negative e. t. s. if all eigenvalues are like that.

Set $\nu = \mathrm{sign}(\dot{\Lambda}(t))$. The matrix of the operator $R_\Lambda(t)$ is equal to the matrix of the quadratic form $\nu\dot{\Lambda}(t)$ in the coordinates where the Euclidean structure $|\dot{\Lambda}(t)|$ is presented by the unit matrix. In particular, assumptions of Theorem 2 are satisfied if the quadratic form $\dot{\Lambda}(t)$ is sign-definite and $R_\Lambda(t) \leq 0$.

It is important that the construction of the curvature is intrinsic and thus survives symplectic transformations. In other words, if A is a linear symplectic transformation of Σ and $A\Lambda : t \mapsto A(\Lambda(t))$, then the operator $R_{A\Lambda}(t)$ is similar to the operator $R_\Lambda(t)$ and has the same eigenvalues.

Definition 4. Let $\Lambda_z\Lambda(\cdot)$, $z \in T*M$, be the Jacobi curves of the Hamiltonian field $\mathbf{H}$; then $R_{\lambda_z}(0) \stackrel{def}{=} R_z^H$ is called the curvature operator of $\mathbf{H}$ at z.

Recall that $\dot{\Lambda}_z(0) = -\frac{\partial^2 H}{\partial p^2}(z)$, $z = (p,q)$. If H is strongly convex with respect to p, then $\dot{\Lambda}_z(0) < 0$, $\forall z$. It was proved in Lecture III that the inequality $\dot{\Lambda}_z(0) \leq 0$, $\forall z$ implies $\dot{\Lambda}_z(t) \leq 0$, $\forall t$. We can repeat that proof and see that the result remains valid if one substitute the non-strong inequality by the strong one. In fact, we can see much more if we analyze the proof. Indeed, the germ at t of the curve $\Lambda_z(\cdot)$ is the image of the germ at 0 of the curve $\Lambda_{e^{t\mathbf{H}}(z)}(\cdot)$ under the fixed symplectic transformation $e_*^{-t\mathbf{H}}$. Hence all invariant quantities of these germs are equal. In particular, the operator $R_{\Lambda_z}(t)$ has the same eigenvalues as the curvature operator of $\mathbf{H}$ at the point $e^{t\mathbf{H}}(z)$. This fact gives a great advantage because the curvature of $\mathbf{H}$ is just a (rather complicated but quite explicit) differential operator of H, we do not need to solve differential equations in order to compute this curvature.

Combining all together, we obtain

Theorem 3. *If the restrictions of H to the fibers T_q^*, $q \in M$, are strongly convex functions and $R_z^H \leq 0$, $\forall z \in M$, then the trajectories of the flow $e^{t\mathbf{H}}$ do not have conjugate points and, moreover, this flow possesses two invariant Lagrangian distributions $\Lambda_z(\pm\infty)$, $z \in T^*M$.*

Let us discuss what happens when $R_\Lambda \geq 0$. First of all, let us see how changes the curvature if we re-parameterize the curve $\Lambda(t)$.

Remark. Clearly, the presence of conjugate point does not depend on the parameterization.

Let $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be a change of parameter; we assume that $\dot\varphi(t) \neq 0$. Denote

$$R_\varphi(t) = \frac{1}{2}\dot\varphi^{-1}(t)\,\dddot\varphi\,(t)(t) - \frac{3}{4}\left(\dot\varphi^{-1}\ddot\varphi(t)\right)^2,$$

the Schwartzian derivative of $\varphi(t)$. Let

$$\Lambda_\varphi : t \to \Lambda(\varphi(t))$$

be the re-parameterized curve.

Proposition 2 $R_{\Lambda_\varphi}(t) = \dot\varphi(t)^2 R_\Lambda(\varphi(t)) + R_\varphi(t)Id.$

This formula (the chain rule) can be checked by the direct calculation which we omit here.

Example. Take $\varphi(t) = \frac{1}{\sqrt{c}}\arctan(\sqrt{c}t)$. Then:

$$R_\varphi(t) = \frac{-c}{(ct^2+1)^2}, \quad \dot\varphi(t) = \frac{1}{ct^2+1},$$

$$R_{\Lambda_\varphi}(t) = \frac{1}{(ct^2+1)^2}\left(R_\Lambda(\varphi(t)) - cId\right).$$

Theorem 4 (Comparison theorem). *Let $t \mapsto \Lambda(t)$ be a smooth strongly monotone curve in the Lagrange Grassmannian and c be a nonnegative constant.*

1. *If $R_\Lambda(t) \le cId$, $\forall t$, then any pair of conjugate points t_1 and t_2 satisfies the inequality $|t_1 - t_2| \ge \frac{\pi}{\sqrt{c}}$.*
2. *If $\frac{1}{n}\mathrm{trace}R_\Lambda(t) \ge c$, then any segment $\left[t, t+\frac{\pi}{\sqrt{c}}\right]$ contains a conjugate to 0 point.*

Remark. If $c \to 0$ in the statement 1., then $|t_1 - t_2| \to \infty$ which correlates with our previous result.

Proof. Statement 1. Re-parameterization and the chain rule reduces everything to the case of non-positive curvature (see above Example).

Statement 2. Assume that $\Lambda(t)$ is transversal to some fixed Lagrangian subspace (for instance, to $\Lambda(0)$) for any $t \in [t_1, t_2]$. Then the segment $\Lambda(t),\ t \in [t_1, t_2]$ of the curve is contained in the fixed coordinate chart of the Lagrange Grassmannian and can be presented in the matrix form:

$$\Lambda(t) = \{(p, S_t p) : p \in \mathbb{R}^n\}, \quad t \in [t_1, t_2].$$

Hence

$$R_\Lambda(t) = \frac{1}{2}\dot S_t^{-1}\,\dddot S_t - \frac{3}{4}(\dot S_t^{-1}\ddot S_t)^2.$$

Now set $W_t = \frac{1}{2}\dot{S}_t^{-1}\ddot{S}_t$; we have, $R_\Lambda(t) = \dot{W}(t) - W(t)^2$. What remain is to find the low bound for the blow-up time of the solutions to the differential inequality $\mathrm{trace}\dot{W} \geq \mathrm{trace}(W^2) + nc$. This is an easy task due to the fact that $(\mathrm{trace}W)^2 \leq n\,\mathrm{trace}(W^2)$.

In order to conclude we recall that in the Riemannian case the Hamiltonian flow is the geodesic flow. Actually, the Riemannian structure identifies TM and T^*M. So, in this case $\frac{\partial^2 H}{\partial p^2}$ is the Riemannian metric. The curvature R_z^H, where $z \in T_q^*$, is essentially the sectional curvature at q in the two-dimensional directions which include $z \in T_q^*M \cong T_qM$.

References

1. Agrachev, A.A.: Geometry of optimal control problems and Hamiltonian systems. In: C.I.M.E. Lecture Notes in Mathematics, Springer Verlag, Berlin (to appear)
2. Agrachev, A.A., Sachkov, Yu.L.: Control Theory from the Geometric Viewpoint. Springer Verlag, Berlin (2004)
3. Gamkrelidze, R.V.: Principles of Optimal Control Theory. Plenum Publishing Corporation, New York (1978)
4. Jurdjevic, V.: Geometric Control Theory. Cambridge University Press (1997)
5. Pontryagin, L.S., Boltyanskij, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: The Mathematical Theory of Optimal Processes. Pergamon Press, Oxfrod (1964)